Maths Methods Week 1: Vector Spaces

http://www-thphys.physics.ox.ac.uk/people/JohnMagorrian/mm
john.magorrian@physics.ox.ac.uk

This part is largely a reminder of the material in last year's Vectors & Matrices course. I won't cover all of it in my lectures, but will focus on the parts that are most relevant to complex vector spaces, namely the defintion of inner product and dual space in §2 and the definition of the adjoint operator §3.3. You should make sure you're comfortable with *everything* here though.

1 Linear vector spaces

A **linear vector space** (or just **vector space** for short) consists of

- a set $\mathcal{V}$ of vectors (the elements of which we'll usually denote by $\mathbf{a}$, $\mathbf{b}$, ..., $\vec{a}$, $\vec{b}$, ... or $|a\rangle$, $|b\rangle$, ...);
- a field $\mathcal{F}$ of scalars (scalars denoted by α, β, a, b,...),
- a rule for adding two vectors to produce another vector,
- a rule for multiplying vectors by scalars,

that together satisfy 10 conditions. The four most interesting conditions are the following.

(i) The set $\mathcal{V}$ of vectors is closed under addition, i.e.,

$$\mathbf{a}+\mathbf{b}\in\mathcal{V}\quad\text{for all }\mathbf{a},\,\mathbf{b}\in\mathcal{V};\tag{1.1}$$

(ii) $\mathcal{V}$ is also closed under multiplication by scalars, i.e.,

$$\alpha\mathbf{a}\in\mathcal{V}\quad\text{for all }\mathbf{a}\in\mathcal{V}\text{ and }\alpha\in\mathcal{F}.\tag{1.2}$$

(iii) $\mathcal{V}$ contains a special zero vector, $\mathbf{0}\in\mathcal{V}$, for which

$$\mathbf{a}+\mathbf{0}=\mathbf{a}\quad\text{for all }\mathbf{a}\in\mathcal{V};\tag{1.3}$$

(iv) Every vector has an additive inverse: for all $\mathbf{a}\in\mathcal{V}$ there is some $\mathbf{a}'\in\mathcal{V}$ for which

$$\mathbf{a}+\mathbf{a}'=\mathbf{0}.\tag{1.4}$$

The other six conditions are more technical. The addition operator must be commutative and associative:

$$\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a},\tag{1.5}$$
$$(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c}).\tag{1.6}$$

The multiplication-by-scalar operation must be distributive with respect to vector and scalar addition, consistent with the operation of multiplying two scalars and must satisfy the multiplicative identity:

$$\alpha(\mathbf{a}+\mathbf{b})=\alpha\mathbf{a}+\alpha\mathbf{b}\tag{1.7}$$
$$(\alpha+\beta)\mathbf{a}=\alpha\mathbf{a}+\beta\mathbf{a}\tag{1.8}$$
$$\alpha(\beta\mathbf{a})=(\alpha\beta)\mathbf{a}\tag{1.9}$$
$$1\mathbf{a}=\mathbf{a}.\tag{1.10}$$

For our purposes the scalars $\mathcal{F}$ will usually be either the set $\mathbb{R}$ of all real numbers (in which case we have a **real vector space**) or the set $\mathbb{C}$ of all complex numbers (giving a **complex vector space**).

1.1 Basic ideas

In a "raw" vector space there is no notion of the length of a vector or the angle between two vectors. Nevertheless, there are many important ideas that follow by applying the basic rules (1.1 – 1.10) above to **linear combinations** of vectors, i.e., weighted sums such as

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots. \tag{1.11}$$

A set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is said to be **linearly independent** (abbreviated **LI**) if the only solution to the equation

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots \alpha_n \mathbf{v}_n = 0 \tag{1.12}$$

is if all scalar coefficients $\alpha_i = 0$. Otherwise the set is linearly dependent. The **dimension** of a vector space is the maximum number of LI vectors in the space.

The **span** of a list of vectors $\mathbf{v}_1, \dots, \mathbf{v}_m$ is the set of all possible linear combinations $\{\alpha_1 \mathbf{v}_1 + \cdots + \alpha_m \mathbf{v}_m : \alpha_1, \dots, \alpha_m \in \mathcal{F}\}$. A list $\mathbf{e}_1, \mathbf{e}_2, \dots \mathbf{e}_n$ of vectors forms a **basis** for the space $\mathcal{V}$ if the elements of the list are LI and span $\mathcal{V}$. Then any $\mathbf{a} \in \mathcal{V}$ can be expressed as

$$\mathbf{a} = \sum_{i=1}^{n} a_i \mathbf{e}_i, \tag{1.13}$$

and the coefficients $(a_1, \dots, a_n)$ for which (1.13) holds are known as the **components** or **coordinates** of **a** with respect to the basis vectors $\mathbf{e}_i$.

> **Claim:** Given a basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ the coordinates a_i of **a** are unique.
> **Proof:** suppose that there is another set of coordinates a'_i. Then we can express **a** in two ways:
>
> $$\begin{aligned} \mathbf{a} &= a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + \cdots + a_n \mathbf{e}_n \\ &= a'_1 \mathbf{e}_1 + a'_2 \mathbf{e}_2 + \cdots + a'_n \mathbf{e}_n. \end{aligned} \tag{1.14}$$
>
> Subtracting,
>
> $$0 = (a_1 - a'_1)\mathbf{e}_1 + (a_2 - a'_2)\mathbf{e}_2 + \cdots + (a_n - a'_n)\mathbf{e}_n. \tag{1.15}$$
>
> But we are told that the $\{\mathbf{e}_i\}$ are a basis. By definition then they must be LI. Therefore the only way of satisfying the equation above is if all $a_i - a'_i = 0$. So $a'_i = a_i$: the coordinates are unique.

A subset $\mathcal{W} \subseteq \mathcal{V}$ is a **subspace** of $\mathcal{V}$ if it satisfies conditions (1.1–1.4). That is: it must be closed under addition of vectors and multiplication by scalars; it must contain the zero vector; the additive inverse of each element must be included. Conditions (1.5–1.10) are automatically satisfied because they depend only on the definition of the addition and multiplication operations.

Let $\mathcal{U}$ and $\mathcal{W}$ be subspaces of $\mathcal{V}$. Their **sum**, written $\mathcal{U} + \mathcal{W}$, is the subspace of $\mathcal{V}$ consisting of all sums $\mathbf{u} + \mathbf{w}$ for $\mathbf{u} \in \mathcal{U}$ and $\mathbf{w} \in \mathcal{W}$. If every element of $\mathbf{v} \in \mathcal{V}$ can be written as $\mathbf{v} = \mathbf{u} + \mathbf{w}$ with unique elements $\mathbf{u} \in \mathcal{U}$, $\mathbf{w} \in \mathcal{W}$ for each such $\mathbf{v}$, then $\mathcal{V}$ is the **direct sum** of $\mathcal{U}$ and $\mathcal{W}$, written $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$.

> **Exercise:** (i) If $\mathcal{V} = \mathcal{U} + \mathcal{W}$ and $\mathcal{U} \cap \mathcal{W} = \{0\}$, show that $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$. (ii) If $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$ show that $\dim \mathcal{V} = \dim \mathcal{U} + \dim \mathcal{W}$.

> **Exercise:** Let $\mathcal{U}$ and $\mathcal{W}$ be any two vector spaces over the same field of scalars $\mathcal{F}$ (i.e., they are not necessarily subspaces of some "larger" space $\mathcal{V}$). Their **direct product**[†] $\mathcal{U} \times \mathcal{W}$ is the set of all pairs $(\mathbf{u}, \mathbf{w})$ for $\mathbf{u} \in \mathcal{U}$, $\mathbf{w} \in \mathcal{W}$. Defining addition of such pairs as
>
> $$(\mathbf{u}_1, \mathbf{w}_1) + (\mathbf{u}_2, \mathbf{w}_2) = (\mathbf{u}_1 + \mathbf{u}_2, \mathbf{w}_1 + \mathbf{w}_2) \tag{1.16}$$
>
> and multiplication by scalars $\alpha \in \mathcal{F}$ as
>
> $$\alpha(\mathbf{u}, \mathbf{w}) = (\alpha \mathbf{u}, \alpha \mathbf{w}), \tag{1.17}$$
>
> show that $\mathcal{U} \times \mathcal{W}$ is a vector space and that $\dim(\mathcal{U} \times \mathcal{W}) = \dim \mathcal{U} + \dim \mathcal{W}$.

[†] Note that, whereas the direct sum is written $\mathcal{U} \oplus \mathcal{W}$, the direct product is denoted $\mathcal{U} \times \mathcal{W}$, *not* $\mathcal{U} \otimes \mathcal{W}$. For the meaning of the latter see §A.3 below.

1.2 Examples

Example: Three-dimensional column vectors with real coefficients The set of column vectors $(x_1, x_2, x_3)^{\mathrm{T}}$ with $x_i \in \mathbb{R}$ forms a real vector space under the usual rules of vector addition and multiplication by scalars. This space is usually known as $\mathbb{R}^3$.

To confirm that this really is a vector space, let's check the conditions (1.1–1.10). The usual rules of vector algebra satisfy conditions (1.5–1.10). For the conditions (1.1–1.4) note that:

(i) For any $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \in \mathbb{R}^3$, the sum $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \equiv \begin{pmatrix} a_1+b_1 \\ a_2+b_2 \\ a_3+b_3 \end{pmatrix} \in \mathbb{R}^3$.

(ii) Multiplying any vector $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ by a real scalar α gives $\alpha \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \equiv \begin{pmatrix} \alpha a_1 \\ \alpha a_2 \\ \alpha a_3 \end{pmatrix} \in \mathbb{R}^3$.

(iii) There is a zero element, $(0,0,0)^T \in \mathbb{R}^3$.

(iv) Each vector $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ has an additive inverse $\begin{pmatrix} -a_1 \\ -a_2 \\ -a_3 \end{pmatrix} \in \mathbb{R}^3$.

So, all conditions (1.1–1.10) are satisfied. Here are two possible bases for this space:

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \quad \text{or} \quad \left\{ \begin{pmatrix} 1 \\ \pi \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 6 \end{pmatrix} \right\}. \tag{1.18}$$

Each of these basis sets has three LI elements that span $\mathbb{R}^3$. Therefore the dimension of $\mathbb{R}^3$ is 3.

> **Exercise:** The set of all 3-dimensional column vectors with real coefficients cannot form a complex vector space. Why not? (Which of the conditions 1.1–1.10 is broken?)

Example: $\mathbb{R}^n$ and $\mathbb{C}^n$ Similarly, the set of all n-dimensional column vectors with real (complex) elements forms a real (complex) vector space under the usual rules of vector addition and multiplication by scalars.

Example: Arrows on a plane The set of all arrows on a plane with the obvious definitions of addition of arrows and multiplication (stretching/shrinking) of arrows by scalars forms a real two-dimensional vector space.

Example: The set of all $m \times n$ matrices with complex coefficients forms a complex vector space with dimension mn. The most natural basis is

$$\left\{ \begin{pmatrix} 1 & 0 & \cdots \\ 0 & 0 & \cdots \\ \vdots & \vdots & \ldots \end{pmatrix}, \begin{pmatrix} 0 & 1 & \cdots \\ 0 & 0 & \cdots \\ \vdots & \vdots & \ldots \end{pmatrix}, \ldots, \begin{pmatrix} 0 & 0 & \cdots \\ 1 & 0 & \cdots \\ \vdots & \vdots & \ldots \end{pmatrix}, \begin{pmatrix} 0 & 0 & \cdots \\ 0 & 1 & \cdots \\ \vdots & \vdots & \ldots \end{pmatrix}, \ldots \right\}. \tag{1.19}$$

Example: n^{th}-order polynomials The set of all n^{th}-order polynomials in a complex variable z forms an $n+1$ dimensional complex vector space. A natural basis is the set of monomials $\{1, z, z^2, \ldots, z^n\}$.

Example: Trigonometric polynomials Given n distinct (mod 2π) complex constants $\lambda_1, \ldots, \lambda_n$, the set of all linear combinations of $\mathrm{e}^{\mathrm{i}\lambda_n z}$ forms an n-dimensional complex vector space.

Example: Functions The set $L_2(a,b)$ of all complex-valued functions

$$f : [a,b] \to \mathbb{C} \tag{1.20}$$

for which the integral

$$\int_a^b \mathrm{d}x |f(x)|^2 \tag{1.21}$$

exists forms a complex vector space under the usual operations of addition of functions and multiplication of functions by scalars. This space has an infinite number of dimensions. We postpone the issue of identifying a suitable basis until §4 later.

> **Exercise:** Consider the set of coupled linear, homogeneous differential equations
>
> $$\dot{\mathbf{x}} = A(t)\mathbf{x}, \tag{1.22}$$
>
> where $\mathbf{x}$ is an n-dimensional vector and $A(t)$ is an $n \times n$ matrix whose coefficients might depend on t. Show that the set of solutions to this equation is a vector space. What is its dimension?

1.3 Linear maps

A mapping $A : \mathcal{V} \to \mathcal{W}$ from one vector space $\mathcal{V}$ to another $\mathcal{W}$ is a **linear map** if it satisfies

$$\begin{aligned} A\big(\mathbf{v}_1 + \mathbf{v}_2\big) &= A\mathbf{v}_1 + A\mathbf{v}_2, \\ A\big(\alpha\mathbf{v}_1\big) &= \alpha A\mathbf{v}_1, \end{aligned} \tag{1.23}$$

for all $\mathbf{v}_1$, $\mathbf{v}_2 \in \mathcal{V}$ and scalars $\alpha \in \mathcal{F}$. Notice that for this definition to make sense, both vector spaces must be defined over the same type of scalars $\mathcal{F}$. A **linear operator** is the special case of a linear map of a vector space $\mathcal{V}$ to itself.

Recall that the **image** (some times referred to as the **range**) of A is the set of all possible output vectors,

$$\operatorname{Im} A = \{A\mathbf{v} \,:\, \mathbf{v} \in \mathcal{V}\}, \tag{1.24}$$

while the **kernel** (or **nullspace**) of A is the set of all inputs that give zero output:

$$\ker A = \{\mathbf{v} \,:\, A\mathbf{v} = \mathbf{0},\, \mathbf{v} \in \mathcal{V}\}. \tag{1.25}$$

Both image and kernel are themselves vector spaces: the image is a subspace of $\mathcal{W}$, while the kernel is a subspace of $\mathcal{V}$. Their dimensions are related by the **dimension theorem**:

$$\dim \mathcal{V} = \dim \operatorname{Im} A + \dim \ker A. \tag{1.26}$$

The first term, $\dim \operatorname{Im} A$ is also known as the **rank** of the map A.

An **isomorphism** between $\mathcal{V}$ and $\mathcal{W}$ is a linear map $\phi : \mathcal{V} \to \mathcal{W}$ that is invertible (i.e., has zero kernel). By virtue of ϕ each element of $\mathcal{V}$ is paired up up with precisely one element of $\mathcal{W}$ and vice versa, the pairing satisying the linearity condition (1.23), so that $\phi(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) = \alpha_1\phi(\mathbf{v}_1) + \alpha_2\phi(\mathbf{v}_2)$. The spaces $\mathcal{V}$ and $\mathcal{W}$ are **isomorphic** if we can construct an isomorphism between them.

> **Exercise:** Suppose that $\mathcal{V}$ and $\mathcal{W}$ are two vector spaces over scalars $\mathcal{F}$. Explain why $\mathcal{V}$ and $\mathcal{W}$ cannot be isomorphic unless they have the same dimension. Show that there is an infinite number of isomorphisms between them if they do have the same dimension.

1.4 Representation of vectors and linear maps by matrices

Any n-dimensional vector space $\mathcal{V}$ over scalars $\mathcal{F}$ is isomorphic to $\mathcal{F}^n$, the set of n-dimensional column vectors whose elements are drawn from $\mathcal{F}$. To see this, choose any basis $\mathbf{e}_1$,....,$\mathbf{e}_n$ for $\mathcal{V}$ and construct the isomorphism ϕ by identifying

$$\mathbf{e}_1 \text{ with } \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 \text{ with } \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad \mathbf{e}_n \text{ with } \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \tag{1.27}$$

Then any vector $\mathbf{v}$ can be expressed as

$$\mathbf{v} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \cdots + a_n\mathbf{e}_n, \text{ which corresponds to } \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \tag{1.28}$$

and linear combinations of such vectors are paired up with corresponding linear combinations of the corresponding column vectors. So, your deepest thoughts about n-dimensional column vectors carry over directly to any n-dimensional vector space.

To spell out the connection between linear maps $A : \mathcal{V} \to \mathcal{W}$ and matrices, let n be the dimension of $\mathcal{V}$ and m the dimension of $\mathcal{W}$. Choose any basis $\mathbf{e}_1^{\mathcal{V}}, \ldots, \mathbf{e}_n^{\mathcal{V}}$ for $\mathcal{V}$ and another, $\mathbf{e}_1^{\mathcal{W}}, \ldots, \mathbf{e}_m^{\mathcal{W}}$, for $\mathcal{W}$. Any vector $\mathbf{v} \in \mathcal{V}$ can be expressed as $\mathbf{v} = \sum_{i=1}^n a_j \mathbf{e}_j^{\mathcal{V}}$. Using the properties (1.23) we have that the image of $\mathbf{v}$ under the linear map A is

$$A\mathbf{v} = \sum_{j=1}^n a_j \left(A\mathbf{e}_j^{\mathcal{V}}\right). \tag{1.29}$$

As this holds for any $\mathbf{v} \in \mathcal{V}$, we see that the map A is completely determined by the images $A\mathbf{e}_1^{\mathcal{V}}$, ..., $A\mathbf{e}_n^{\mathcal{V}}$ of $\mathcal{V}$'s basis vectors. Each of these images $A\mathbf{e}_i^{\mathcal{V}}$ is a vector that lives in $\mathcal{W}$ and so can be expressed in terms of the basis $\mathbf{e}_1^{\mathcal{W}}, \ldots, \mathbf{e}_m^{\mathcal{W}}$ as

$$A\mathbf{e}_j^{\mathcal{V}} = \sum_{i=1}^m A_{ij}\mathbf{e}_i^{\mathcal{W}}, \tag{1.30}$$

where the coefficient A_{ij} is the i^{th} component in the $\mathbf{e}_1^{\mathcal{W}}, \ldots, \mathbf{e}_m^{\mathcal{W}}$ basis of the vector $A\mathbf{e}_j^{\mathcal{V}}$. Substituting this into (1.29),

$$A\mathbf{v} = \sum_{i=1}^m \left[\sum_{j=1}^n A_{ij}a_j\right] \mathbf{e}_i^{\mathcal{W}}. \tag{1.31}$$

That is, a vector in $\mathcal{V}$ with components $a_1, \ldots, a_n$ maps under A to another vector in $\mathcal{W}$ whose i^{th} component is given by $\sum_{j=1}^n A_{ij}a_j$. The values of the coefficients A_{ij} depend on the bases chosen for $\mathcal{V}$ and $\mathcal{W}$.

Let us again identify $\mathbf{e}_1^{\mathcal{V}}, ..., \mathbf{e}_n^{\mathcal{V}}$ with the natural n-dimensional column vector basis (1.27) and similarly for $\mathbf{e}_1^{\mathcal{W}}, ..., \mathbf{e}_m^{\mathcal{W}}$. Then the map A corresponds to the matrix

$$A = \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1n} \\ A_{21} & A_{22} & \ldots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{pmatrix}, \tag{1.32}$$

and the vector $A\mathbf{v} \in \mathcal{W}$ to

$$\begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1n} \\ A_{21} & A_{22} & \ldots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} A_{11}a_1 + A_{12}a_2 + \cdots + A_{1n}a_n \\ A_{21}a_1 + A_{22}a_2 + \cdots + A_{2n}a_n \\ \vdots \\ A_{m1}a_1 + A_{m2}a_2 + \cdots + A_{mn}a_n \end{pmatrix} \tag{1.33}$$

in accordance with the familiar rules of matrix multiplication.

Linear operators (that is, linear maps of a vector space to itself) are represented by square $n \times n$ matrices.

Further reading

Linear vector spaces are introduced in RHB §8.1 and linear maps in RHB §8.2. DK §II is another good starting point. You've already had a thorough, clear introduction to vector spaces in last year's "Vectors and Matrices" course. There you learned to distinguish between linear operators $f : \mathcal{V} \to \mathcal{V}$ and the matrices that represent those operators, $A = \varphi^{-1} \circ f \circ \varphi$, where $\varphi(\mathbf{v}) = \sum_i v_i \mathbf{e}_i$ is an invertible "coordinate map" that translates column vectors of coordinates into the vectors that the map f understands. I'm not as careful in these lectures.

Most maths-for-physicists books introduce inner products (see §2 below) at the same time as vector spaces. Nevertheless, pausing to work out the consequences of the unadorned conditions (1.1–1.10) is an supremely useful introduction to mathematical reasoning: many of the statements that we take as self-evident from our experience in manipulating vectors and matrices are not easy to prove without some practice. For more on this, see the first-year "Vectors and Matrices" course, or, e.g., *Linear Algebra* by Lang or similar books for mathematicians.

2 Inner-product spaces

The conditions (1.1–1.10) do not allow us to say whether two vectors are orthogonal, or even what the length of a vector is. To do these, we need to introduce some additional structure on the space, namely the idea of an inner product. This is a straightforward generalization of the familiar scalar product.

An **inner product** is a mapping $\mathcal{V} \times \mathcal{V} \to \mathcal{F}$ that takes two vectors and returns a scalar and satisfies the following conditions for all $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c} \in \mathcal{V}$ and $\alpha \in \mathcal{F}$:

$$\langle \mathbf{c}, \mathbf{a} + \mathbf{b} \rangle = \langle \mathbf{c}, \mathbf{a} \rangle + \langle \mathbf{c}, \mathbf{b} \rangle; \tag{2.1}$$

$$\langle \mathbf{c}, \alpha \mathbf{a} \rangle = \alpha \langle \mathbf{c}, \mathbf{a} \rangle; \tag{2.2}$$

$$\langle \mathbf{a}, \mathbf{b} \rangle = \langle \mathbf{b}, \mathbf{a} \rangle^\star; \tag{2.3}$$

$$\begin{aligned} \langle \mathbf{a}, \mathbf{a} \rangle &= 0, \quad \text{only if } \mathbf{a} = 0, \\ &> 0, \quad \text{otherwise.} \end{aligned} \tag{2.4}$$

Notice that the inner product is linear in the *second* argument, but not necessarily in the first.

An inner-product space is simply a vector space $\mathcal{V}$ on which an inner product $\langle \mathbf{a}, \mathbf{b} \rangle$ has been defined.

Some definitions:

- The inner product of a vector with itself, $\langle \mathbf{a}, \mathbf{a} \rangle$, is real and non-negative. The **length** or **norm** of the vector $\mathbf{a}$ is $|\mathbf{a}| \equiv \sqrt{\langle \mathbf{a}, \mathbf{a} \rangle}$.
- The vectors $\mathbf{a}$ and $\mathbf{b}$ are **orthogonal** if $\langle \mathbf{a}, \mathbf{b} \rangle = 0$.
- A set of vectors $\{\mathbf{v}_i\}$ of $\mathcal{V}$ is **orthonormal** if $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{ij}$.

The condition (2.3) is essential if we want lengths of vectors to be real numbers, but a consequence is that in general the inner product is not linear in both arguments.

Exercise: Use the properties (2.1–2.4) above to show that

$$\langle \alpha \mathbf{a} + \beta \mathbf{b}, \mathbf{c} \rangle = \alpha^\star \langle \mathbf{a}, \mathbf{c} \rangle + \beta^\star \langle \mathbf{b}, \mathbf{c} \rangle. \tag{2.5}$$

Some books use the term “sesquilinear” to describe this property. Under what conditions is the scalar product linear in both arguments?

Exercise: Show that if $\langle \mathbf{a}, \mathbf{v} \rangle = 0$ for all $\mathbf{v} \in \mathcal{V}$ then $\mathbf{a} = 0$.

2.1 Orthonormal bases

An **orthonormal basis** for $\mathcal{V}$ is a set of basis vectors $\mathbf{e}_1, ..., \mathbf{e}_n$ that satisfy

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij}. \tag{2.6}$$

Exercise: Show that any n orthonormal vectors in an n-dimensional inner-product space form a basis. The converse is not true.

Every n-dimensional vector space $\mathcal{V}$ has an orthonormal basis: given any list of n LI vectors $\mathbf{v}_1, ..., \mathbf{v}_n \in \mathcal{V}$ we can construct an orthonormal basis using the following **Gram–Schmidt** procedure.

(1) Start with the first vector from the list, $\mathbf{v}_1$. The first basis vector $\mathbf{e}_1$ is defined via

$$\begin{aligned} \mathbf{e}'_1 &= \mathbf{v}_1, \\ \mathbf{e}_1 &= \mathbf{e}'_1 / |\mathbf{e}'_1|. \end{aligned} \tag{2.7}$$

(2) Take the next vector $\mathbf{v}_2$. Subtract any component that is parallel to the previously constructed basis vector $\mathbf{e}_1$. Normalise the result to get $\mathbf{e}_2$.

$$\begin{aligned}\mathbf{e}_2' &= \mathbf{v}_2 - \langle \mathbf{e}_1, \mathbf{v}_2\rangle \mathbf{e}_1,\\ \mathbf{e}_2 &= \mathbf{e}_2'/|\mathbf{e}_2'|.\end{aligned} \tag{2.8}$$

(i) Similarly, work along the remaining $\mathbf{v}_i$, $i = 3, \ldots, n$, subtracting from each one any component that is parallel to any of the previously constructed basis vectors $\mathbf{e}_1, \ldots, \mathbf{e}_{i-1}$. That is,

$$\begin{aligned}\mathbf{e}_i' &= \mathbf{v}_i - \sum_{j=1}^{i-1} \langle \mathbf{e}_j, \mathbf{v}_i\rangle \mathbf{e}_j,\\ \mathbf{e}_i &= \mathbf{e}_i'/|\mathbf{e}_i'|.\end{aligned} \tag{2.9}$$

It is easy to see that taking the inner product of (2.9) with any $\mathbf{e}_k$, $k < i$, yields the equation $\langle \mathbf{e}_k, \mathbf{e}_i\rangle = 0$: by construction each new $\mathbf{e}_i$ is orthogonal to all the preceding ones.

The same procedure can be used to construct an orthonormal basis for the space spanned by a list of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ of any length, including cases where the list is not LI: if $\mathbf{v}_i$ is linearly dependent on the preceding $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}$ then $\mathbf{e}_i' = 0$ and so that particular $\mathbf{v}_i$ does not produce a new basis vector.

2.2 Representing the inner product by matrices

Consider an n-dimensional inner-product space $\mathcal{V}$ for which the vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are an orthonormal basis. Let

$$\mathbf{a} = \sum_{i=1}^{n} a_i \mathbf{e}_i, \quad \text{and} \quad \mathbf{b} = \sum_{i=1}^{n} b_i \mathbf{e}_i \tag{2.10}$$

be any two vectors in $\mathcal{V}$. Using properties (2.1) and (2.2) of the inner product together with the orthonormality of the basis vectors, we have that the projection of $\mathbf{a}$ onto the j^{th} basis vector

$$\langle \mathbf{e}_j, \mathbf{a}\rangle = \sum_{i=1}^{n} a_i \langle \mathbf{e}_j, \mathbf{e}_i\rangle = \sum_{i=1}^{n} a_i \delta_{ji} = a_j \tag{2.11}$$

and similarly $\langle \mathbf{e}_j, \mathbf{b}\rangle = b_j$. Therefore the inner product of $\mathbf{a}$ and $\mathbf{b}$ is

$$\langle \mathbf{a}, \mathbf{b}\rangle = \sum_{i=1}^{n} b_i \langle \mathbf{a}, \mathbf{e}_i\rangle = \sum_{i=1}^{n} b_i \langle \mathbf{e}_i, \mathbf{a}\rangle^\star = \sum_{i=1}^{n} a_i^\star b_i. \tag{2.12}$$

Note that the i^{th} element of $\mathbf{a}$ is given by $\langle \mathbf{e}_i, \mathbf{a}\rangle$, not $\langle \mathbf{a}, \mathbf{e}_i\rangle$.

As in §1.4 we may identify each $\mathbf{e}_i$ with the n-dimensional column vector (1.27) that has 1 in its i^{th} row and zeros everywhere else. Then

$$\langle \mathbf{a}, \mathbf{b}\rangle = \begin{pmatrix} a_1^\star & \cdots & a_n^\star \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \mathbf{a}^\dagger \mathbf{b}, \tag{2.13}$$

where $\mathbf{a}^\dagger$ is the **Hermitian conjugate** of the column vector $\mathbf{a}$.

2.3 Dual space

Every vector space $\mathcal{V}$ has a corresponding dual space $\mathcal{V}^\star$. The elements of $\mathcal{V}^\star$ are linear maps between $\mathcal{V}$ and scalars $\mathcal{F}$. An inner product on $\mathcal{V}$ sets up a natural correspondence between elements of $\mathcal{V}^\star$ and elements of $\mathcal{V}$.

Consider the set $\mathcal{V}^\star$ of all linear maps from vectors $\mathbf{v} \in \mathcal{V}$ to scalars $\mathcal{F}$. Applying any $L \in \mathcal{V}^\star$ to an element $\mathbf{v} = b_1\mathbf{e}_1 + \cdots + b_n\mathbf{e}_n$ of $\mathcal{V}$, we have, by the linearity of L, that

$$L(\mathbf{v}) = L\left(\sum_{i=1}^n b_i\mathbf{e}_i\right) = \sum_{i=1}^n b_i L(\mathbf{e}_i). \tag{2.14}$$

So, given a basis $\{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$ for $\mathcal{V}$, any $L \in \mathcal{V}^\star$ is completely defined by the n scalar values $L(\mathbf{e}_1)$,…,$L(\mathbf{e}_n)$.

Comparing (2.14) with (2.13), once $\mathcal{V}$ has an inner product then for each $\mathbf{a} \in \mathcal{V}$ there is a corresponding $L_\mathbf{a} \in \mathcal{V}^\star : \mathcal{V} \to \mathcal{F}$ defined by

$$L_\mathbf{a}(\mathbf{b}) \equiv \langle \mathbf{a}, \mathbf{b}\rangle. \tag{2.15}$$

We can turn $\mathcal{V}^\star$ into a vector space. Define the sum of two maps to be the new map given by

$$L_{\mathbf{a}_1+\mathbf{a}_2}(\mathbf{b}) = L_{\mathbf{a}_1}(\mathbf{b}) + L_{\mathbf{a}_2}(\mathbf{b}), \tag{2.16}$$

and multiplication by scalars through

$$L_{\alpha\mathbf{a}}(\mathbf{b}) = \alpha^\star L_\mathbf{a}(\mathbf{b}). \tag{2.17}$$

Exercise: Verify that the maps $L_\mathbf{a}$ constructed this way satisfy the conditions (1.1–1.10).

Bases for $\mathcal{V}^\star$ Given any orthonormal basis $\mathbf{e}_1$,...,$\mathbf{e}_n$ for $\mathcal{V}$, we can immediately define corresponding dual vectors $L_{\mathbf{e}_1}$,...,$L_{\mathbf{e}_n}$ by

$$L_{\mathbf{e}_i}\mathbf{e}_j = \delta_{ij}. \tag{2.18}$$

These $L_{\mathbf{e}_i}$ are LI because, given any $\mathbf{e}_j$, by virtue of (2.18) the only way of producing $\left(\sum_i \alpha_i L_{\mathbf{e}_i}\right)\mathbf{e}_j = 0$ is when all $\alpha_i = 0$. The $L_{\mathbf{e}_i}$ also span $\mathcal{V}^\star$: any $L \in \mathcal{V}^\star$ applied to an arbitrary $\mathbf{v} = \sum_i \alpha_i\mathbf{e}_i$ yields

$$L\mathbf{v} = L\left(\sum_{i=1}^n \alpha_i\mathbf{e}_i\right) = \sum_{i=1}^n \alpha_i L(\mathbf{e}_i) = \sum_{i=1}^n \alpha_i\beta_i = \sum_{i=1}^n \beta_i L_{\mathbf{e}_i}(\mathbf{v}), \tag{2.19}$$

where we have written the result of applying L to $\mathbf{e}_i$ as $\beta_i \equiv L\mathbf{e}_i$. So, any $L \in \mathcal{V}^\star$ can be expressed as a linear combination of the $L_{\mathbf{e}_i}$ and these $L_{\mathbf{e}_i}$ are LI. Therefore the $L_{\mathbf{e}_1}$, ..., $L_{\mathbf{e}_n}$ defined by (2.18) constitute a basis for $\mathcal{V}^\star$. A vector space $\mathcal{V}$ and its dual $\mathcal{V}^\star$ have the same dimension.

Exercise: We have shown that introducing an inner product on $\mathcal{V}$ defines a natural mapping between $\mathcal{V}^\star$ and $\mathcal{V}$. Show that this mapping is an isomorphism: there are many isomorphisms beween $\mathcal{V}$ and $\mathcal{V}^\star$, but the choice of inner product identifies one as special.

The space $\mathcal{V}^\star$ is the **dual** to the space $\mathcal{V}$. Elements of $\mathcal{V}^\star$ are known as “dual vectors”, “covectors”, “linear forms”, “1-forms”, or “bras”.

2.4 Bra-ket notation

Let $\mathcal{V}$ be an inner product space. In Dirac's bra-ket notation, vectors $\mathbf{v} \in \mathcal{V}$ are denoted by $|v\rangle$, pronounced “ket v”. Inhabitants of the vector space $\mathcal{V}^\star$ are “bras” and are written $\langle a|$, $\langle b|$ etc instead of the $L_\mathbf{a}$, $L_\mathbf{b}$ notation used above. For every ket $|v\rangle$ there is a corresponding bra, written $\langle v|$, and vice versa. The addition and multiplication rules (2.16) and (2.17) above mean that

$$\boxed{\begin{aligned} \text{the ket}\quad & |v\rangle = \alpha\,|a\rangle + \beta\,|b\rangle \\ \text{has dual}\quad & \langle v| = \alpha^\star\langle a| + \beta^\star\langle b|\,. \end{aligned}} \tag{2.20}$$

This offers a convenient alternative way of carrying out calculations that involve the inner product. For example, for any $\mathbf{a} = |a\rangle = \sum_{i=1}^n a_i\,|e_i\rangle$ and $\mathbf{b} = |b\rangle = \sum_{i=1}^n b_i\,|e_i\rangle$, the dual to $|a\rangle$ is $\langle a| = \sum_{i=1}^n a_i^\star\langle e_i|$ and we may define

$$\begin{aligned}\langle a|b\rangle \equiv \langle a|\,|b\rangle &= \left(\sum_{i=1}^n a_i^\star\langle e_i|\right)\left(\sum_{j=1}^n b_j\,|e_j\rangle\right)\\ &= \sum_{i=1}^n\sum_{j=1}^n a_i^\star b_j\langle e_i|\,|e_j\rangle = \sum_{i=1}^n\sum_{j=1}^n a_i^\star b_j\delta_{ij} = \sum_{i=1}^n a_i^\star b_i = \langle\mathbf{a},\mathbf{b}\rangle,\end{aligned} \tag{2.21}$$

in agreement with the expression (2.13) from the previous section. It is easy to confirm that this alternative definition of $\langle a|b\rangle$ as the result of operating on the ket vector $|b\rangle$ by the bra vector $\langle a| \in V^\star$ satisfies the conditions (2.1–2.4) for an inner product.

Connection to matrices Here is a summary of the preceding results. If we have an orthonormal basis in which we represent kets by column vectors (§1.4),

$$\begin{aligned}|v\rangle &= \alpha_1\,|e_1\rangle + \alpha_2\,|e_2\rangle + \cdots + \alpha_n\,|e_n\rangle\\ &= \alpha_1\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} + \alpha_2\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix} + \cdots + \alpha_n\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix} = \begin{pmatrix}\alpha_1\\\alpha_2\\\vdots\\\alpha_n\end{pmatrix},\end{aligned} \tag{2.22}$$

then the bra dual to $|v\rangle$ is represented by the Hermitian conjugate of this column vector:

$$\begin{aligned}\langle v| &=\alpha_1^\star\langle e_1| + \alpha_2^\star\langle e_2| + \cdots + \alpha_n^\star\langle e_n|\\ &=\alpha_1^\star\begin{pmatrix}1 & 0 & \dots & 0\end{pmatrix} + \alpha_2^\star\begin{pmatrix}0 & 1 & \dots & 0\end{pmatrix} + \cdots + \alpha_n^\star\begin{pmatrix}0 & 0 & \dots & 1\end{pmatrix}\\ &=\begin{pmatrix}\alpha_1^\star & \alpha_2^\star & \dots & \alpha_n^\star\end{pmatrix}.\end{aligned} \tag{2.23}$$

The inner product $\langle a|b\rangle$ of the vectors $|a\rangle = (a_1,\dots,a_n)^{\mathrm{T}}$ and $|b\rangle = (b_1,\dots,b_n)^{\mathrm{T}}$ is obtained by premultiplying $|b\rangle$ by the dual vector to $|a\rangle$ under the usual rules of matrix multiplication:

$$\langle a|b\rangle \equiv \langle a|\;\;|b\rangle = \begin{pmatrix}a_1^\star & a_2^\star & \dots & a_n^\star\end{pmatrix}\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix} = \sum_{i=1}^n a_i^\star b_i. \tag{2.24}$$

2.5 Example: application of the Gram–Schmidt procedure using bras and kets

Consider the list $|v_1\rangle = (0,\mathrm{i},\mathrm{i},0)^{\mathrm{T}}$, $|v_2\rangle = (0,2,2,1)^{\mathrm{T}}$, $|v_3\rangle = (1,1,1,1)^{\mathrm{T}}$ and $|v_4\rangle = (2,1,1,0)^{\mathrm{T}}$. We want to construct an orthonormal basis for the space spanned by these vectors.

From $|v_1\rangle$ we immediately have that

$$|e_1\rangle = \frac{1}{\sqrt{2}}(0,\mathrm{i},\mathrm{i},0)^{\mathrm{T}}. \tag{2.25}$$

The corresponding basis bra is the row vector

$$\langle e_1| = \frac{1}{\sqrt{2}}(0,-\mathrm{i},-\mathrm{i},0). \tag{2.26}$$

The inner product $\langle e_1|v_2\rangle = -2\sqrt{2}\mathrm{i}$, so

$$|e_2'\rangle = |v_2\rangle - (-2\sqrt{2}\mathrm{i})\,|e_1\rangle = (0,0,0,1)^{\mathrm{T}} = |e_2\rangle. \tag{2.27}$$

For $|v_3\rangle$ the necessary inner products are $\langle e_1|v_3\rangle = -\sqrt{2}\mathrm{i}$ and $\langle e_2|v_3\rangle = 1$. Then

$$\begin{aligned} |e_3'\rangle &= |v_3\rangle - (-\sqrt{2}\mathrm{i})\,|e_1\rangle - |e_2\rangle \\ &= (1,0,0,0)^{\mathrm{T}} = |e_3\rangle. \end{aligned} \tag{2.28}$$

Finally, notice that $|e_4'\rangle = 0$ because $|v_4\rangle = 2\,|e_3\rangle - \sqrt{2}\mathrm{i}\,|e_1\rangle$. Therefore the four vectors $|v_1\rangle, \ldots, |v_4\rangle$ span a three-dimensional subspace of the original four-dimensional space. The kets $|e_1\rangle$, $|e_2\rangle$ and $|e_3\rangle$ constructed above are one possible orthonormal basis for this subspace.

2.6 Some important relations involving the inner product

Recall that $|\mathbf{a}|^2 \equiv \langle \mathbf{a}, \mathbf{a}\rangle$. In bra-ket notation, $|a|^2 \equiv \langle a|a\rangle$.

Pythagoras if $\langle \mathbf{a}, \mathbf{b}\rangle = 0$ (or, $\langle a|b\rangle = 0$) then

$$\begin{aligned} |\mathbf{a}+\mathbf{b}|^2 &= |\mathbf{a}|^2 + |\mathbf{b}|^2\,, \\ \big|\,|a\rangle + |b\rangle\big|^2 &= \big|\,|a\rangle\big|^2 + \big|\,|b\rangle\big|^2\,. \end{aligned} \tag{2.29}$$

Parallelogram law

$$\begin{aligned} |\mathbf{a}+\mathbf{b}|^2 + |\mathbf{a}-\mathbf{b}|^2 &= 2\left(|\mathbf{a}|^2 + |\mathbf{b}|^2\right), \\ \big|\,|a\rangle + |b\rangle\big|^2 + \big||a\rangle - |b\rangle\big|^2 &= 2\left(||a\rangle|^2 + ||b\rangle|^2\right). \end{aligned} \tag{2.30}$$

Triange inequality

$$\begin{aligned} |\mathbf{a}+\mathbf{b}| &\le |\mathbf{a}| + |\mathbf{b}|\,, \\ \big|\,|a\rangle + |b\rangle\big| &\le \big|\,|a\rangle\big| + \big|\,|b\rangle\big|\,. \end{aligned} \tag{2.31}$$

Cauchy–Schwarz inequality

$$\begin{aligned} |\langle \mathbf{a}, \mathbf{b}\rangle|^2 &\le \langle \mathbf{a}, \mathbf{a}\rangle\langle \mathbf{b}, \mathbf{b}\rangle, \\ |\langle a|b\rangle|^2 &\le \langle a|a\rangle\langle b|b\rangle. \end{aligned} \tag{2.32}$$

Proof of (2.32): Let $|d\rangle = |a\rangle + c\,|b\rangle$, where c is a scalar whose value we choose later. Then $\langle d| = \langle a| + c^\star\langle b|$. By the properties of the inner product,

$$0 \le \langle d|d\rangle = \langle a|a\rangle + c^\star\langle b|a\rangle + c\langle a|b\rangle + |c|^2\langle b|b\rangle. \tag{2.33}$$

Now choose $c = -\langle b|a\rangle/\langle b|b\rangle$. Then $c^\star = -\langle a|b\rangle/\langle b|b\rangle$ and (2.33) becomes

$$0 \le \langle a|a\rangle - |\langle a|b\rangle|^2/\langle b|b\rangle, \tag{2.34}$$

which on rearrangement gives the required result.

Further reading

Much of the material in this section is covered in §8 of RHB and §II of DK. For another introduction to the concept of dual vectors see §1.3 of Shankar's *Principles of Quantum Mechanics*. (The first chapter of Shankar gives a succinct summary of the first half of this course.)

Beware in that most books written for mathematicians the inner product $\langle \mathbf{a}, \mathbf{b}\rangle$ is defined to be linear in the *first* argument.

3 Linear operators

Throughout this section I assume that $\mathcal{V}$ is an n-dimensional inner-product space and that $|e_1\rangle, \dots, |e_n\rangle$ are an orthonormal basis for this space.

Recall that linear operators are mappings of a vector space $\mathcal{V}$ to itself that satisfy the conditions (1.23). Let A be a linear operator and suppose that

$$|b\rangle = A\,|a\rangle. \tag{3.1}$$

In §1.4 we saw how this is equivalent to the matrix equation $\mathbf{b} = A\mathbf{a}$, where $\mathbf{a}$ and $\mathbf{b}$ are the column vectors that represent $|a\rangle$ and $|b\rangle$ respectively. Armed with our inner product we can now obtain an explicit expression for the elements of the matrix representing the operator A. Expanding $|a\rangle = \sum_{k=1}^n a_k\,|e_k\rangle$ and $|b\rangle = \sum_{i=1}^n b_i\,|e_i\rangle$, equation (3.1) becomes

$$\sum_{i=1}^n b_i\,|e_i\rangle = \sum_{k=1}^n a_k A\,|e_k\rangle. \tag{3.2}$$

Now choose any basis bra $\langle e_j|$ and apply it to both sides:

$$\begin{aligned}\langle e_j|\left(\sum_{i=1}^n b_i\,|e_i\rangle\right) &= \langle e_j|\left(\sum_{k=1}^n a_k A\,|e_k\rangle\right)\\ \sum_{i=1}^n b_i \underbrace{\langle e_j|e_i\rangle}_{\delta_{ji}} &= \sum_{k=1}^n \langle e_j|\,A\,|e_k\rangle a_k.\end{aligned} \tag{3.3}$$

Therefore equation (3.1) can be represented by the matrix equation

$$b_j = \sum_{k=1}^n A_{jk} a_k, \tag{3.4}$$

where $A_{jk} \equiv \langle e_j|\,A\,|e_k\rangle = \langle \mathbf{e}_j, A\mathbf{e}_k\rangle$ are the **matrix elements** of the operator A in the $|e_1\rangle, \dots, |e_n\rangle$ basis.

3.1 The identity operator

The identity operator I defined through $I\,|v\rangle = |v\rangle$ for all $|v\rangle \in \mathcal{V}$ is clearly a linear operator. Less obviously, it can be written as

$$I = \sum_{i=1}^n |e_i\rangle\langle e_i|\,. \tag{3.5}$$

This is sometimes known as **resolution of the identity**.

Proof: Any $|v\rangle \in \mathcal{V}$ can be expressed as $|v\rangle = \sum_{i=1}^n \alpha_i\,|e_i\rangle$. Using the expression (3.5) for I, we have that

$$\begin{aligned} I\,|v\rangle &= \sum_{i=1}^n |e_i\rangle\langle e_i| \sum_{j=1}^n \alpha_j\,|e_j\rangle \\ &= \sum_{i=1}^n |e_i\rangle \sum_{j=1}^n \alpha_j \underbrace{\langle e_i|e_j\rangle}_{\delta_{ij}} \\ &= \sum_{i=1}^n |e_i\rangle \alpha_i = |v\rangle. \end{aligned} \tag{3.6}$$

The individual terms $|e_i\rangle\langle e_i|$ appearing in the sum (3.5) are known as **projection operators**: if we apply $P_i \equiv |e_i\rangle\langle e_i|$ to a vector $|a\rangle = \sum_j a_j |e_j\rangle$ the result $P_i |a\rangle = a_i |e_i\rangle$. Similarly, $\langle a| P_i = \langle e_i| a_i^\star$. We have already seen a use of projection operators in the Gram–Schmidt procedure earlier (equation 2.9).

3.2 Combining operators

The composition of two linear operators A and B is another linear operator. In case it is not obvious how to show this, let us write $C = AB$ for the result of applying B first, then A. Now notice that C is a mapping from $\mathcal{V}$ to $\mathcal{V}$ and that conditions (1.23)

$$\begin{aligned} C\big(|a\rangle + |b\rangle\big) &= A\big(B(|a\rangle + |b\rangle)\big) = A\big(B|a\rangle\big) + A\big(B|b\rangle\big) = C|a\rangle + C|b\rangle, \\ C\big(\alpha |A\rangle\big) &= A\big(B(\alpha|a\rangle)\big) = A\big(\alpha B|a\rangle\big) = \alpha\big(AB|a\rangle\big) = \alpha C|a\rangle \end{aligned} \tag{3.7}$$

hold for any $|a\rangle$, $|b\rangle \in \mathcal{V}$ and $\alpha \in \mathcal{F}$.

Exercise: Show that the matrix representing C is identical to the matrix obtained by multiplying the matrix representations of the operators A and B.

Exercise: Show that sum of two linear operators is another linear operator.

In general $AB \neq BA$: the order of composition matters. The difference

$$AB - BA \equiv [A, B] \tag{3.8}$$

is another linear operator, known as the **commutator** of A and B.

Functions of operators

We can construct new linear operators by composition and addition. For example, given a linear operator A, let us define $\exp A$ in the obvious way as

$$\begin{aligned} \exp A &\equiv \lim_{N\to\infty} \left(I + \frac{1}{N}A\right)^N \\ &= I + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \cdots + \frac{1}{m!}A^m + \cdots, \end{aligned} \tag{3.9}$$

in which I is the identity operator, $A^0 = I$, $A^1 = A$, $A^2 = AA$, $A^3 = AAA$ and so on. It might not be obvious that this $\exp A$ is a linear operator, but notice that it is a mapping from $\mathcal{V}$ to itself and that for any $|a\rangle$, $|b\rangle \in \mathcal{V}$ and $\alpha \in \mathcal{F}$ we have that

$$\begin{aligned} (\exp A)(|a\rangle + |b\rangle) &= (\exp A)|a\rangle + (\exp A)|b\rangle, \\ (\exp A)(\alpha|a\rangle) &= \alpha(\exp A)|a\rangle. \end{aligned} \tag{3.10}$$

Example: what is $\exp(\alpha G)$, where $G = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$?

First note that $G^2 = -I$. Therefore $G^{2m} = (-1)^m I$ and $G^{2m+1} = (-1)^m G$. We can use this to split the expansion of the exponential into a sum of even and odd terms:

$$\begin{aligned} \exp(\alpha G) &= \sum_{k=0}^{\infty} \frac{1}{k!}(\alpha G)^k \\ &= \sum_{m=0}^{\infty} \frac{1}{(2m)!}\alpha^{2m}G^{2m} + \sum_{m=0}^{\infty} \frac{1}{(2m+1)!}\alpha^{2m+1}G^{2m+1} \\ &= \sum_{m=0}^{\infty} \frac{1}{(2m)!}(-1)^m\alpha^{2m}I + \sum_{m=0}^{\infty} \frac{1}{(2m+1)!}(-1)^m\alpha^{2m+1}G \\ &= \cos\alpha I + \sin\alpha G \\ &= \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}. \end{aligned} \tag{3.11}$$

For this reason G is known as the **generator** of two-dimensional rotations: for $\epsilon \ll 1$ the operator $I+\epsilon G$ is a rotation by an angle ϵ; from the definition (3.9) the operator $\exp(\alpha G)$ is obtained by chaining together many such infinitesmal rotations.

3.3 Adjoint of an operator

For any operator $A:\mathcal{V}\to\mathcal{V}$ there is another operator $A^\dagger:\mathcal{V}\to\mathcal{V}$, called the **adjoint** to A, that satisfies

$$\langle \mathbf{a}, A\mathbf{b}\rangle = \langle A^\dagger \mathbf{a}, \mathbf{b}\rangle. \tag{3.12}$$

Using property (2.3) of the scalar product, this is equivalent to

$$\langle \mathbf{a}, A\mathbf{b}\rangle = \langle \mathbf{b}, A^\dagger \mathbf{a}\rangle^\star, \tag{3.13}$$

or, in bra-ket notation

$$\langle a|\, A\, |b\rangle = \langle b|\, A^\dagger\, |a\rangle^\star. \tag{3.14}$$

Now let us show that the adjoint $A^\dagger$ exists, is unique, and is linear. Consider the map $L_{\mathbf{a}}:\mathcal{V}\to\mathcal{F}$ defined by

$$L_{\mathbf{a}}(\mathbf{b}) = \langle \mathbf{a}, A\mathbf{b}\rangle. \tag{3.15}$$

As the inner product is linear in its second argument, this $L_{\mathbf{a}}$ is a linear map from vectors $\mathbf{b}$ to scalars $\mathcal{F}$. That is, it is a dual vector: for each choice of $\mathbf{a}$ and A there is a unique $\mathbf{a}'\in\mathcal{V}$ for which $L_{\mathbf{a}}(\mathbf{b}) = \langle \mathbf{a}, A\mathbf{b}\rangle = \langle \mathbf{a}', \mathbf{b}\rangle$. Following (3.12) we define $A^\dagger:\mathcal{V}\to\mathcal{V}$ to be the mapping that returns this $\mathbf{a}'$ given $\mathbf{a}$. So $A^\dagger$ exists and is unique. To show that it is linear, take the complex conjugate of (3.12) and replace $\mathbf{a}$ by $\alpha_1\mathbf{a}_1+\alpha_2\mathbf{a}_2$:

$$\begin{aligned} \langle \mathbf{b}, A^\dagger(\alpha_1\mathbf{a}_1 + \boldsymbol{\alpha}_2\mathbf{a}_2)\rangle &= \langle A\mathbf{b}, \alpha_1\mathbf{a}_1 + \boldsymbol{\alpha}_2\mathbf{a}_2\rangle \\ &= \alpha_1\langle A\mathbf{b}, \mathbf{a}_1\rangle + \alpha_2\langle A\mathbf{b}, \mathbf{a}_2\rangle \\ &= \alpha_1\langle \mathbf{b}, A^\dagger\mathbf{a}_1\rangle + \alpha_2\langle \mathbf{b}, A^\dagger\mathbf{a}_2\rangle. \end{aligned} \tag{3.16}$$

Since this holds for any $\mathbf{b}$ we have that $A^\dagger(\alpha_1\mathbf{a}_1+\alpha_2\mathbf{a}_2) = \alpha_1 A^\dagger\mathbf{a}_1 + \alpha_2 A^\dagger\mathbf{a}_2$: the $A^\dagger$ defined by (3.12) is a linear operator.

Setting $\mathbf{a}=\mathbf{e}_i$ and $\mathbf{b}=\mathbf{e}_j$ in (3.13) shows that the matrix representating the operator $A^\dagger$ has elements $(A^\dagger)_{ij} = A^\star_{ji}$: the matrix for the adjoint operator $A^\dagger$ is the Hermitian transpose of that for the original operator A.

> **Exercise:** Show that the dual to the vector $A\,|v\rangle$ is $\langle v|\, A^\dagger$.

3.4 Hermitian, unitary and normal operators

An operator A is **Hermitian** if it is self-adjoint: $A^\dagger = A$.

Unitary operators U are those for which $\langle U\mathbf{a}, U\mathbf{b}\rangle = \langle \mathbf{a}, \mathbf{b}\rangle$ for all $\mathbf{a},\mathbf{b}\in\mathcal{V}$: applying U to any pair of vectors preserves the inner product. From (3.12) this is equivalent to $UU^\dagger = U^\dagger U = I$.

> **Exercise:** Show that the composition of two unitary operators produces another unitary operator. Does the same result hold for Hermitian operators? If not, find a condition under which it does hold.
>
> **Exercise:** For real vector spaces, show that hermitian operators correspond to symmetric matrices and unitary operators to orthogonal matrices.

Hermitian and unitary operators are special cases of the more general class of **normal** operators for which $[A, A^\dagger] = 0$.

3.5 Change of basis

Calculations inevitably involve choosing a basis. Sometimes it is convenient to do parts of a calculation in one basis and the rest in another. Therefore it is important to know how to transform vector coordinates and matrix elements from one basis to another.

Suppose we have two different orthonormal basis sets, $\{|e_1\rangle,\ldots,|e_n\rangle\}$ and $\{|e'_1\rangle,\ldots,|e'_n\rangle\}$, which are related via $|e'_i\rangle = U|e_i\rangle$, where U is some operator whose properties we leave open for the moment. It follows then that (exercise: show this!)

$$|e'_i\rangle = \sum_{j=1}^n U_{ji}\,|e_j\rangle. \tag{3.17}$$

That is, the i^{th} column of the matrix U gives the representation of $|e'_i\rangle$ in the unprimed basis. The corresponding relationship for the basis bras is

$$\langle e'_i| = \sum_{j=1}^n \langle e_j|\,U^\star_{ji}. \tag{3.18}$$

The orthonormality of the bases places constraints on the the transformation matrix U:

$$\delta_{ij} = \langle e'_i|e'_j\rangle = \left(\sum_{k=1}^n \langle e_k|\,U^\star_{ki}\right)\left(\sum_{l=1}^n U_{lj}\,|e_l\rangle\right) = \sum_{k=1}^n\sum_{l=1}^n U^\star_{ki}U_{lj}\langle e_k|e_l\rangle = \sum_{k=1}^n U^\star_{ki}U_{kj} = \sum_{k=1}^n (U^\dagger)_{ik}U_{kj}, \tag{3.19}$$

or, in matrix form, $I = U^\dagger U$: the matrix U describing the coordinate transformation **must be unitary**: unitary matrices are the generalization of real orthogonal transformations (i.e., rotations and reflections) to complex vector spaces.

Transformation of vector components Taking an arbitrary vector $|a\rangle = \sum_{j=1}^n a_j\,|e_j\rangle = \sum_{j=1}^n a'_j\,|e'_j\rangle$ and using $\langle e'_i|$ to project $|a\rangle$ along the i^{th} primed basis vector gives

$$a'_i = \langle e'_i|a\rangle = \sum_{k=1}^n \langle e_k|\,U^\star_{ki}\sum_{j=1}^n a_j\,|e_j\rangle = \sum_{k=1}^n U^\star_{ki}\sum_{j=1}^n a_j\langle e_k|e_j\rangle = \sum_{j=1}^n U^\star_{ji}a_j = \sum_{j=1}^n (U^\dagger)_{ij}a_j. \tag{3.20}$$

In matrix form, the components in the primed basis are given by $\mathbf{a}' = U^\dagger\mathbf{a}$.

Transformation of matrix elements In the primed basis, the operator A has matrix elements

$$A'_{ij} \equiv \langle e'_i|\,A\,|e'_j\rangle = \left(\sum_{k=1}^n \langle e_k|\,U^\star_{ki}\right) A \left(\sum_{l=1}^n U_{lj}\,|e_l\rangle\right) = \sum_{k=1}^n U^\star_{ki}\sum_{l=1}^n \langle e_k|\,A\,|e_l\rangle U_{lj} = \sum_{k,l=1}^n (U^\dagger)_{ik}A_{kl}U_{lj}, \tag{3.21}$$

so that $A' = U^\dagger AU$. What does this mean? When applying the matrix A' to a vector whose coordinates $\mathbf{a}'$ are given with respect to the $|e'_i\rangle$ basis, think of U as transforming from $\mathbf{a}'$ to $\mathbf{a}$. Then we apply the matrix A to the result before transforming back to the primed basis with $U^\dagger$.

[More generally, matrices A and B are said to be **similar** if they related by $B = P^{-1}AP$ for some invertible matrix P. That is, they represent the same linear operator under two different (possibly non-orthonormal) bases, with P being the matrix that effects the change in basis.]

Exercise: Derive the change-of-basis formulae (3.20) and (3.21) by resolving the identity (3.5). That is, write

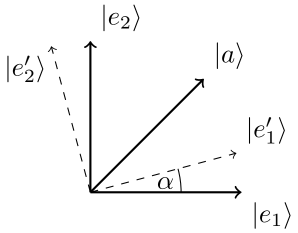
$$\begin{aligned} a'_i &= \langle e'_i|a\rangle = \langle e'_i|\,I\,|a\rangle \\ A'_{ij} &= \langle e'_i|\,A\,|e'_j\rangle = \langle e'_i|\,IAI\,|e'_j\rangle \end{aligned} \tag{3.22}$$

and use the fact that I can be expressed as $I = \sum_k |e_k\rangle\langle e_k| = \sum_l |e_l\rangle\langle e_l|$. You should find that

$$\begin{aligned} a'_i &= \sum_j \langle e'_i|e_j\rangle a_j, \\ A'_{ij} &= \sum_k \langle e'_i|e_k\rangle \sum_l \langle e_k|\, A\, |e_l\rangle\langle e_l|e'_j\rangle, \end{aligned} \tag{3.23}$$

where $\langle e'_i|e_j\rangle$ is the projection of $|e_j\rangle$ onto the $|e'_i\rangle$ basis – that is, a matrix whose j^{th} column expresses $|e_j\rangle$ in the primed basis. How is this matrix related to U introduced above?

Example: Two-dimensional rotations



Suppose that the $(|e'_1\rangle, |e'_2\rangle)$ basis is related to the $(|e_1\rangle, |e_2\rangle)$ basis by a rotation through an angle α. From the diagram, the basis vectors are related through

$$\begin{aligned} |e'_1\rangle &= \cos\alpha\,|e_1\rangle + \sin\alpha\,|e_2\rangle \\ |e'_2\rangle &= -\sin\alpha\,|e_1\rangle + \cos\alpha\,|e_2\rangle \end{aligned} \quad \Rightarrow \quad \begin{pmatrix} |e'_1\rangle \\ |e'_2\rangle \end{pmatrix} = U^{\mathrm{T}} \begin{pmatrix} |e_1\rangle \\ |e_2\rangle \end{pmatrix}, \tag{3.24}$$

with

$$U^{\mathrm{T}} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}, \quad \text{so that} \quad U = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}. \tag{3.25}$$

The i^{th} column of U expresses $|e'_i\rangle$ in the $|e_j\rangle$ basis: $U_{ji} = \langle e_j|e'_i\rangle$. Clearly the coordinates $\mathbf{a}$, $\mathbf{a}'$ of a vector $|a\rangle$ in the two bases are related through $\mathbf{a}' = U^{\mathrm{T}}\mathbf{a}$.

3.6 Trace

The trace $\operatorname{tr} A$ of an $n \times n$ matrix A is the sum of its diagonal elements:

$$\operatorname{tr} A = \sum_{i=1}^{n} A_{ii}. \tag{3.26}$$

The trace satisfies

$$\operatorname{tr}(AB) = \operatorname{tr}(BA) \tag{3.27}$$

because

$$\operatorname{tr}(AB) = \sum_{i=1}^{n} \underbrace{\left(\sum_{j=1}^{n} A_{ij} B_{ji} \right)}_{(AB)_{ii}} = \sum_{j=1}^{n} \underbrace{\left(\sum_{i=1}^{n} B_{ji} A_{ij} \right)}_{(BA)_{jj}} = \operatorname{tr}(BA). \tag{3.28}$$

Taking $A = A_1$, $B = A_2 A_3 \cdots A_m$ we see that $\operatorname{tr}(A_1 A_2 \cdots A_m) = \operatorname{tr}(A_2 \cdots A_m A_1)$: the trace is invariant under cyclic permutations.

Exercise: Show that the trace is independent of basis. Explain then how, given a 3×3 rotation matrix, one can find the rotation angle directly from the trace.

3.7 Determinant

Suppose $\mathcal{V}_1, ..., \mathcal{V}_k$ are vector spaces over a common field of scalars $\mathcal{F}$. A map $f : \mathcal{V}_1 \times \cdots \times \mathcal{V}_k \to \mathcal{F}$ is **multilinear**, specifically k-linear, if it is linear in each variable separately:

$$f(\mathbf{v}_1, ..., \alpha\mathbf{v}_i + \alpha'\mathbf{v}'_i, ..., \mathbf{v}_k) = \alpha f(\mathbf{v}_1, ..., \mathbf{v}_i, ..., \mathbf{v}_k) + \alpha' f(\mathbf{v}_1, ..., \mathbf{v}'_i, ..., \mathbf{v}_k). \tag{3.29}$$

For the special case $k = 2$ the map is called bilinear. The inner product of two real vectors is an example of a bilinear map.

> **Exercise:** Given two k-linear maps, f_1 and f_2, show that any linear combination, $\alpha_1 f_1 + \alpha_2 f_2$, is also k-linear.

A multilinear map is **alternating** if it returns zero whenever two of its arguments are equal:

$$f(\mathbf{v}_1, ..., \mathbf{v}_i, ..., \mathbf{v}_i, ..., \mathbf{v}_k) = 0. \tag{3.30}$$

> **Exercise:** Use conditions (3.29) and (3.30) to show that the output of a multilinear alternating map changes sign when two of its arguments are exchanged.

The **determinant** is the (unique) mapping from $n \times n$ matrices to scalars that is n-linear alternating in the columns, and takes the value 1 for the identity matrix. Some immediate consequences of this definition are that

(i) If two columns of A are identical then $\det A = 0$.
(ii) Swapping two columns of A changes the sign of $\det A$.
(iii) If B is obtained from A by multiplying a single column of A by a factor c then $\det B = c \det A$.
(iv) If one column of A consists entirely of zeros then $\det A = 0$.
(v) Adding a multiple of one column to another does not change $\det A$.

Before showing how to construct this unique mapping, we develop some ideas to simplify keeping track of the sign changes that occur when columns are swapped.

Permutations A permutation of the list $(1, 2, \ldots, m)$ is another list that contains each of the numbers 1, 2, $\ldots m$ exactly once. In other words, it is a straightforward shuffling of the order of the elements. There are $m!$ permutations of an m-element list.

Given a permutation P we write $P(1)$ for the first element in the shuffled list, $P(2)$ for the second, etc. Then P can be written as $(P(1), P(2), \ldots, P(m))$. An alternative notation is

$$P = \begin{pmatrix} 1 & 2 & \ldots & m \\ P(1) & P(2) & \ldots & P(m) \end{pmatrix}, \tag{3.31}$$

which emphasises that P is a mapping from the set $(1, \ldots, m)$ (top row) to itself (values given on bottom row). From any two permutation mappings P and Q we can compose a new one PQ defined through $(PQ)(i) = P(Q(i))$. There is an identity mapping (for which $P(i) = i$) and every P has an inverse

$$P^{-1} = \begin{pmatrix} P(1) & P(2) & \ldots & P(m) \\ 1 & 2 & \ldots & m \end{pmatrix}, \tag{3.32}$$

which is well defined because each number $1, 2, \ldots, m$ appears exactly once in the top row of (3.32).

Any permutation $(P(1), P(2), \ldots, P(m))$ can be constructed from $(1, 2, \ldots, m)$ by a sequence of pairwise element exchanges. **Even** (**odd**) permutations require an even (odd) number of exchanges. The **sign** of a permutation is defined as

$$\operatorname{sgn}(P) = \begin{cases} +1, & \text{if } P \text{ is an even permutation of } (1, 2, \ldots, m), \\ -1, & \text{if } P \text{ is an odd permutation of } (1, 2, \ldots, m). \end{cases} \tag{3.33}$$

Given two permutations P, Q, we have $\mathrm{sgn}(PQ) = \mathrm{sgn}(P)\,\mathrm{sgn}(Q)$. The identity permutation is even. Therefore $+1 = \mathrm{sgn}(P^{-1}P) = \mathrm{sgn}(P^{-1})\,\mathrm{sgn}\,P$, showing that $\mathrm{sgn}(P^{-1}) = \mathrm{sgn}\,P$.

The following table shows all 6 permutations of the 3-elements list $(1, 2, 3)$:

	$P(1)$	$P(2)$	$P(3)$	$\mathrm{sgn}(P)$
P_1	1	2	3	+1
P_2	2	1	3	-1
P_3	2	3	1	+1
P_4	3	2	1	-1
P_5	3	1	2	+1
P_6	1	3	2	-1

Leibniz' expansion of the determinant Now we obtain an explicit expression for the determinant and show that it is unique. We can express each column vector $\mathbf{A}_j$ of the matrix A as a linear combination $\mathbf{A}_j = \sum_i A_{ij}\mathbf{e}_i$ of the column's basis vectors $\mathbf{e}_1 = (1, 0, 0, ..)^T, ..., \mathbf{e}_n = (0, .., 0, 1)^T$. For any k-linear map δ we have that, by definition,

$$\delta(\mathbf{A}_1, ..., \mathbf{A}_k) = \delta\left(\sum_{i_1=1}^{n} A_{i_1,1}\mathbf{e}_{i_1}, ..., \sum_{i_k=1}^{n} A_{i_k,k}\mathbf{e}_{i_k}\right) = \sum_{i_1=1}^{n}\cdots\sum_{i_k=1}^{n} A_{i_1,1}\cdots A_{i_k,k}\delta(\mathbf{e}_{i_1}, ..., \mathbf{e}_{i_k}), \tag{3.34}$$

showing that the map is completely determined by the n^k possible ways of applying δ to the basis vectors. Imposing the condition that δ be alternating means that that $\delta(\mathbf{e}_{i_1}, ..., \mathbf{e}_{i_n})$ vanishes if two or more of the i_k are equal. Therefore we need consider only those $(i_1, ..., i_n)$ that are permutations P of $(1, ..., n)$. The change of sign under pairwise exchanges implied by the alernating condition means that $\delta(\mathbf{e}_{P(1)}, ..., \mathbf{e}_{P(n)}) = \mathrm{sgn}(P)\delta(\mathbf{e}_1, ..., \mathbf{e}_n)$. Finally the condition that $\det I = 1$ sets $\delta(\mathbf{e}_1, ..., \mathbf{e}_n) = 1$, completely determining δ. The result is that

$$\det A = \sum_P \mathrm{sgn}(P) A_{P(1),1}A_{P(2),2}\cdots A_{P(n),n}. \tag{3.35}$$

For example, for the case $n = 3$

$$\det A = A_{11}A_{22}A_{33} - A_{21}A_{12}A_{33} + A_{21}A_{32}A_{13} - A_{31}A_{22}A_{13} + A_{31}A_{12}A_{23} - A_{11}A_{32}A_{23}, \tag{3.36}$$

where in the sum (3.35) I have taken the permutations P in the order given in the table above.

Reasoning about the properties of determinants can be sometimes be simplified if we increase the number of terms in (3.35) from a mere $n!$ up to a more substantial n^n by writing (3.35) as

$$\det A = \sum_{l_1=1}^{n}\cdots\sum_{l_n=1}^{n} \epsilon_{l_1\cdots l_n} A_{l_1,1}\cdots A_{l_n,n}, \tag{3.37}$$

where the Levi-Civita (or alternating) symbol

$$\epsilon_{l_1,...,l_n} \equiv \begin{cases} \mathrm{sgn}(l_1, ..., l_n), & \text{if } (l_1, ..., l_n) \text{ is a permutation of } (1, ..., n), \\ 0, & \text{otherwise,} \end{cases} \tag{3.38}$$

kills off the $n^n - n!$ choices of $(l_1, ..., l_n)$ in which one or more indices are repeated.

Rows versus columns For each term in (3.35) let Q be the inverse of P: if $P(i) = j$ then $Q(j) = i$. Then $A_{P(i),i} = A_{j,Q(j)}$ and we have that the product $\prod_{i=1}^n A_{P(i),i} = \prod_{j=1}^n A_{j,Q(j)}$. Since $\mathrm{sgn}(P) = \mathrm{sgn}(Q)$ and there is precisely one P for each Q and vice versa, we can rearrange the order of the terms in the sum to obtain

$$\begin{aligned} \det A &= \sum_Q \mathrm{sgn}(Q) A_{1,Q(1)}A_{2,Q(2)}\cdots A_{n,Q(n)} \\ &= \sum_Q \mathrm{sgn}(Q)(A^T)_{Q(1),1}(A^T)_{Q(2),2}\cdots(A^T)_{Q(n),n} \\ &= \det A^T, \end{aligned} \tag{3.39}$$

using (3.35). That is, the determinant is also multilinear alternating in the rows of A and the properties (i)–(v) listed above apply with the word "column" replaced by "row".

Determinant of products Another important property is that $\det(AB) = \det A \det B$. This takes a little more work to show than properties (i)–(v) above. First note that applying a permutation Q to the columns in the Leibniz expansion (3.35) gives

$$\operatorname{sgn}(Q)\det A = \sum_P \operatorname{sgn}(P) A_{P(1),Q(1)} \cdots A_{P(n),Q(n)}. \tag{3.40}$$

Now apply the Leibniz expansion to $(AB)_{ij} = \sum_k A_{ik}B_{kj}$:

$$\begin{aligned}\det(AB) &= \sum_P \sum_{k_1} \cdots \sum_{k_n} \operatorname{sgn}(P) A_{P(1),k_1} B_{k_1,1} \cdots A_{P(n),k_n} B_{k_n,n} \\ &= \sum_P \sum_Q \operatorname{sgn}(P) A_{P(1),Q(1)} B_{Q(1),1} \cdots A_{P(n),Q(n)} B_{Q(n),n},\end{aligned} \tag{3.41}$$

introducing the sum over permutations Q after observing that the right-hand side of the first line vanishes unless the k_i are distinct. Then take the $B_{Q(i),i}$ factors outside the sum over P and use (3.40):

$$\begin{aligned}\det(AB) &= \sum_Q B_{Q(1),1} \cdots B_{Q(n),n} \sum_P \operatorname{sgn}(P) A_{P(1),Q(1)} \cdots A_{P(n),Q(n)} \\ &= \sum_Q B_{Q(1),1} \cdots B_{Q(n),n} \operatorname{sgn}(Q) \det(A) \\ &= \det A \det(B).\end{aligned} \tag{3.42}$$

Laplace expansion of the determinant You are probably more familiar with another expression for the determinant. Given an $n \times n$ matrix A, let $A_{(i,j)}$ be the $(n-1)\times(n-1)$ matrix obtained by omitting the i^{th} row and j^{th} column of A. Suppose that we are given a function $\delta^{(n-1)}$ which is an alternating $(n-1)$-linear map on such $(n-1)\times(n-1)$ matrices. Then, for any choice of $i = 1, ..., n$, the new function

$$\delta_i^{(n)}(A) = \sum_{j=1}^n A_{ij}(-1)^{i+j}\delta^{(n-1)}(A_{(i,j)}) \tag{3.43}$$

is an alternating n-linear map on the columns of $n \times n$ matrices.

> **Proof:** (Multilinearity) By construction, $A_{(i,j)}$ is independent of the j^{th} column of A; it is therefore an $(n-1)$-linear function of each column of A, except for the j^{th}. Hence $A_{ij}\delta^{(n-1)}(A_{(i,j)})$ is an n-linear function of the columns of of $n \times n$ matrices. Any linear combination of n-linear functions is itself n-linear. Therefore the sum $\delta_i^{(n)}(A)$ defined by (3.43) is n-linear.
> (Alternating) Suppose that columns k and $k+l$ (assume $l > 0$) of A are equal. Then the submatrices $A_{(i,j)}$ for $j \neq k$ or $k+l$ have two equal columns, which means that applying the alternating map $\delta^{(n-1)}$ produces 0: that is, $\delta^{(n-1)}(A_{(i,j)}) = 0$ unless $j = k$ or $j = k+l$. Eq. (3.43) reduces to
>
> $$\delta_i^{(n)}(A) = A_{ik}(-1)^{i+k}\delta^{(n-1)}(A_{(i,k)}) + A_{i,k+l}(-1)^{i+k+l}\delta^{(n-1)}(A_{(i,k+l)}). \tag{3.44}$$
>
> But since columns k and $k+l$ of A are equal we have that $A_{ik} = A_{i,k+l}$ and the matrix $A_{(i,k+l)}$ can be obtained from $A_{(i,k+l)}$ by $(l-1)$ pairwise column exchanges. Then $\delta^{(n-1)}(A_{(i,k+l)}) = (-1)^{l-1}\delta^{(n-1)}(A_{(i,k)})$ and so (3.44) vanishes: the mapping defined by (3.43) is alternating.

If we define the map $\delta^{(1)}(A)$ on 1×1 matrices to return the single element A_{11}, then the $\delta_j^{(n)}$ defined by (3.43) returns 1 when fed the $n \times n$ identity matrix. So, in addition to the Leibniz expansion (3.35) of $\det A$ as a sum of permutations of A's rows/columns, we have another $i = 1, ..., n$ explicit expressions for the same quantity, given recursively in terms of lower-order submatrices:

$$\begin{aligned} \det A &= \sum_{j=1}^{n} A_{ij} c_{ij} \\ &= \sum_{j=1}^{n} A_{ij} \, (\mathrm{adj} A)_{ji}, \end{aligned} \tag{3.45}$$

where $c_{ij} = (-1)^{i+j} \det(A_{(i,j)})$ is known as the **cofactor** matrix of A and its transpose, adj$A$, is the **adjugate** matrix or **classical adjoint** of A.

A useful generalisation of (3.45) is that

$$\delta_{ij} \det A = \sum_{k=1}^{n} A_{ik} c_{jk} = \sum_{k=1}^{n} c_{ki} A_{kj} \tag{3.46}$$

or, in matrix notation,

$$(\det A) I = A \, \mathrm{adj} A = (\mathrm{adj} A) A, \tag{3.47}$$

where I is the $n \times n$ identity matrix and the **adjugate matrix** or **classical adjoint** is the transpose of the cofactor matrix: $(\mathrm{adj}A)_{ji} = (-1)^{i+j} \det(A_{(i,j)})$. From this follows the expression $A^{-1} = \frac{1}{\det A} \mathrm{adj} A$ for the inverse of A and Cramer's rule.

Proof of (3.46): Consider element (i, j) of the product $(\mathrm{adj}A)A$. Using the Leibniz expansion of the *rows* of the cofactor matrix $A_{(i,k)}$ we find

$$\begin{aligned} \sum_{k=1}^{n} c_{ki} A_{kj} &= \sum_{k=1}^{n} (-1)^{j+k} \left[\sum_{P} \mathrm{sgn}(P) A_{1,P(1)} \cdots A_{k-1,P(k-1)} A_{k+1,P(k+1)} \cdots A_{nP(n)} \right] A_{kj} \\ &= \sum_{k=1}^{n} (-1)^{j+k} \sum_{P} \mathrm{sgn}(P) A_{1,P(1)} \cdots A_{k-1,P(k-1)} A_{kj} A_{k+1,P(k+1)} \cdots A_{nP(n)} \end{aligned} \tag{3.48}$$

in which the permutation P maps the ordered list of integers $\{1, ..., n\} - \{k\}$ to $\{1, ..., n\} - \{i\}$. We can combine the sum over k and the permutations P into a sum over permutations P' from the list $(k, 1..., k-1, k+1, ..., n)$ to the list $(j, 1, ..., i-1, i+1, ..., n)$. Note that the latter has a repeated element if $i \neq j$. Clearly $\mathrm{sgn}(P) = \mathrm{sgn}(P')$. It takes $(-1)^{k+1}$ pairwise exchanges to put the domain of P' into the order $(1, ..., n)$ and $(-1)^{j+1}$ to put its range in the order $(1, ..., i-1, j, i+1, ..., n)$. So, the permutation P' can be written as a permutation Q of $(1, ..., n)$ to $(1, ..., i-1, j, i+1, ..., n)$ having $\mathrm{sgn}(Q) = (-1)^{j+k} \mathrm{sgn}(P') = (-1)^{j+k} \mathrm{sgn}(P)$. Writing (3.48) as a sum over such Q,

$$\begin{aligned} \sum_{k=1}^{n} c_{ki} A_{kj} &= \sum_{Q} \mathrm{sgn}(Q) A_{1,Q(1)} \cdots A_{n,Q(n)} \\ &= \begin{cases} \det A, & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{3.49}$$

The proof for $A \, \mathrm{adj} A$ is similar.

Exercise: Use the properties above together with results derived in §3.5 to show that the determinant is independent of (orthonormal) basis.

Exercise: Use the Laplace expansion to show that

$$\frac{\partial \det A}{\partial A_{ij}} = (\mathrm{adj} A)_{ij}. \tag{3.50}$$

Hence show that for matrices $A(\alpha)$ whose elements depend on a parameter α we have that

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\det A = (\det A)\,\mathrm{tr}\left(A^{-1}\frac{\mathrm{d}A}{\mathrm{d}\alpha}\right). \tag{3.51}$$

Geometrical meaning of the determinant In n-dimensional Euclidean space the **block** spanned by the n vectors $\mathbf{v}_1, ..., \mathbf{v}_n$ is the set of points $\lambda_1\mathbf{v}_1 + \cdots + \lambda_n\mathbf{v}_n$ for $0 \le \lambda_1, ..., \lambda_n \le 1$. Let us write $\mathrm{Vol}(\mathbf{v}_1, ..., \mathbf{v}_n)$ for the volume of space occupied by this block and define the **oriented volume** of the block to be

$$\mathrm{Vol}^0(\mathbf{v}_1, ..., \mathbf{v}_n) = \pm\mathrm{Vol}(\mathbf{v}_1, ..., \mathbf{v}_n), \tag{3.52}$$

the sign being chosen according to the sign of $\det(\mathbf{v}_1, ..., \mathbf{v}_n)$, the determinant of the matrix whose i^{th} column is $\mathbf{v}_i$.

Claim: $\mathrm{Vol}^0(\mathbf{v}_1, ..., \mathbf{v}_n) = \det(\mathbf{v}_1, ..., \mathbf{v}_n)$.

> **Justification:** We need to show that the oriented volume (LHS) is multilinear, alternating and takes the value $+1$ when fed the vectors $\mathbf{v}_i = \mathbf{e}_i$ (RHS). If the $\mathbf{v}_i$ are linearly dependent then the volume is zero and so too is the determinant. For the LI case, consider first the case $\mathbf{v}_i = \mathbf{e}_i$. The block B is then the unit cube, which has $\mathrm{Vol}^0 = +1$, in agreement with RHS. The definition (3.52) of Vol^0 is clearly alternating. All that remains is to prove that Vol^0 is linear in each argument. This last step is intuitively plausible, but the proof is more involved than we need. See, e.g., XX,§4 of Lang's *Undergraduate analysis* for the details.

Therefore the determinant, $\det A$, of a linear operator A is simply the change in (oriented) volume effected by the operator. In particular, $\det A = +1$ if A preserves volume and orientation, while $\det A = -1$ if A preserves volume but flips the orientation (i.e., reflects).

3.8 Reduction to triangular matrices

Recall that any matrix can be put in to upper (or lower) triangular form by carrying out a sequence of **elementary row operations**:

- swap two rows;
- multiply one row by a non-zero constant;
- add a multiple of one row to another.

Each such operation produces a new matrix whose columns are linear combinations of the columns of the original matrix. A simple way of calculating the rank of a matrix is to use elementary row operations to reduce it to either an **upper triangular matrix** or a **lower triangular matrix** and counting the number of LI columns in the result.

Examples:

$$\mathrm{rank}\begin{pmatrix}1&0&0\\0&1&0\\0&0&3\end{pmatrix} = 3, \quad \mathrm{rank}\begin{pmatrix}0&0&0\\0&1&-1\\0&1&1\end{pmatrix} = 2, \quad \mathrm{rank}\begin{pmatrix}0&0&0\\0&2&1\\0&2&1\end{pmatrix} = 1. \tag{3.53}$$

Every elementary row operation can be expressed as an **elementary matrix**. For example, starting with a 3 × 3 matrix A and adding α times row 3 to row 2, then swapping the first two rows gives a new matrix E_2E_1A, where the elementary matrices E_1 and E_2 are

$$E_1 = \begin{pmatrix}1&0&0\\0&1&\alpha\\0&0&1\end{pmatrix}, \quad E_2 = \begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}. \tag{3.54}$$

This reduction to triangular form is usually the sanest simple way of carrying out real calculations.

Solving linear equations To solve the linear equation $A\mathbf{x} = \mathbf{b}$ apply elementary row operations E_1, E_2, …to A to reduce it to upper triangular form. That is,

$$\begin{aligned} & A\mathbf{x} = \mathbf{b} \\ \Rightarrow \quad & (E_m \cdots E_2 E_1)A\mathbf{x} = (E_m \cdots E_2 E_1)\mathbf{b} \\ \Rightarrow \quad & \begin{pmatrix} A'_{11} & A'_{12} & A'_{13} & \cdots & A'_{1,n-1} & A'_{1n} \\ 0 & A'_{22} & A'_{23} & \cdots & A'_{2,n-1} & A'_{2n} \\ 0 & 0 & A'_{33} & \cdots & A'_{3,n-1} & A'_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_{n-1,n-1} & A_{n-1,n} \\ 0 & 0 & 0 & \cdots & 0 & A_{n,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} b'_1 \\ b'_2 \\ b'_3 \\ \vdots \\ b'_{n-1} \\ b'_n \end{pmatrix}, \end{aligned} \tag{3.55}$$

where $A'_{ij} = (E_m \cdots E_2 E_1 A)_{ij}$ and $b'_i = (E_m \cdots E_2 E_1 \mathbf{b})_i$. Then the x_i can be found by backsubstitution.

Matrix inverse To find the inverse (assuming that it exists), of a square matrix A, apply elementary row operations to reduce $A \to E_1 A \to E_2 E_1 A \to E_m \cdots E_2 E_1 A = I$ while simultaneously applying the same operations to the identity I: $I \to E_1 I \to E_2 E_1 I \to E_m \cdots E_2 E_1 I$. Then $(E_m \cdots E_1)A = I$, so that $A^{-1} = (E_m \cdots E_1)I$.

Determinant Similarly, following on from the example immediately above, the determinant of A is given by $(\det A)^{-1} = (\det E_m) \cdots (\det E_1)$. This is useful because calculating the determinants of an elementary matrix is trivial. The resulting expression involves multiplying just $O(n)$ numbers instead summing of the $n!$ terms that appear in the Leibniz and Laplace expansions of the determinant.

3.9 Eigenvectors and diagonalisation

Recall that $|v\rangle \neq 0$ is an eigenvector of an operator A with eigenvalue λ if it satisfies the **eigenvalue equation**

$$A\,|v\rangle = \lambda\,|v\rangle. \tag{3.56}$$

To find $|v\rangle$ we first find λ by rewriting eq above as

$$A\,|v\rangle - \lambda\,|v\rangle = (A - \lambda I)\,|v\rangle = 0. \tag{3.57}$$

Clearly $A - \lambda I$ can't be invertible: if it were, then we could operate on 0 with $(A - \lambda I)^{-1}$ to get $|v\rangle$. So, we must have that

$$\det(A - \lambda I) = 0, \tag{3.58}$$

which is known as the **characteristic equation** for A. The characteristic equation is an n^{th}-order polynomial in λ which can always be written in the form $(\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n) = 0$, where the n roots (i.e., eigenvalues) $\lambda_1, \ldots, \lambda_n$ are in general complex and not necessarily distinct. If a particular value of λ appears $k > 1$ times then that eigenvalue is said to be k-fold **degenerate**. (The integer k is also sometimes called the **multiplicity** of the eigenvalue.)

> **Exercise:** Let A be a linear operator with eigenvector $|v\rangle$ and corresponding eigenvalue λ. Show that $|v\rangle$ is also an eigenvector of the linear operator $\exp(\alpha A)$, but with eigenvalue $\exp(\alpha\lambda)$.

Example: find the eigenvalues and eigenvectors of

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}. \tag{3.59}$$

The characteristic equation is $(1 - \lambda)(2 - \lambda)(1 - \lambda) - (2 - \lambda) = 0$, which, when factorised, is $(\lambda - 2)^2\lambda = 0$. Therefore the eigenvalues are $\lambda = 0, 2, 2$: the eigenvalue $\lambda = 2$ is doubly degenerate.

To find the eigenvector $|v_1\rangle$ corresponding to the first eigenvalue with $\lambda = \lambda_1 = 0$, take $|v_1\rangle = (x_1, x_2, x_3)^{\mathrm{T}}$ and substitute into eigenvalue equation (3.56) to find $x_1 = -x_3$ and $x_2 = 0$. Any vector $|v\rangle$ that satisfies these constraints is an eigenvector of A with eigenvalue 0. For example, we could take $|v_1\rangle = (1, 0, -1)^{\mathrm{T}}$ or $(-\pi, 0, \pi)^{\mathrm{T}}$ or even $(\mathrm{i}, 0, -\mathrm{i})^{\mathrm{T}}$ (assuming we have a complex vector space). It is usually most convenient, however, to choose them to make $|v\rangle$ normalized. Therefore we choose $|v_1\rangle = \frac{1}{\sqrt{2}}(1, 0, -1)^{\mathrm{T}}$.

Taking $\lambda = \lambda_2 = \lambda_3 = 2$ and substituting $\mathbf{v} = (x_1, x_2, x_3)^{\mathrm{T}}$ into the eigenvalue equation yields the constraints $-x_1 + x_3 = 0$ and $x_1 - x_3 = 0$. So, we must set $x_1 = x_3$, but we are free to choose anything for x_2. For $|v_2\rangle$ let us take $x_1 = x_3 = 1/\sqrt{2}$ and $x_2 = 0$, while for $|v_3\rangle$ we choose $x_1 = x_3 = 0$ and $x_2 = 1$. To summarise, we have found the following eigenvalues and eigenvectors for the matrix A:

$$\lambda_1 = 0, \quad |v_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \lambda_2 = 2, \quad |v_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \lambda_3 = 2, \quad |v_3\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \tag{3.60}$$

Diagonalisation Notice that for the 3×3 matrix A above we were able to find three eigenvectors $|v_1\rangle$, $|v_2\rangle$, $|v_3\rangle$. These three eigenvectors turn out to be orthogonal. Therefore they can be used as a basis for the 3-dimensional vector space on which A operates. In this **eigenbasis** the matrix representing the operator A takes on a particularly simple form: it is **diagonal**, with matrix elements

$$\langle v_i|\, A\, |v_j\rangle = \lambda_i \delta_{ij}. \tag{3.61}$$

3.10 The complex spectral theorem*

A fundamental problem of linear algebra is to find the conditions under which a given matrix A can be diagonalised, i.e., whether there exists an invertible coordinate transformation P for which $P^{-1}AP$ becomes diagonal. For our purposes, we need only accept that a sufficient condition is that A be Hermitian (see next section).

If you don't like to accept such things, let A be an operator on a *complex* inner-product space $\mathcal{V}$. It is relatively easy to show that $\mathcal{V}$ has an orthonormal basis consisting of eigenvectors of A if and only if A is **normal**. Recall that normal matrices satisfy $AA^\dagger = A^\dagger A$: Hermitian and unitary matrices are special cases of normal matrices.

Here is an outline sketch of the proof. Let $\mathbf{v}_1, ..., \mathbf{v}_n$ be the eigenvectors of A and $\lambda_1, ..., \lambda_n$ the corresponding eigenvalues.

First suppose that $\mathcal{V}$ has an orthonormal basis consisting of the eigenvectors $\mathbf{v}_i$. In this basis the matrix representing A is diagonal with matrix elements $A_{ij} = \langle \mathbf{v}_i, A\mathbf{v}_j\rangle$. The matrix representing $A^\dagger$ has elements $(A^\dagger)_{ij} = \left\langle \mathbf{v}_i, A^\dagger \mathbf{v}_j \right\rangle = \langle \mathbf{v}_j, A\mathbf{v}_i\rangle^\star$, which is also diagonal. Diagonal matrices commute. Therefore $AA^\dagger = A^\dagger A$: A is normal.

Now the converse: we need to show that if A is normal then $\mathcal{V}$ has an orthonormal basis consisting of eigenvectors of A. Any matrix can be reduced to upper triangular form by applying an appropriate set of elementary row operations (see §3.8). In particular, there is an orthonormal basis $\mathbf{e}'_1, ..., \mathbf{e}'_n$ in which the matrix representing A becomes

$$\left\langle \mathbf{e}'_i, A\mathbf{e}'_j \right\rangle = \begin{pmatrix} a'_{11} & \cdots & a'_{1n} \\ & \ddots & \vdots \\ 0 & & a'_{nn} \end{pmatrix}. \tag{3.62}$$

From (3.62) we see that

$$\begin{aligned} \left|A\mathbf{e}'_1\right|^2 &= |a'_{11}|^2 \quad \text{and} \\ \left|A^\dagger \mathbf{e}'_1\right|^2 &= |a'_{11}|^2 + |a'_{12}|^2 + \cdots + |a'_{1n}|^2. \end{aligned} \tag{3.63}$$

But A is normal, so $\left|A\mathbf{e}'_1\right| = \left|A^\dagger \mathbf{e}'_1\right|$. This means that all entries in the first row of the matrix in the RHS of (3.62) vanish, except possibly for the first.

Similarly, from equation (3.62) we have that

$$\begin{aligned} \left|A\mathbf{e}'_2\right|^2 &= |a'_{22}|^2 + |a'_{12}|^2 \\ &= |a'_{22}|^2, \quad \text{and} \\ \left|A^\dagger \mathbf{e}'_1\right|^2 &= |a'_{22}|^2 + |a'_{23}|^2 + \cdots + |a'_{1n}|^2, \end{aligned} \tag{3.64}$$

using the result that $a'_{12} = 0$ from the previous paragraph. Because A is normal we have that $\left|A\mathbf{e}'_2\right| = \left|A^\dagger \mathbf{e}'_2\right|$: the whole second row must equal zero, except possibly for the diagonal element a'_{22}.

Repeating this procedure for $A\mathbf{e}'_3$ to $A\mathbf{e}'_n$ shows that all of the off-diagonal elements in (3.62) must be zero. Therefore the orthonormal basis vectors $\mathbf{e}'_1$,...,$\mathbf{e}'_n$ are eigenvectors of A.

The corresponding result for *real* vector spaces is harder to prove. It states that a real vector space $\mathcal{V}$ has an orthonormal basis consisting of eigenvectors of A if and only if A is *symmetric* (which is equivalent to Hermitian for the special case of a real vector space).

3.11 Hermitian operators

Hermitian operators are particularly important. Here are two central results. If A is Hermitian then:

(1) its eigenvalues are real
(2) its eigenvectors are orthogonal (and therefore form a basis of $\mathcal{V}$)

Proof of (1): the eigenvalues of a Hermitian operator are real Let $|v\rangle$ be a eigenvector of A with eigenvalue λ. Then the eigenvalue equation (3.56) and its dual are

$$\begin{aligned} A\,|v\rangle &= \lambda\,|v\rangle \\ \langle v|\,A^\dagger &= \lambda^\star \langle v|\,. \end{aligned} \tag{3.65}$$

As A is Hermitian we have $A^\dagger = A$. Operate on the first of (3.65) with $\langle v|$ and use the second to operate on $|v\rangle$. Subtracting, the result is

$$0 = (\lambda - \lambda^\star)\langle v|v\rangle. \tag{3.66}$$

But $\langle v|v\rangle > 0$. Therefore $\lambda = \lambda^\star$: the eigenvalues are real.

Proof of (2): the eigenvectors of a Hermitian operator are orthogonal Let $|v_1\rangle$ and $|v_2\rangle$ be two eigenvectors with corresponding eigenvalues λ_1, λ_2. The eigenvalue equations are

$$\begin{aligned} A\,|v_1\rangle &= \lambda_1\,|v_1\rangle, \\ A\,|v_2\rangle &= \lambda_2\,|v_2\rangle \end{aligned} \tag{3.67}$$

For simplicity, let us first consider the case $\lambda_1 \neq \lambda_2$. Operating on the first of (3.67) with $\langle v_2|$ and on the second with $\langle v_1|$ results in

$$\begin{aligned} \langle v_2|\,A\,|v_1\rangle &= \lambda_1\langle v_2|v_1\rangle, \\ \langle v_1|\,A\,|v_2\rangle &= \lambda_2\langle v_1|v_2\rangle. \end{aligned} \tag{3.68}$$

Taking the complex conjugate of the second of these gives

$$\begin{aligned} \langle v_1|\,A\,|v_2\rangle^\star &= \lambda_2^\star\langle v_1|v_2\rangle^\star \\ \Rightarrow \quad \langle v_2|\,A\,|v_1\rangle &= \lambda_2\langle v_2|v_1\rangle, \end{aligned} \tag{3.69}$$

since $\lambda_2 = \lambda_2^\star$ and $\langle v_1|\,A\,|v_2\rangle^\star = \langle v_2|\,A^\dagger\,|v_1\rangle = \langle v_2|\,A\,|v_1\rangle$. Now subtract (3.69) from the first of (3.68):

$$0 = (\lambda_1 - \lambda_2)\langle v_1|v_2\rangle. \tag{3.70}$$

Under our assumption that $\lambda_1 \neq \lambda_2$ we must have $\langle v_1|v_2\rangle = 0$: the eigenvectors are orthogonal. If all n eigenvalues are distinct, then it is clear that the eigenvectors span $\mathcal{V}$ and therefore form a basis.

If $\lambda_1 = \lambda_2$ then we can use the Gram–Schmidt procedure to construct an orthonormal pair from $|v_1\rangle$ and $|v_2\rangle$: the complex spectral theorem (§3.10) guarantees that a Hermitian operator on an n-dimensional space always has n orthogonal eigenvectors and so there must exist appropriate linear combinations of $|v_1\rangle$ and $|v_2\rangle$ that are orthogonal.

Exercise: Show that (i) the eigenvalues of a **unitary** operator are complex numbers of unit modulus and (ii) the eigenvectors of a unitary operator are mutually orthogonal.

How to diagonalise a Hermitian operator Let A be a Hermitian operator with normalised eigenvectors $|v_1\rangle, ..., |v_n\rangle$. When the matrix representing A is expressed in terms of its eigenbasis then the matrix elements are

$$\langle v_i| A |v_j\rangle = \lambda_i \delta_{ij}, \tag{3.71}$$

where λ_i is the eigenvalue corresponding to $|v_i\rangle$. In our standard $|e_1\rangle, ..., |e_n\rangle$ basis, we have that the matrix elements of A are given by

$$\begin{aligned}\langle e_i| A |e_j\rangle &= \langle e_i| IAI |e_j\rangle \\ &= \sum_{k=1}^{n}\sum_{l=1}^{n} \langle e_i|v_k\rangle\langle v_k| A |v_l\rangle\langle v_l|e_j\rangle,\end{aligned} \tag{3.72}$$

resolving the identity (3.5) through $I = \sum_k |v_k\rangle\langle v_k| = \sum_l |v_l\rangle\langle v_l|$. Written as a matrix equation, this is

$$A = U \operatorname{diag}(\lambda_1, ..., \lambda_n) U^\dagger, \tag{3.73}$$

where $U_{ji} = \langle e_j|v_i\rangle$ so that the i^{th} column of U expresses $|v_i\rangle$ in terms of the $|e_1\rangle, ..., |e_n\rangle$ basis.

Exercise: Using $UU^\dagger = U^\dagger U = I$ show the following:

$$\begin{aligned} U^\dagger A U &= \operatorname{diag}(\lambda_1, \ldots, \lambda_n), \\ A^m &= U \operatorname{diag}(\lambda_1^m, \ldots, \lambda_n^m) U^\dagger, \\ \operatorname{tr} A &= \sum_{i=1}^{n} \lambda_i, \\ \det A &= \prod_{i=1}^{n} \lambda_i. \end{aligned} \tag{3.74}$$

Simultaneous diagonalisation of two Hermitian matrices Let A and B be two Hermitian operators. There exists a basis in which A and B are both diagonal if and only if $[A, B] = 0$. Some comments before proving this:

(1) Because the eigenvectors of Hermitian operators are orthogonal, any basis in which such an operator is diagonal must be an eigenbasis.

(2) An equivalent statement is therefore that "A and B both have the same eigenvectors if and only if $[A, B] = 0$."

(3) In this eigenbasis the only difference between A and B is the values of their diagonal elements (i.e., their eigenvalues).

Proof: We first show that if there is basis in which A and B are both diagonal then $[A, B] = 0$. This is obvious: diagonal matrices commute.

The converse is that if $[A, B] = 0$ then there is a basis in which both A and B are diagonal. To prove this, note that because A is Hermitian we can find a basis in which $A = \operatorname{diag}(a_1, \ldots, a_n)$, where the a_i are the eigenvalues of A. In this basis B will be represented by some matrix

$$B = \begin{pmatrix} B_{11} & B_{12} & \ldots & B_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ B_{n1} & B_{n2} & \ldots & B_{nn} \end{pmatrix}. \tag{3.75}$$

The commutator

$$AB - BA = \begin{pmatrix} 0 & (a_1 - a_2)B_{12} & (a_1 - a_3)B_{13} & \ldots \\ (a_2 - a_1)B_{21} & 0 & (a_2 - a_3)B_{23} & \ldots \\ (a_3 - a_1)B_{31} & (a_3 - a_2)B_{32} & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}. \tag{3.76}$$

By assumption $[A, B] = 0$. From the matrix (3.76) we see that this means that $B_{ij} = B_{ji} = 0$ for all indices (i, j) for which $a_i \neq a_j$: if all of A's eigenvalues are distinct then B must be diagonal.

If some of the $\{a_i\}$ aren't distinct we have just a little more work to do. Take, for example, the case $a_1 = a_2$. Then we have that

$$B = B^\dagger = \begin{pmatrix} B_{11} & B_{12} & 0 & 0 & \dots \\ B_{12}^\star & B_{22} & 0 & 0 & \dots \\ 0 & 0 & B_{33} & 0 & \dots \\ 0 & 0 & 0 & B_{44} & \dots \\ \vdots & & & & \end{pmatrix} = \begin{pmatrix} \bar{B}_2 & \\ & \bar{B}_{n-2} \end{pmatrix}, \tag{3.77}$$

using $\bar{B}_2 = \begin{pmatrix} B_{11} & B_{12} \\ B_{12}^\star & B_{22} \end{pmatrix}$ and $\bar{B}_{n-2} = \operatorname{diag}(B_{33}, ..., B_{nn})$ to denote block submatrices of B. Now introduce a unitary change-of-basis matrix

$$U = \begin{pmatrix} U_{11} & U_{12} & 0 & \dots \\ U_{21} & U_{22} & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & & & \end{pmatrix} = \begin{pmatrix} \bar{U}_2 & 0 \\ 0 & \bar{I}_{n-2} \end{pmatrix}. \tag{3.78}$$

Then in this new basis B will be represented by the matrix

$$U^\dagger B U = \begin{pmatrix} \bar{U}_2^\dagger \bar{B}_2 \bar{U}_2 & \\ & \bar{B}_{n-2} \end{pmatrix}. \tag{3.79}$$

Notice that U changes only the upper-left corner of B. We can always choose $\bar{U}_2$ to make $\bar{U}_2^\dagger \bar{B}_2 \bar{U}_2 = \operatorname{diag}(b_1, b_2)$, so that the matrix $U^\dagger B U$ becomes diagonal. The basis change U has no effect on the matrix representing A because $\bar{A}_2 = \operatorname{diag}(a_1, a_2) = a_1 \operatorname{diag}(1, 1)$ is proportional to the identity, we have that

$$U^\dagger A U = \begin{pmatrix} \bar{U}_2^\dagger \bar{A}_2 \bar{U}_2 & \\ & \bar{A}_{n-2} \end{pmatrix} = \begin{pmatrix} \bar{A}_2 & \\ & \bar{A}_{n-2} \end{pmatrix} = A. \tag{3.80}$$

3.12 An application: coupled linear first-order ODEs

Let $\mathbf{x}(t)$ be a vector in an n-dimensional vector space that evolves with time t according to

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathbf{x}(t) = A \mathbf{x}(t), \tag{3.81}$$

where $A(t)$ is a (possibly time-dependent) linear operator on $\mathbf{x}$. The full time evolution of this set of n coupled ODEs is completely determined given the n values $\mathbf{x}(t_0)$ at some initial time t_0. An example of such an equation is the damped harmonic oscillator $\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} + k \frac{\mathrm{d}x}{\mathrm{d}t} + \omega^2 x = 0$: introducing the auxilliary variable $\dot{x}(t)$ and defining $\mathbf{x} = (x, \dot{x})^T$ it can be rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{pmatrix} x(t) \\ \dot{x}(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\omega^2 & k \end{pmatrix} \begin{pmatrix} x(t) \\ \dot{x}(t) \end{pmatrix}, \tag{3.82}$$

in which the coefficients of the 2×2 matrix A do not depend explicitly on time t.

If A is independent of time then the solution to (3.81) is $\mathbf{x}(t) = \exp[tA]\mathbf{x}(0)$. Now consider an initial condition $\mathbf{x}(0)$ and n nearby points displaced by $\Delta\mathbf{x}_1(0), ..., \Delta\mathbf{x}_n(0)$. This set of $n + 1$ points defines an n-dimensional parallelepiped having (oriented) volume $V(0) = \det(\Delta\mathbf{x}_1(0), ..., \Delta\mathbf{x}_n(0))$. From the solution $\mathbf{x}(t) = \exp[tA]\mathbf{x}(0)$ it follows that the (oriented) volume of the parallelepiped evolves as

$$V(t) = \det\left(\exp[tA]\right) = \mathrm{e}^{t \operatorname{tr} A}. \tag{3.83}$$

This is known as **Liouville's formula**. A direct consequence is that if $\operatorname{tr} A = 0$ then the evolution operator $\exp[tA]$ preserves volume. Returning to the harmonic oscillator example, the ODE (3.83) defines a flow on the $(x, \dot{x})$ plane, which preserves area if there is no damping ($k = 0$). Adding damping ($k > 0$) makes the flow contract, in this case towards $(x, \dot{x}) = (0, 0)$.

3.13 Odds and ends

Quadratic forms Expressions such as $ax^2+2bxy+cy^2$, or, more generally,

$$\sum_{i=1}^{n} A_{ii}x_i^2 + 2\sum_{i=1}^{n}\sum_{j<i} A_{ij}x_ix_j = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}x_ix_j \tag{3.84}$$

are known as (homogeneous) quadratic forms. They can be written in matrix form as $\mathbf{x}^\mathrm{T}A\mathbf{x}$, where A is a symmetric matrix with elements $A_{ij}=A_{ji}$. For example,

$$4x^2+6xy+7y^2 = \begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} 4 & 3 \\ 3 & 7 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}. \tag{3.85}$$

> **Exercise:** Explain why under a certain change of basis $\mathbf{x}\to\mathbf{x}'$ the quadratic form (3.84) can be expressed as $\lambda_1 x_1'^2+\cdots+\lambda_n x_n'^2$, where the λ_i are the eigenvalues of A. What is the relationship between $\mathbf{x}$ and $\mathbf{x}'$? Do you need to make any assumptions about the elements A_{ij} or x_i?

Lorentz transformations Our definition of scalar product includes a condition (2.4) that scalar products are positive unless one of the vectors involved is zero. This condition is relaxed in the special theory of relativity, where the scalar product of two four-vectors $\mathbf{x}=(ct,x,y,z)$ and $\bar{\mathbf{x}}=(c\bar{t},\bar{x},\bar{y},\bar{z})$ is defined to be

$$\bar{\mathbf{x}}\cdot\mathbf{x} = \begin{pmatrix} c\bar{t} & \bar{x} & \bar{y} & \bar{z} \end{pmatrix}\begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}\begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \bar{\mathbf{x}}^\mathrm{T}\eta\mathbf{x}, \tag{3.86}$$

where the **metric** $\eta=\mathrm{diag}(1,-1,-1,-1)$. The (square of the) "length" of a spacetime interval $\mathrm{d}\mathbf{x}=(c\mathrm{d}t,\mathrm{d}x,\mathrm{d}y,\mathrm{d}z)$ is then

$$(\mathrm{d}s)^2 = (c\mathrm{d}t)^2-(\mathrm{d}x)^2-(\mathrm{d}y)^2-(\mathrm{d}z)^2, \tag{3.87}$$

which can be positive, negative or zero, depending on whether the interval is time-like, space-like or light-like.

A Lorentz transformation is a change of basis $\mathbf{x}\to\mathbf{x}'$, $\bar{\mathbf{x}}\to\bar{\mathbf{x}}'$ that preserves the scalar product (3.86). An example familiar from the first-year course is $\mathbf{x}'=\Lambda\mathbf{x}$, where

$$\Lambda = \begin{pmatrix} \gamma & -\beta\gamma & & \\ -\beta\gamma & \gamma & & \\ & & 1 & \\ & & & 1 \end{pmatrix}, \tag{3.88}$$

with $\beta=v/c$ and $\gamma=1/\sqrt{1-\beta^2}$. It is easy to confirm that $\Lambda^\mathrm{T}\eta\Lambda=\eta$.

Jordan normal form Not all matrices can be diagonalised, but it turns out that for every matrix A there is an invertible transformation P for which $P^{-1}AP$ takes on the block-diagonal form

$$P^{-1}AP = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_n \end{pmatrix}, \tag{3.89}$$

where the blocks are

$$A_i = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix}, \tag{3.90}$$

having ones immediately above the main diagonal. Each λ_i here is an eigenvalue of A with multiplicity given by the number of diagonal elements in the corresponding A_i.

Singular value decomposition A generalisation of the eigenvector decomposition we have been applying to square matrices holds to any $m \times n$ matrix A: any such A can be factorized as

$$A = UDV^\dagger, \tag{3.91}$$

where U is an $m \times m$ unitary matrix, D is an $m \times n$ diagonal matrix consisting of the so-called singular values of A, and V is an $n \times n$ unitary matrix.

Further reading

See RHB§8, DR§II.

Maths Methods Week 2: Functions as vectors

http://www-thphys.physics.ox.ac.uk/people/JohnMagorrian/mm
john.magorrian@physics.ox.ac.uk

Having dealt with finite-dimensional vector spaces, we now show how certain classes of functions can be treated as members of an infinite-dimensional vector space. Provided we're careful about the sort of function we admit, such function spaces share many of the properties of finite-dimensional vector spaces.

4 Functions as vectors

The set of all functions $f : [a, b] \to \mathbb{C}$ for which the integral

$$\int_a^b |f(x)|^2 w(x)\, \mathrm{d}x \tag{4.1}$$

converges is a vector space under the natural rules of addition of functions and multiplication of functions by scalars. Each such space is defined by the choices of a, b and the function $w(x)$. The **weight function** $w(x)$ measures how densely we think points should be sampled along $[a, b]$. Therefore we impose the condition that $w(x) > 0$ for $x \in (a, b)$. Given two functions $f(x)$ and $g(x)$ from this space, we define their inner product to be

$$\langle f|g\rangle \equiv \langle f, g\rangle = \int_a^b f^\star(x)g(x)w(x)\, \mathrm{d}x, \tag{4.2}$$

which satisfies all of the conditions (2.1–2.4). Similarly, the bra corresponding to the ket $f(x)$ is the linear mapping

$$\langle f| \bullet = \int_a^b \mathrm{d}x\, w(x)f^\star(x) \bullet . \tag{4.3}$$

This inner-product space is sometimes known as $L^2_w(a, b)$. It will be the focus of most of the rest of this course.

In addition to continuous functions on $[a, b]$, this space includes piecewise-continuous functions that undergo finite-sized jumps. For example, the function $f : [-1, 1] \to \mathbb{C}$ defined by

$$f(x) = \begin{cases} 1, & \text{if } |x| < \frac{1}{2}, \\ 0, & \text{otherwise,} \end{cases} \tag{4.4}$$

is continuous except at $x = \pm\frac{1}{2}$. The integral (4.1) converges for $w(x) = 1$ and so this $f(x)$ is a member of the space $L^2_1(-1, 1)$.

Completeness A (possibly) interesting technical point is that we *must* include functions with finite jump discontinuities in the vector space $L^2_w(a, b)$; we cannot restrict our vector space $\mathcal{V}$ to the set of continuous functions. Consider a sequence of functions $|f_k\rangle \in \mathcal{V}$ that satisfies the condition

$$\lim_{k,l\to\infty} D^2(|f_k\rangle, |f_l\rangle) = 0. \tag{4.5}$$

It is not hard to show that this condition means that the sequence converges to some well-defined $|f\rangle$: the sequence $|f_k\rangle$ is said to "converge in the mean" to $|f\rangle$. For example, the sequence $f_k : [-1, 1] \to \mathbb{R}$ of continuous functions defined by

$$f_k(x) = \begin{cases} 0, & x < -\frac{1}{k}, \\ \frac{1}{2}(kx + 1), & -\frac{1}{k} < x < \frac{1}{k}, \\ 1, & x > \frac{1}{k}, \end{cases} \tag{4.6}$$

satisfies the condition (4.5). The limit $f(x) = \lim_{k\to\infty} f_k(x)$ is not continuous, however: it has a jump discontinuity at $x = 0$.

It is natural to require that our function space $\mathcal{V}$ admit all such limiting $|f\rangle$. If all sequences $|f_n\rangle \in \mathcal{V}$ that satisfy the condition (4.5) have limits $|f\rangle$ that themselves are members of $\mathcal{V}$ then $\mathcal{V}$ is said to be **complete**. The example just given shows that the set of all continuous functions is not complete. On the other hand, it can be shown that the space $L^2_w(a, b)$ *is* complete (Riesz–Fischer theorem).

4.1 Generalized Fourier series

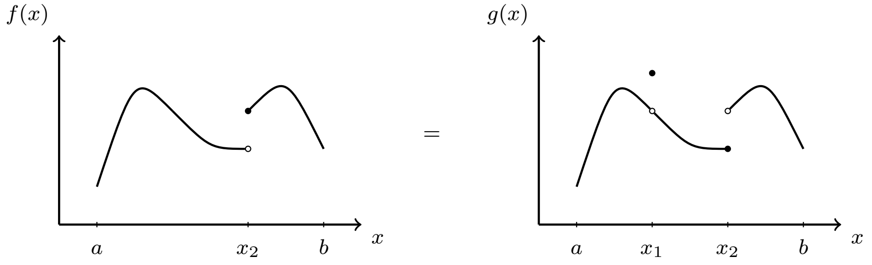
A natural question to ask about $L^2_w(a,b)$ is: what kind of basis "vectors" does it have? Before answering that, however, we need to identify how we expect our basis function expansion should work.

Suppose then that we have some $f(x) \in L^2_w(a,b)$ and we often want to approximate it by some *finite* linear combination $f_n(x) = \sum_{i=0}^n a_i e_i(x)$ of simpler functions $e_i(x)$. The most natural way of quantifying how "good" such an approximation is by using the norm of $f(x) - f_n(x)$ to measure the distance D of $f_n(x)$ from $f(x)$. As for finite-dimensional spaces, this D is given by

$$D^2(f, f_n) \equiv \langle f - f_n, f - f_n \rangle = \int_a^b |f(x) - f_n(x)|^2 w(x)\, \mathrm{d}x. \tag{4.7}$$

If two functions $f(x)$ and $g(x)$ have $D^2(f,g) = 0$ then we say that $f(x)$ and $g(x)$ are **equal (almost everywhere)**.

$D^2(f,g) = 0$ can occur only if $f(x) = g(x)$ for all but a set of isolated points $x_i \in [a,b]$. In the rest of these notes, I abuse notation and take the statements $f = g$ or $|f\rangle = |g\rangle$ to mean that the distance $D^2(f,g) = 0$: any two functions f and g for which $f(x) = g(x)$ everywhere except for a finite set of isolated points $x_i \in [a,b]$ will be "equal" according to this definition. For example, the diagram below shows two piecewise-continuous functions $f(x)$ and $g(x)$ that are identical except at the points $x = x_1$ and $x = x_2$. They are "equal" to one another because the distance $D^2(f,g) = 0$.



Orthonormal bases Now suppose that the functions $e_i(x)$ appearing in the approximations $f_n(x) = \sum_{i=0}^n a_i e_i(x)$ are orthonormal: $\langle e_i, e_j \rangle = \delta_{ij}$. The set $\{e_i(x)\}$ constitutes an orthonormal basis for $L^2_w(a,b)$ if for any $f \in L^2_w(a,b)$ we can find some n and expansion coefficients $a_1, ..., a_n$ that make $D^2(f, f_n)$ arbitrarily small. Then any $f \in L^2_w(a,b)$ can be expanded as

$$\begin{aligned} f(x) &= \lim_{n\to\infty} f_n(x) = \sum_{i=0}^{\infty} a_i e_i(x), \quad \text{(almost everywhere)} \\ \text{with} \quad a_i &= \langle e_i, f \rangle = \int_a^b e_i^\star(x) f(x) w(x)\, \mathrm{d}x. \end{aligned} \tag{4.8}$$

Such an expansion is known as a generalized Fourier series. The coordinates $a_i = \langle e_i, f \rangle$ that give the position of the vector $f(x)$ in the space with respect to the basis $e_1(x), e_2(x), ...$ are often known as **Fourier coefficients**.

Proof that $a_i = \langle e_i, f \rangle$: we have

$$\begin{aligned} D^2(f, f_n) &= \left(\langle f| - \sum_{i=0}^n a_i^\star \langle e_i| \right) \left(|f\rangle - \sum_{j=0}^n a_j\, |e_j\rangle \right) \\ &= \langle f|f\rangle - \sum_{j=0}^n a_j \langle f|e_j\rangle - \sum_{i=0}^n a_i^\star \langle e_i|f\rangle + \sum_{i=0}^n \sum_{j=0}^n a_i^\star a_j \langle e_i|e_j\rangle \\ &= \langle f|f\rangle + \sum_{i=0}^n |a_i - \langle e_i|f\rangle|^2 - \sum_{i=0}^n |\langle e_i|f\rangle|^2, \end{aligned} \tag{4.9}$$

which is clearly minimised by the choice $a_i = \langle e_i|f\rangle$, independent of the value of $n \geq i$. Because the $e_i(x)$ (by assumption) form a basis, the distance $D^2 \to 0$ as $n \to \infty$ and so the "equals" sign in the expansion (4.8) is justified.

Notice from either (4.8) or from (4.9) that the (squared) norm of f is given by

$$\|f\|^2 \equiv \langle f|f\rangle = \sum_{i=0}^{\infty} |a_i|^2, \tag{4.10}$$

where $a_i = \langle e_i|f\rangle$ is the coefficient of the i^{th} term in the expansion (4.8). This result, known as **Parseval's identity**, is essentially Pythagoras' theorem for spaces of functions.

4.2 Basis

Claim: The monomials x^0, x^1, x^2,... are a basis for the space $L^2_w(a,b)$.

Comment: the monomials are countable.

The Weierstrass approximation theorem states that any continuous function $f : [a,b] \to \mathbb{C}$ can be approximated to arbitrary accuracy by a sufficiently high-order polynomial. That is, for any desired accuracy $\epsilon > 0$, there is always a polynomial,

$$g(x) = \sum_{i=0}^{n} a_i x^i, \tag{4.11}$$

of some finite order n for which $D^2(f,g) < \epsilon$. (In general this $n \to \infty$ as $\epsilon \to 0$, but n is finite for any $\epsilon > 0$.) This means that the infinite set of monomials $\{x^0, x^1, x^2, ...\}$ is a basis for the space of continuous functions on the finite interval $[a,b]$ to $\mathbb{C}$: the monomials are LI and in the limit $n \to \infty$ they span the space.

The space $L^2_w(a,b)$ includes piecewise-continuous functions: i.e., functions with a number of finite-sized jump discontinuities. These can be approximated to any desired accurarcy ϵ by continuous functions. Therefore the monomials are a basis for both continuous and piecewise-continuous functions.

4.3 The Gram–Schmidt procedure for functions

We can use the Gram–Schmidt algorithm (§2.1) to construct an orthonormal basis for $L^2_w(a,b)$ from the monomials given choices of a, b and $w(x)$. The procedure is almost the same as for the case of a finite-dimensional vector space; the only difference is that there is now an infinite number of basis elements. As an example, consider the case $(a,b) = (-1,1)$ and $w(x) = 1$. Applying the procedure to the list x^0, x^1, x^2, ... in that order results in:

$$
\begin{aligned}
& e_0'(x) = x^0 \\
& \|e_0'\|^2 = \int_{-1}^{1} |x^0|^2 \,\mathrm{d}x = 2 \\
\Rightarrow\quad & e_0(x) = \frac{1}{\sqrt{2}}. \\
\\
& e_1'(x) = x^1 - \langle e_0|x^1\rangle e_0(x) \\
& \qquad = x - \left[\int_{-1}^{1} x \frac{1}{\sqrt{2}} \,\mathrm{d}x\right] \frac{1}{\sqrt{2}} = x \\
& \|e_1'\|^2 = \int_{-1}^{1} x^2 \,\mathrm{d}x = \frac{2}{3} \\
\Rightarrow\quad & e_1(x) = \sqrt{\frac{3}{2}}x. \\
\\
& e_2'(x) = x^2 - \langle e_0|x^2\rangle e_0(x) - \langle e_1|x^2\rangle e_1(x) \\
& \qquad = x^2 - \left[\int_{-1}^{1} x^2 \sqrt{\frac{3}{2}} x \,\mathrm{d}x\right] \sqrt{\frac{3}{2}} x - \left[\int_{-1}^{1} x^2 \frac{1}{\sqrt{2}} \,\mathrm{d}x\right] \frac{1}{\sqrt{2}} \\
& \qquad = x^2 - \frac{1}{3} \\
& \|e_2'\|^2 = \int_{-1}^{1} \left(x^2 - \frac{1}{3}\right)^2 \,\mathrm{d}x = \frac{8}{45} \\
\Rightarrow\quad & e_2(x) = \sqrt{\frac{5}{8}}(3x^2 - 1). \\
\\
& e_3'(x) = x^3 - \langle e_0|x^3\rangle e_0(x) - \langle e_1|x^3\rangle e_1(x) - \langle e_2|x^3\rangle e_2(x), \quad \text{etc.}
\end{aligned}
\tag{4.12}
$$

The next two elements of this infinite list of orthonormal basis functions turn out to be

$$
e_3(x) = \sqrt{\frac{7}{8}}(5x^3 - 3x), \quad e_4(x) = \frac{3}{8\sqrt{2}}(35x^4 - 30x^2 + 3). \tag{4.13}
$$

These are normalised versions of the **Legendre Polynomials**, which we will encounter later when we discover much easier ways of finding orthogonal bases for any (a, b, w). The first few $e_i(x)$ are plotted on Figure 4-1.

Example: Triangle function Consider the triangular function $f : [-1, 1] \to \mathbb{C}$ defined by

$$
f(x) = 1 - |x|. \tag{4.14}
$$

The Fourier coefficients of this function are given by (4.8)

$$
a_l = \langle e_l|f\rangle = \int_{-1}^{1} e_l^\star(x) f(x) \,\mathrm{d}x. \tag{4.15}
$$

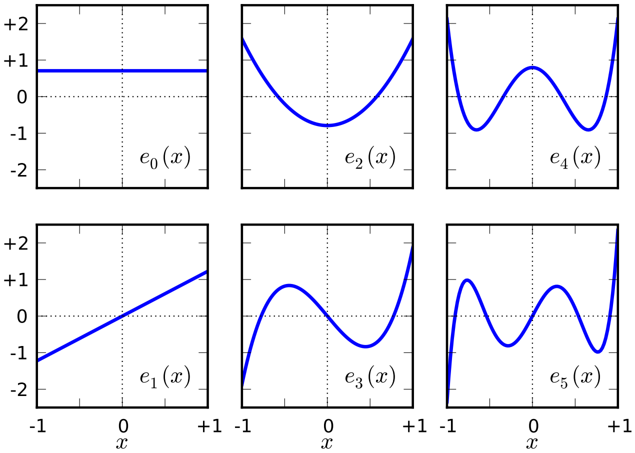


Figure 4-1. The first few orthonormal basis functions for the space $L^2_1(-1,1)$ constructed using the Gram–Schmidt procedure (§4.3).

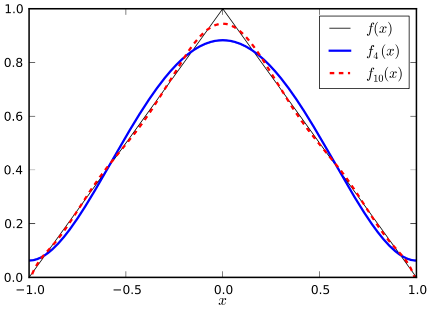


Figure 4-2. Fourier–Legendre approximation to the function $f(x) = 1 - |x|$ using the orthonormal basis (Figure 4-1) constructed in §4.3 and including terms up to order 4 (solid blue curve) and 10 (dashed red curve) in the expansion (4.8).

Notice that $f(x)$ is an even function of x, whereas the odd- (even-) numbered basis functions $e_l(x)$ are odd (even) functions of x. Therefore all odd $a_l = 0$. The first few even a_l are $a_0 = 2^{-1/2}$, $a_2 = -\sqrt{5/32}$, $a_4 = 2^{-7/2}$. Figure 4-2 illustrates the reconstruction of $f(x)$ using the generalized Fourier series (4.8) with these coefficients multiplying the basis functions $e_l(x)$.

4.4 Fourier Series

We have so far been considering functions defined on an interval $[a,b]$ of the real line. A very important special case is that of functions $f : S \to \mathbb{C}$, where S is the unit circle around which we label points by their angular coordinates $\theta \in [-\pi,\pi]$. Such functions are **periodic** with $f(\theta) = f(\theta+2\pi)$. They clearly form a vector space. A basis for this space is

$$e_m(\theta) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{im\theta}, \qquad m \in \mathbb{Z}. \tag{4.16}$$

These $e_m(\theta)$ are orthonormal with $\langle e_m|e_n\rangle = \delta_{mn}$ for weight function $w(\theta)=1$.

DK III§11 presents one way of proving that the $e_m(x)$ defined in (4.16) form a basis. The monomials $x^m y^n$ are a basis for the square $[-1,1]\times[-1,1]$. Therefore any piecewise-continuous $f(x,y)$ can written as $f(x,y) = \sum_{mn} a_{mn} x^m y^n$. Let $x=\cos\theta$, $y=\sin\theta$ be the Cartesian coordinates of points along the unit circle. Then, around this circle, $f(\theta) = \sum_{mn} a_{mn}\cos^m\theta\sin^n\theta$, which can rewritten as a single sum over $e_0(\theta), e_{\pm1}(\theta), \ldots$.

Using (4.8) with the basis (4.16), any piecewise-continuous periodic function $f : [-\pi,\pi] \to \mathbb{C}$ can be expanded as

$$\begin{aligned} f(\theta) &= \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} c_n \mathrm{e}^{in\theta}, \\ \text{with coefficients} \quad c_n &= \langle e_n|f\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \mathrm{e}^{-in\theta} f(\theta)\mathrm{d}\theta. \end{aligned} \tag{4.17}$$

This is the **complex Fourier expansion** of the function f.

It is often more convenient to use real basis functions. Let

$$\begin{aligned} c_n(\theta) &= \begin{cases} e_0(\theta) = \frac{1}{\sqrt{2\pi}}, & n=0, \\ \frac{1}{\sqrt{2}}\left[e_n(\theta)+e_{-n}(\theta)\right] = \frac{1}{\sqrt{\pi}}\cos n\theta, & n = 1,\ 2,\ 3,\ \ldots, \end{cases} \\ s_n(\theta) &= \frac{1}{\sqrt{2}i}\left[e_n(\theta)-e_{-n}(\theta)\right] = \frac{1}{\sqrt{\pi}}\sin n\theta, \quad n = 1,2,3,\ldots \end{aligned} \tag{4.18}$$

be a new basis constructed by taking appropriate linear combination of the $e_m(\theta)$. Rewriting the expansion (4.17) in terms of this new basis, we have

Any piecewise-continuous periodic function $f : [-\pi,\pi] \to \mathbb{C}$ or $\mathbb{R}$ can be expanded as

$$\begin{aligned} f(\theta) &= \frac{a_0}{2} + \sum_{n=1}^{\infty}\left[a_n\cos n\theta + b_n \sin n\theta\right], \qquad (\star) \\ \text{with} \quad a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\cos n\theta\,\mathrm{d}\theta, \\ b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\sin n\theta\,\mathrm{d}\theta. \end{aligned} \tag{4.19}$$

This known as the **Fourier expansion** of the function f.

Comments:

- If $f(\theta)$ is continuous at the point $\theta=\theta_0$ then both expansions converge to $f(\theta_0)$ at $\theta=\theta_0$. If not, they converge to $\frac{1}{2}(f(\theta_0-\epsilon)+f(\theta_0+\epsilon))$.
- The coordinates $\theta=-\pi$ and $\theta=\pi$ correspond to the same point on the circle. Therefore at these points the expansions converge to $\frac{1}{2}(f(-\pi)+f(\pi))$.

- If $f(\theta)$ is an even function, then all $b_n = 0$ in (4.19) and ($\star$) becomes a **Fourier cosine series**. On the other hand, if $f(\theta)$ is odd then all $a_n = 0$ and ($\star$) becomes a **Fourier sine series**.

Substituting $\theta = 2\pi x/L$ in (4.19), we see that any piecewise continuous function $f(x)$ that has period L, so that $f(x+L) = f(x)$, can be expressed as

$$\begin{aligned} f(x) &= \frac{a_0}{2} + \sum_{n=1}^{\infty}\left[a_n \cos\left(\frac{2\pi n}{L}x\right) + b_n \sin\left(\frac{2\pi n}{L}x\right)\right], \\ \text{with}\quad a_n &= \frac{2}{L}\int_{-L/2}^{L/2} f(x)\cos\left(\frac{2\pi n}{L}x\right)\,\mathrm{d}x, \\ b_n &= \frac{2}{L}\int_{-L/2}^{L/2} f(x)\sin\left(\frac{2\pi n}{L}x\right)\,\mathrm{d}x. \end{aligned} \tag{4.20}$$

Additional comments:

- $f(x)$ is assumed to be periodic. Therefore the limits of the integrals for a_n and b_n can be changed from $(-L/2, L/2)$ to, e.g., $(0, L)$.
- Although the derivation of (4.20) comes from considering functions that map from the unit circle to scalars, it can be used to expand any function $f : [-L/2, L/2] \to \mathbb{C}$ or $\mathbb{R}$ that satisfies $f(-L/2) = f(L/2)$. Similarly, it can be used to expand any $f[0, L] \to \mathbb{C}$ or $\mathbb{R}$ that satisfies $f(0) = f(L)$.

Example: Triangle function Let us consider the function (4.14) defined for $x \in [-1, 1]$ as

$$f(x) = 1 - |x|. \tag{4.21}$$

Take $L = 2$ and notice that $f(-L/2) = f(L/2)$. All $b_n = 0$ because $f(x)$ is even. The a_n are given by

$$a_n = 2\int_0^1 (1-x)\cos(n\pi x)\,\mathrm{d}x = \begin{cases} 1, & n = 0, \\ \frac{2}{n^2\pi^2}(1 - \cos n\pi) = \begin{cases} \frac{4}{n^2\pi^2}, & \text{odd } n, \\ 0, & \text{even } n > 0. \end{cases} \end{cases} \tag{4.22}$$

Therefore $f(x)$ can be expressed as the Fourier cosine expansion

$$f(x) = 1 - |x| = \frac{1}{2} + \sum_{k=1}^{\infty} \frac{4}{(2k+1)^2\pi^2}\cos((2k+1)\pi x). \tag{4.23}$$

Figure 4-3 illustrates how well this series converges.

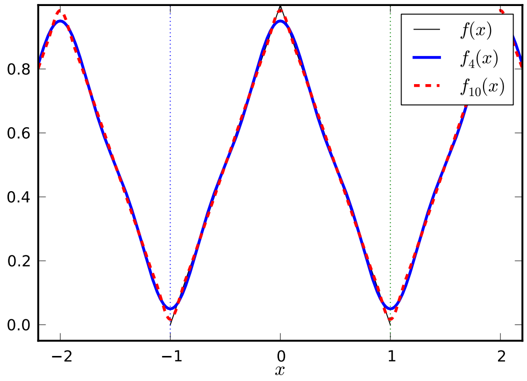


Figure 4-3. Fourier expansion of the function $f(x) = 1 - |x|$ including terms up to 4th (solid blue curve), 10th order (dashed red curve) in the series (4.20) with $L = 2$.

Example: Sawtooth Here is an example of how (generalized) Fourier series expansions behave at discontinuities. The function

$$f(x) = x \tag{4.24}$$

defined on $[-1, 1]$ has Fourier coefficients (4.20) $a_n = 0$ (because the function is odd) and

$$\begin{aligned} b_n = \int_{-1}^{1} x \sin(n\pi x)\,\mathrm{d}x &= \frac{1}{\pi^2}\int_{-\pi}^{\pi} y \sin ny\,\mathrm{d}y \\ &= \frac{1}{\pi^2}\left[\left[-\frac{1}{n} y \cos ny\right]_{-\pi}^{\pi} + \frac{1}{n}\int_{-\pi}^{\pi} \cos ny\,\mathrm{d}y\right] \\ &= \frac{2}{n\pi}(-1)^{n+1}. \end{aligned} \tag{4.25}$$

So, the function $f(x) = x$ defined on $[-1, 1]$ can be represented as the Fourier sine series

$$x = \sum_{n=1}^{\infty} \frac{2}{n\pi}(-1)^{n+1} \sin(n\pi x)\,. \tag{4.26}$$

Figure 4-4 illustrates the convergence of this series. Notice that the truncated series rings at the points $x = \pm 1$: this Fourier expansion is based on the assumption that $f(x)$ is periodic with $f(x) = f(x + 2)$ and so sees a discontinuity in $f(x)$ at the point $x = \pm 1$. This is an example of a **Gibbs phenomenon**.

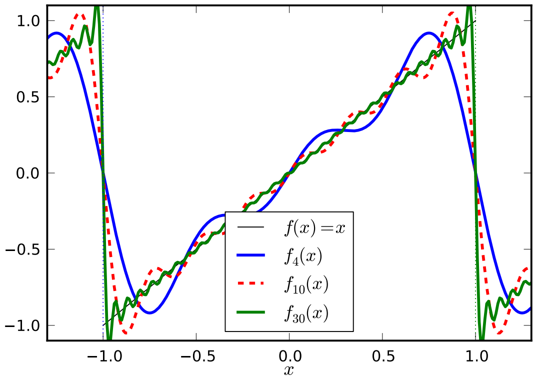


Figure 4-4. Fourier expansion of the function $f(x) = x$ including terms up to 4th (solid blue curve) and 10th (dashed red curve) and 30th order in the series (4.20) with $L = 2$.

Further reading

DK III§§1-9 give a deeper introduction to the ideas I have only outlined here. The classic reference for this material is Courant & Hilbert's *Methods of Mathematical Physics*, Vol I, ch. 2. You might need to brush up on the various notions of sequences and their convergence (e.g., uniform convergence) before you can benefit fully from either of these. None of this is essential reading, but many of you will find it interesting.

More importantly, RHB§12 has plenty of examples of the use of Fourier series.

5 Fourier transforms

Recall that any well-behaved function $f : [-\pi, \pi] \to \mathbb{C}$ can be expanded as (4.17)

$$\begin{aligned} f(x') &= \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} c_n \mathrm{e}^{\mathrm{i}nx'}, \\ \text{with} \quad c_n &= \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}nx'} f(x') \mathrm{d}x'. \end{aligned} \tag{5.1}$$

Let us stretch out the $[-\pi, \pi]$ domain to $[-L/2, L/2]$ by introducing $x = (L/2\pi)x'$ and let us label the Fourier coefficients c_n by the **wavenumber** $k = 2\pi n/L$ instead of n. Then $kx = nx'$ and $\mathrm{d}x' = (2\pi/L)\mathrm{d}x$, so that

$$c_{n(k)} = \frac{2\pi}{L} \frac{1}{\sqrt{2\pi}} \int_{-L/2}^{L/2} \mathrm{e}^{-\mathrm{i}kx} f(x) \mathrm{d}x. \tag{5.2}$$

We define

$$F(k) \equiv \frac{L}{2\pi} c_{n(k)} = \frac{1}{\sqrt{2\pi}} \int_{-L/2}^{L/2} \mathrm{e}^{-\mathrm{i}kx} f(x) \mathrm{d}x, \tag{5.3}$$

in terms of which

$$f(x) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \frac{2\pi}{L} F(k) \mathrm{e}^{\mathrm{i}kx}. \tag{5.4}$$

Note that the spacing between successive values of $k_n = 2\pi n/L$ in the sum (5.4) is $\Delta k = k_{n+1} - k_n = 2\pi/L$.

Taking the limit $L \to \infty$ we define the **Fourier transform** of a function $f : \mathbb{R} \to \mathbb{C}$ as

$$F(k) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx} f(x) \mathrm{d}x. \tag{5.5}$$

Having $F(k)$ we can recover the original $f(x)$ by the taking the **inverse transform**,

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(k) \mathrm{e}^{\mathrm{i}kx} \, \mathrm{d}k. \tag{5.6}$$

5.1 Examples

Example: Sharply truncated sine/cosine wave Consider the function

$$f(x) = \begin{cases} \mathrm{e}^{\mathrm{i}\omega x}, & |x| < l/2, \\ 0, & \text{otherwise.} \end{cases} \tag{5.7}$$

Its Fourier transform is given by

$$\begin{aligned} F(k) &= \frac{1}{\sqrt{2\pi}} \int_{-l/2}^{l/2} \mathrm{e}^{-\mathrm{i}kx} \mathrm{e}^{\mathrm{i}\omega x} \, \mathrm{d}x = \frac{1}{\sqrt{2\pi}} \int_{-l/2}^{l/2} \mathrm{e}^{\mathrm{i}(\omega - k)x} \, \mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{\mathrm{i}(\omega - k)} \left[\mathrm{e}^{\mathrm{i}(\omega-k)l/2} - \mathrm{e}^{-\mathrm{i}(\omega-k)l/2} \right] \\ &= \frac{1}{\sqrt{2\pi}} \frac{2}{\omega - k} \sin\left((\omega - k) \frac{l}{2} \right) \\ &= \frac{l}{\sqrt{2\pi}} \operatorname{sinc}\left((\omega - k) \frac{l}{2} \right). \end{aligned} \tag{5.8}$$

As $l \to \infty$ this tends to a sharp spike $k = \omega$, the area under the spike remaining constant.

Example: Gaussian The normalised Gaussian of dispersion (or standard deviation) a is

$$g(x) = \frac{1}{\sqrt{2\pi}a} \exp\left[-\frac{1}{2}\left(\frac{x}{a}\right)^2\right]. \tag{5.9}$$

Its Fourier transform is given by

$$\begin{aligned} G(k) &= \frac{1}{2\pi a}\int_{-\infty}^{\infty} \exp\left[-\frac{x^2}{2a^2} - ikx\right] \mathrm{d}x \\ &= \frac{1}{2\pi a}\int_{-\infty}^{\infty} \exp\left[-\frac{(x+ika^2)^2}{2a^2} - \frac{k^2a^2}{2}\right] \mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}a} \exp\left[-\frac{k^2a^2}{2}\right] \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp\left[-\frac{(x+ika^2)^2}{2a^2}\right] \mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}a} \exp\left[-\frac{k^2a^2}{2}\right] \frac{1}{\sqrt{2\pi}}\int_{-\infty+\mathrm{i}ka^2}^{\infty+\mathrm{i}ka^2} \exp\left[-\frac{x^2}{2a^2}\right] \mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}a} \exp\left[-\frac{k^2a^2}{2}\right] \underbrace{\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp\left[-\frac{x^2}{2a^2}\right] \mathrm{d}x}_{a} \\ &= \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{k^2a^2}{2}\right]. \end{aligned} \tag{5.10}$$

[See notes for Short Option S1 to understand how the fifth line of (5.10) follows from the fourth.] So, the FT of a Gaussian of dispersion a is $1/a$ times another Gaussian of dispersion $1/a$.

5.2 Properties of Fourier transforms

The following table gives some important relationships between a function $f(x)$ and its Fourier transform $F(k)$:

function	Fourier transform	
$f(ax)$	$\frac{1}{a}F(k/a)$	scale
$f(a+x)$	$\mathrm{e}^{\mathrm{i}ka}F(k)$	phase shift
$\mathrm{e}^{iqx}f(x)$	$F(k-q)$	phase shift
$\frac{\mathrm{d}f}{\mathrm{d}x}$	$\mathrm{i}kF(k)$	derivative
$xf(x)$	$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}k}F(k)$	derivative

To prove the first of these, let $F_a(k)$ be the Fourier transform of $f(ax)$, so that

$$\begin{aligned} F_a(k) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx} f(ax)\,\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}ky/a} f(y)\,\frac{\mathrm{d}y}{a} \\ &= \frac{1}{a}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}(k/a)y} f(y)\,\mathrm{d}y = \frac{1}{a}F(k/a), \end{aligned} \tag{5.11}$$

using the substitution $y = ax$ to go from the first to the second line. Similarly, the second follows on substituting $y = x + a$.

The third follows by replacing k in the definition (5.5) of $F(k)$ by $k-q$.

To prove the second last one, note that the Fourier transform of $\mathrm{d}f/\mathrm{d}x$ is

$$\begin{aligned} \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}kx}\frac{\mathrm{d}f}{\mathrm{d}x}\mathrm{d}x &= \frac{1}{\sqrt{2\pi}}\underbrace{\left[f(x)\mathrm{e}^{-\mathrm{i}kx}\right]_{-\infty}^{\infty}}_{0}+\mathrm{i}k\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}kx}f(x)\,\mathrm{d}x \\ &= \mathrm{i}kF(k), \end{aligned} \tag{5.12}$$

using integration by parts. The final one is left as an exercise.

5.3 Multi-dimensional Fourier transforms

The Fourier transform of a three-dimensional function $f(x,y,z)$ is another function $F(k_x,k_y,k_z)$ obtained by first Fourier transforming in z, then Fourier transforming the result in y and finally Fourier transforming in z. That is,

$$\begin{aligned} F(k_x,k_y,k_z) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}x\,\mathrm{e}^{-\mathrm{i}k_x x}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}y\,\mathrm{e}^{-\mathrm{i}k_y y}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}z\,\mathrm{e}^{-\mathrm{i}k_z z}f(x,y,z) \\ &= \frac{1}{(2\pi)^{3/2}}\int_{-\infty}^{\infty}\mathrm{d}x\int_{-\infty}^{\infty}\mathrm{d}y\int_{-\infty}^{\infty}\mathrm{d}z\mathrm{e}^{-\mathrm{i}(k_x x+k_y y+k_z z)}f(x,y,z). \end{aligned} \tag{5.13}$$

Notice that the order in which the transforms are carried out does not matter. In vector notation, we may write

$$F(\mathbf{k}) = \frac{1}{(2\pi)^{3/2}}\int \mathrm{d}^3\mathbf{x}\,\mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{x}}f(\mathbf{x}), \tag{5.14}$$

where $\mathbf{x}=(x,y,z)$ and $\mathbf{k}=(k_x,k_y,k_z)$. The generalisation to a function $f(\mathbf{x})=f(x_1,...,x_n)$ over an n-dimensional space is obvious:

$$F(\mathbf{k}) = \frac{1}{(2\pi)^{n/2}}\int \mathrm{d}^n\mathbf{x}\,\mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{x}}f(\mathbf{x}); \qquad f(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}}\int \mathrm{d}^n\mathbf{k}\,\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}F(\mathbf{k}). \tag{5.15}$$

5.4 Convolution theorem

The **convolution** of two functions $g(x)$ and $h(x)$ is a new function $f=g*h$ given by

$$f(x)=(g*h)(x)=\int_{-\infty}^{\infty}h(x-x')g(x')\,\mathrm{d}x'. \tag{5.16}$$

That is, f is obtained by "smearing" one function by the other.

Exercise: Show that $g*h=h*g$ and that $f*(\alpha_1 g_1+\alpha_2 g_2)=\alpha_1 f*g_1+\alpha_2 f*g_2$.

Let $G(k)$ and $H(k)$ be the Fourier transforms of the functions $g(x)$ and $h(x)$, respectively. Then the Fourier transform of their convolution $f=g*h=h*g$ is simply

$$F(k)=\sqrt{2\pi}H(k)G(k). \tag{5.17}$$

Proof

$$\begin{aligned} F(k) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}x\,\mathrm{e}^{-\mathrm{i}kx}\int_{-\infty}^{\infty}h(x-x')g(x')\,\mathrm{d}x' \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}x\int_{-\infty}^{\infty}\mathrm{d}x'\,\mathrm{e}^{-\mathrm{i}k(x-x')}\mathrm{e}^{-\mathrm{i}kx'}h(x-x')g(x'). \end{aligned} \tag{5.18}$$

Now change variables to $u = x'$, $v = x - x'$. Then $\mathrm{d}x\mathrm{d}x' = \mathrm{d}u\mathrm{d}v$, which each of u and v spanning the whole real line. So,

$$\begin{aligned} F(k) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}u \int_{-\infty}^{\infty} \mathrm{d}v\, \mathrm{e}^{-\mathrm{i}kv} \mathrm{e}^{-\mathrm{i}ku} h(v) g(u) \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}v\, \mathrm{e}^{-\mathrm{i}kv} h(v) \int_{-\infty}^{\infty} \mathrm{d}u\, \mathrm{e}^{-\mathrm{i}ku} g(u) = \sqrt{2\pi} H(k) G(k). \end{aligned} \tag{5.19}$$

The generalisation to functions on n-dimensional space is straightforard:

$$f(\mathbf{x}) = (g * h)(\mathbf{x}) = (h * g)(\mathbf{x}) \equiv \int h(\mathbf{x} - \mathbf{x}') g(\mathbf{x})\, \mathrm{d}^n \mathbf{x}, \tag{5.20}$$

which has Fourier transform $(2\pi)^{n/2} H(\mathbf{k}) G(\mathbf{k})$.

5.5 Some applications of the Fourier transform

You'll encounter uses of the Fourier transform in both quantum mechanics and optics this year. Here are some brief examples of how it can be applied.

Example: An integral equation Given the functions $f_0(x)$ and $g(x)$, what $f(x)$ satisfies

$$f(x) = f_0(x) + \int_{-\infty}^{\infty} \mathrm{d}y\, g(x-y) f(y) \tag{5.21}$$

subject to the boundary conditions that $f(x)$ vanishes as $|x| \to \infty$? Taking the Fourier transform of (5.21), recognising that the integral on the RHS is a convolution, gives

$$\begin{aligned} & F(k) = F_0(k) + \sqrt{2\pi} G(k) F(k) \\ \Rightarrow\quad & F(k) = \frac{F_0}{1 - \sqrt{2\pi} G}, \end{aligned} \tag{5.22}$$

where $F(K)$, $F_0(k)$, $G(k)$ are the Fourier transforms of $f(x)$, $f_0(x)$ and $g(x)$, respectively. The problem reduces to one of finding the inverse Fourier transform of this $F(k)$.

Example: Poisson's equation In Cartesian coordinates Poisson's equation is

$$\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} = 4\pi G \rho(x, y, z), \tag{5.23}$$

with both Φ and ρ vanishing at infinity. It is easy to show that the (3d) Fourier transform of $\partial f/\partial x$ is $-\mathrm{i}k_x \bar{\Phi}$, where $\bar{\Phi}$ is the 3d Fourier transform of $\Phi(x, y, z)$. So, taking the (3d) Fourier transform of (5.23) we have that

$$-\mathbf{k}^2 \bar{\Phi} = 4\pi G \bar{\rho}, \tag{5.24}$$

where $\bar{\rho}(\mathbf{k})$ is the (3d) Fourier transform of $\rho(\mathbf{x})$. Therefore $\bar{\Phi} = -4\pi G \bar{\rho}/|\mathbf{k}|^2$.

Further reading

See RHB§13.1 for more discussion of Fourier transforms and exercises.

6 Dirac delta

The Dirac delta "function" $\delta(x)$ has the following properties: [†]

$$\begin{aligned} \delta(x) &= 0, \quad x \neq 0 \\ \int_{-\infty}^{\infty} \delta(x)\,\mathrm{d}x &= 1, \\ \int_{-\infty}^{\infty} f(x)\delta(x)\,\mathrm{d}x &= f(0). \end{aligned} \tag{6.1}$$

No such function exists! Nevertheless, we can understand

$$\delta(x) = \lim_{\epsilon\to 0} \delta_\epsilon(x) \tag{6.2}$$

as the limit of a sequence of **kernel functions** $\delta_\epsilon(x)$ that in the limit $\epsilon \to 0$ satisfies $\int_{-\infty}^{\infty} f(x)\delta_\epsilon(x)\,\mathrm{d}x \to f(0)$. Some choices for $\delta_\epsilon(x)$ include:

$$\begin{aligned} \delta_\epsilon(x) =& \frac{1}{\sqrt{2\pi}\epsilon} \exp\left[-\frac{x^2}{2\epsilon^2}\right] \quad \text{(Gaussian)}, \\ \text{or} \quad \delta_\epsilon(x) =& \frac{1}{\pi}\frac{\epsilon}{(x^2+\epsilon^2)} \quad \text{(Cauchy–Lorentz)}. \end{aligned} \tag{6.3}$$

These share the properties that (i) $\int \delta_\epsilon(x)\mathrm{d}x = 1$, (ii) $\frac{\mathrm{d}^k}{\mathrm{d}x^k}\delta_\epsilon$ exists and tends to 0 faster than any power of $1/|x|$ as $x \to \pm\infty$ and (iii) $\delta_\epsilon(x)$ becomes more and more concentrated towards $x = 0$ as $\epsilon \to 0$: that is, for any choice of X and mass $m < 1$ enclosed within $|x| < X$, there is always some $\epsilon_{\max}$ for which, for any $\epsilon < \epsilon_{\max}$, we have that

$$\int_{-X}^{X} \delta_\epsilon(x)\,\mathrm{d}x > m. \tag{6.4}$$

I adopt the first (a normalised Gaussian of dispersion ϵ) in the examples that follow.

> **Exercise:** Show that the Dirac delta is the "identity" element for the convolution operation (5.21). That is, $\delta * f = f * \delta = f$.

6.1 Justification

It is clear that the choice

$$\delta_\epsilon(x) = \frac{1}{\sqrt{2\pi}\epsilon} \exp\left[-\frac{x^2}{2\epsilon^2}\right] \tag{6.5}$$

satisfies the first two of conditions (6.1). Here we show that it satisfies

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)\,\mathrm{d}x = f(a), \tag{6.6}$$

which is a slight generalisation of the final condition of (6.1).

According to our interpretation of the meaning of $\delta(x)$, we need to replace the $\delta(x-a)$ in (6.6) by $\delta_\epsilon(x-a)$ and take the limit as $\epsilon \to 0$ of the whole integral. Doing this, the LHS of (6.6) becomes

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)\,\mathrm{d}x = \lim_{\epsilon\to 0} \int_{-\infty}^{\infty} f(x)\delta_\epsilon(x-a)\,\mathrm{d}x. \tag{6.7}$$

[†] The second property is a special case of the third, but is listed separately to emphasise that the Dirac delta has unit "mass".

Taylor expanding the integrand in the RHS of (6.7) we obtain

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)\delta_\epsilon(x-a)\,\mathrm{d}x &= \int_{-\infty}^{\infty}\left[f(a)+(x-a)f'(a)+\frac{1}{2}f''(a)+\cdots\right]\delta_\epsilon(x-a)\,\mathrm{d}x\\ &= f(a)\int_{-\infty}^{\infty}\delta_\epsilon(x-a)\,\mathrm{d}x\\ &\quad+ f'(a)\int_{-\infty}^{\infty}(x-a)\delta_\epsilon(x-a)\,\mathrm{d}x\\ &\quad+\frac{1}{2}f''(a)\int_{-\infty}^{\infty}(x-a)^2\delta_\epsilon(x-a)\,\mathrm{d}x+\cdots.\end{aligned}\tag{6.8}$$

Substitute $y=x-a$ and note that the integrals in each of the terms of this series are all of the form

$$I_n=\int_{-\infty}^{\infty} y^n\delta_\epsilon(y)\,\mathrm{d}y,\tag{6.9}$$

which is just the n^{th} moment of the Gaussian $\delta_\epsilon(x)$. When n is odd $I_n=0$ because the integrand is an odd function. The first two even moments are easy: $I_0=1$ because our Gaussian δ_ϵ is correctly normalized; $I_2=\epsilon^2$ since I_2 is just the variance of δ_ϵ. More generally, it is easy to see that $I_{2n}\propto\epsilon^{2n}$. (In equation (6.9) use the expression (6.5) for $\delta_\epsilon(y)$ then change variables to $y'=y/\epsilon$.)

Substituting these results into the RHS of equation (6.7) we have finally that

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)\delta(x-a)\,\mathrm{d}x &= \lim_{\epsilon\to 0}\left[f(a)+\frac{1}{2}f''(a)\epsilon^2+\mathcal{O}(\epsilon^4)\right]\\ &= f(a),\end{aligned}\tag{6.10}$$

as required.

6.2 Properties

The Dirac delta has the following properties:

$$f(x)\delta(x-a)=f(a)\delta(x-a).\tag{6.11}$$

$$\delta(ax)=\frac{1}{|a|}\delta(x).\tag{6.12}$$

$$\delta(-x)=\delta(x).\tag{6.13}$$

$$\delta(x^2-a^2)=\frac{1}{2|a|}[\delta(x+a)+\delta(x-a)].\tag{6.14}$$

$$\delta(f(x))=\sum_i\frac{1}{|f'(x_i)|}\delta(x-x_i),\quad\text{where } f(x_i)=0.\tag{6.15}$$

$$\delta'(x)f(x)=-\delta(x)f'(x).\tag{6.16}$$

The simplest way of proving any of these is to multiply both sides by an arbitrary function $g(x)$ and then integrate, using the properties (6.1) together with possible a change of variables to show that both sides are equal.

For example, here is how to prove (6.15) that

$$\delta[f(x)] = \sum_i \frac{\delta(x-x_i)}{|f'(x_i)|}, \tag{6.17}$$

where the x_i are the locations of the zeroes of $f(x)$ (i.e., $f(x_i)=0$). We need to show that

$$\int_{-\infty}^{\infty} g(x)\delta[f(x)]\,\mathrm{d}x = \int_{-\infty}^{\infty} g(x) \sum_i \frac{\delta(x-x_i)}{|f'(x_i)|}\,\mathrm{d}x \tag{6.18}$$

for any well-behaved $g(x)$. The integrand in the LHS is nonzero only for tiny regions around each of the x_i. Therefore we can split the full integral into a sum of smaller integrals around each of these ranges:

$$\int_{-\infty}^{\infty} g(x)\delta[f(x)]\,\mathrm{d}x = \sum_i \int_{x_i-\Delta x}^{x_i+\Delta x} g(x)\delta[f(x)]\,\mathrm{d}x, \tag{6.19}$$

where $\Delta x > 0$ is chosen to be small enough to ensure that each integral includes only one of the zeros of $f(x)$. Each of the integrals in the RHS of (6.19) is of the form

$$\int_a^b g(x)\delta[f(x)]\,\mathrm{d}x, \tag{6.20}$$

in which $a = x_i - \Delta x$ and $b = x_i + \Delta x$. Notice that $a < b$ because $\Delta x > 0$. Substituting $y = f(x)$, we have that $\mathrm{d}y = f'(x)\mathrm{d}x$, and so

$$\int_a^b g(x)\delta[f(x)]\,\mathrm{d}x = \int_{f(a)}^{f(b)} g(x(y))\delta(y)\,\frac{\mathrm{d}y}{f'(x)}, \tag{6.21}$$

where the $x(y)$ that appears in the integrand on the RHS is understood to mean the $x \in [x_i-\Delta x, x_i+\Delta x]$ that solves $y = f(x)$: there will be precisely one such x as long as $f' \neq 0$ and Δx is small enough.

If $f'(x_i) > 0$ we have that $f(b) > f(a)$ and the RHS of (6.21) becomes

$$\int_{f(a)}^{f(b)} g(x)\delta(y)\,\frac{\mathrm{d}y}{f'(x)} = \frac{g(x_i)}{f'(x_i)} = \frac{g(x_i)}{|f'(x_i)|} \qquad (\text{if } f(x_i) > 0). \tag{6.22}$$

On the other hand, if $f'(x_i) < 0$, then $f(b) < f(a)$ and so

$$\int_{f(a)}^{f(b)} g(x)\delta(y)\,\frac{\mathrm{d}y}{f'(x)} = -\int_{f(b)}^{f(a)} g(x)\delta(y)\,\frac{\mathrm{d}y}{f'(x)} = -\frac{g(x_i)}{f'(x_i)} = \frac{g(x_i)}{|f'(x_i)|} \qquad (\text{if } f(x_i) < 0). \tag{6.23}$$

Combining these two results and substituting into equation (6.19) we have that

$$\begin{aligned} \int_{-\infty}^{\infty} g(x)\delta[f(x)]\,\mathrm{d}x &= \sum_i \frac{g(x_i)}{|f'(x_i)|} \\ &= \sum_i \int_{-\infty}^{\infty} g(x) \frac{\delta(x-x_i)}{|f'(x_i)|}\,\mathrm{d}x \\ &= \int_{-\infty}^{\infty} g(x) \sum_i \frac{\delta(x-x_i)}{|f'(x_i)|}\,\mathrm{d}x, \end{aligned} \tag{6.24}$$

which is just the equation (6.18) that we have set out to prove. Since this holds for any $g(x)$ we are justified in claiming (6.15). Notice that we have needed to assume that $f'(x_i) \neq 0$ to obtain this result.

6.3 Multidimensional Dirac delta

Let $\mathbf{r} = (x, y, z)$ and introduce the three-dimensional delta function,

$$\delta^{(3)}(\mathbf{r}) \equiv \delta(x)\delta(y)\delta(z). \tag{6.25}$$

It is easy to show that this satisfies (compare to 6.1)

$$\begin{aligned} \delta(\mathbf{r}) &= 0, \quad \mathbf{x} \neq 0 \\ \int \delta^{(3)}(\mathbf{r})\,\mathrm{d}^3\mathbf{r} &= 1 \\ \int f(\mathbf{r})\delta^{(3)}(\mathbf{r}-\mathbf{r}_0)\,\mathrm{d}^3\mathbf{r} &= \int \mathrm{d}x \int \mathrm{d}y \int \mathrm{d}z f(x,y,z)\delta(x-x_0)\delta(y-y_0)\delta(z-z_0) \\ &= f(x_0, y_0, z_0) = f(\mathbf{r}_0). \end{aligned} \tag{6.26}$$

Similarly $\delta^{(n)}(x_1, ..., x_n) \equiv \delta(x_1)\cdots\delta(x_n)$.

6.4 Fourier-space representation of the Dirac delta

Let us adopt the Gaussian form (6.5) for $\delta_\epsilon(x)$. From (5.10) its Fourier transform is

$$\Delta_\epsilon(k) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}\epsilon^2 k^2}. \tag{6.27}$$

Inverting this, we have that

$$\delta_\epsilon(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}k\,\mathrm{e}^{\mathrm{i}kx}e^{-\frac{1}{2}\epsilon^2 k^2}, \tag{6.28}$$

so that, taking $\epsilon \to 0$ and remembering that k is a dummy variable,

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}k\,\mathrm{e}^{\mathrm{i}kx} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}k\,\mathrm{e}^{-\mathrm{i}kx}. \tag{6.29}$$

This representation of the Dirac delta looks strange, but let us check that it works:

$$\begin{aligned} \int_{-\infty}^{\infty}\mathrm{d}x\,f(x)\delta(x-a) &= \int_{-\infty}^{\infty}\mathrm{d}x\,f(x)\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}k\,\mathrm{e}^{-\mathrm{i}k(x-a)} \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}k\,\mathrm{e}^{\mathrm{i}ka}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}x\,\mathrm{e}^{-\mathrm{i}kx}f(x) \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}k\,\mathrm{e}^{\mathrm{i}ka}F(k) \\ &= f(a), \end{aligned} \tag{6.30}$$

where $F(k)$ is the Fourier transform of $f(x)$. So, we can represent the Dirac delta as a superposition of plane waves in k space, $\mathrm{e}^{\mathrm{i}kx}$, with a uniform distribution of amplitudes.

6.5 A basis for Fourier space

The complex Fourier series for a function $f:[-L/2,L/2]\to\mathbb{C}$ uses the discrete orthonormal basis

$$e_n(x)=\frac{1}{\sqrt{L}}\mathrm{e}^{\mathrm{i}n\pi x/L},\quad n\in\mathbb{Z}. \tag{6.31}$$

Recall that we constructed the Fourier transform of functions $f:\mathbb{R}\to\mathbb{C}$ by replacing the discrete index n by $k=2\pi n/L$ and letting $L\to\infty$ so that k becomes continuous. This procedure works, but can we identify a basis for this space? Notice that the individual $e_n(x)\to 0$ as $L\to\infty$, which means that they are no longer useful.

The results of the previous subsection suggest a basis

$$e_k(x)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{\mathrm{i}kx}, \tag{6.32}$$

because then

$$\langle e_k|e_p\rangle=\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}x\,\mathrm{e}^{\mathrm{i}(p-k)x}=\lim_{\epsilon\to 0}\frac{1}{2\pi}\int_{-1/\epsilon}^{1/\epsilon}\mathrm{d}x\,\mathrm{e}^{\mathrm{i}(p-k)x}=\delta(p-k), \tag{6.33}$$

using the representation (6.29) for $\delta(p-k)$. This relation replaces the orthonormality relation $\langle e_k|e_p\rangle=\delta_{kp}$ when k and p are continuous.

6.6 Resolution of the identity, revisited

The following section introduces some formal notation that can be useful, particularly in quantum mechanics. We can think of $f(x_0)$ as being the projection of the vector $|f\rangle$ along the (generalized) function $\delta(x-x_0)$,

$$\begin{aligned} f(x_0) &= \langle x_0|f\rangle \\ &= \int_{-\infty}^{\infty}\mathrm{d}x\,\delta(x-x_0)f(x). \end{aligned} \tag{6.34}$$

Claim: For the present case of functions $f:\mathbb{R}\to\mathbb{C}$ with unit weight function $w(x)=1$, we can *formally* write the identity operator as

$$I=\int_{-\infty}^{\infty}\mathrm{d}x\,|x\rangle\langle x|\,. \tag{6.35}$$

Proof: For any two vectors $|f\rangle$, $|g\rangle$ we have that $\langle f|x\rangle=\langle x|f\rangle^\star=f^\star(x)$ and $\langle x|g\rangle=g(x)$. Therefore

$$\begin{aligned} \langle f|\,I\,|g\rangle &= \langle f|\left[\int_{-\infty}^{\infty}\mathrm{d}x\,|x\rangle\langle x|\right]|g\rangle \\ &= \int_{-\infty}^{\infty}\mathrm{d}x\,\langle f|x\rangle\langle x|g\rangle \\ &= \int_{-\infty}^{\infty}\mathrm{d}x\,f^\star(x)g(x) \\ &= \langle f|g\rangle. \end{aligned} \tag{6.36}$$

So, with the understanding that $\langle x|f\rangle$ means $f(x)$ and $\langle f|x\rangle$ means $f^\star(x)$, we can slip (6.35) into any scalar product such as $\langle f|g\rangle=\langle f|\,I\,|g\rangle$ and obtain a formal expression for it in terms of the "x-basis" for the objects $|f\rangle$ and $|g\rangle$.

Similarly, we can think of the value of the Fourier transform $F(k_0)$ at $k = k_0$ as being equal to the projection of the function $|f\rangle$ along the basis vector $|e_{k_0}\rangle$ given by (6.32):

$$\begin{aligned} F(k_0) &= \langle k_0|f\rangle = \langle e_{k_0}|f\rangle \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{e}^{-\mathrm{i}k_0 x}\, f(x). \end{aligned} \tag{6.37}$$

Claim: We can also express the identity as

$$I = \int_{-\infty}^{\infty} \mathrm{d}k\, |e_k\rangle\langle e_k|\,. \tag{6.38}$$

Proof: for any two $|f\rangle$, $|g\rangle$ we have that

$$\begin{aligned} \langle f|\,I\,|g\rangle &= \int_{-\infty}^{\infty} \mathrm{d}k\, \langle f|e_k\rangle\langle e_k|g\rangle \\ &= \int_{-\infty}^{\infty} \mathrm{d}k\, \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{e}^{\mathrm{i}kx} f^\star(x)\, \mathrm{d}x \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}x'\, \mathrm{e}^{-\mathrm{i}kx'} g(x')\, \mathrm{d}x' \\ &= \int_{-\infty}^{\infty} \mathrm{d}x f^\star(x) \int_{-\infty}^{\infty} \mathrm{d}x'\, g(x') \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}k\, \mathrm{e}^{\mathrm{i}k(x-x')} \\ &= \int_{-\infty}^{\infty} \mathrm{d}x\, f^\star(x) \int_{-\infty}^{\infty} \mathrm{d}x'\, g(x')\delta(x-x') \\ &= \int_{-\infty}^{\infty} \mathrm{d}x\, f^\star(x) g(x) = \langle f|g\rangle, \end{aligned} \tag{6.39}$$

using (6.29) to go from the third to the fourth line and remembering that $\int_{-\infty}^{\infty}$ really means $\lim_{l\to 0}\int_{-l}^{l}$, or, equivalently, $\lim_{\epsilon\to 0}\int_{-1/\epsilon}^{1/\epsilon}$.

As a sanity check, notice that

$$e_k(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{\mathrm{i}kx} = \langle x|e_k\rangle \tag{6.40}$$

and therefore the Fourier transform

$$F(k) = \langle e_k|f\rangle = \langle e_k|\,I\,|f\rangle = \int_{-\infty}^{\infty} \mathrm{d}x \langle e_k|x\rangle\langle x|f\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{e}^{-\mathrm{i}kx} f(x). \tag{6.41}$$

Similarly,

$$f(x) = \langle x|f\rangle = \langle x|\,I\,|f\rangle = \int_{-\infty}^{\infty} \mathrm{d}k \langle x|e_k\rangle\langle e_k|f\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}k\, \mathrm{e}^{\mathrm{i}kx} F(k). \tag{6.42}$$

[Notice that there is an ambiguity in expressions such as $\langle x|e_k\rangle$. Does it mean the projection of the function $|e_k\rangle$ onto the x basis (i.e., what we'd normally call $e_k(x)$)? Or does it mean the projection of $e_k(x)$ onto the function $f(x) = x$? Here, of course, it means the former. Usually the meaning is clear from the context.]

Exercise: Show that

$$\int_{-\infty}^{\infty} f^\star(x) g(x) \mathrm{d}x = \int_{-\infty}^{\infty} \tilde{f}^\star(k) \tilde{g}(k) \mathrm{d}k, \tag{6.43}$$

where $\tilde{f}(k) = \langle k|f\rangle$ and $\tilde{g}(k) = \langle k|g\rangle$ are the Fourier transforms of $f(x)$ and $g(x)$ respectively. This result is known as **Parseval's formula**. It shows that the Fourier transform preserves the inner product: it is unitary.

Exercise: Let $|e_n\rangle$ with $n \in \mathbb{Z}$ be a (discrete) orthonormal basis for the space $L^2_w(a,b)$. Explain why the identity operator can be expressed as

$$I = \sum_n |e_n\rangle\langle e_n| \tag{6.44}$$

and also, applying the formal notation introduced in this subsection, as

$$I = \int_a^b \mathrm{d}x\, w(x)\, |x\rangle\langle x|\,. \tag{6.45}$$

Hence show that

$$\begin{aligned}\frac{1}{\sqrt{w(x)w(x_0)}}\delta(x-x_0) &= \sum_n \langle x|e_n\rangle\langle e_n|x_0\rangle \\ &= \sum_n e_n(x)e_n^\star(x_0).\end{aligned} \tag{6.46}$$

Further reading

DK III§13 explains how the Dirac delta can be viewed as an example of a *generalized function* or *distribution*; see also §21 of Kolmogorov & Fomin's *Introductory real analysis*. DK III§14.7 gives a fuller justification for the formal $f(x) = \langle x|f\rangle$ etc notation just introduced.

Maths Methods Week 3: linear operators on functions

http://www-thphys.physics.ox.ac.uk/people/JohnMagorrian/mm
john.magorrian@physics.ox.ac.uk

7 Linear differential operators

We have seen how functions from the space $L^2_w(a,b)$ share many of the familiar properties of vectors from finite-dimensional spaces. Now we turn to the properties of certain special linear operators on such functions. Recall that an operator A is linear if $A(u+v) = Au + Av$ and $A(\alpha u) = \alpha Au$, where $u(x)$ and $v(x)$ are any members of the class of function that A admits and α is any scalar constant. If A and B are two such operators then $A+B$, αA and the compositions AB and BA are also linear operators.

Most of undergraduate physics involves second-order linear differential equations of the form $Au = g$, where $u(x)$ is some unknown function and A is a linear operator of the form

$$A\bullet = a_2(x)\frac{\mathrm{d}^2\bullet}{\mathrm{d}x^2} + a_1(x)\frac{\mathrm{d}\bullet}{\mathrm{d}x} + a_0(x)\bullet, \tag{7.1}$$

in which $a_0(x)$, $a_1(x)$ and $a_2(x)$ are smooth, real-valued functions. We focus on such operators for the rest of the course, but not without noting that there are interesting linear operators that are not of this form. For, example:

- the multiplication operator L_a defined by $(L_a f)(x) \equiv a(x)f(x)$ for some given $a(x)$;
- the integration operator L_k defined by $(L_k f)(x) \equiv \int_a^b k(x,x')f(x')\mathrm{d}x'$ for some choice of kernel function $k(x,x')$.

The convolution operator $g \star \bullet$ in (5.16) is an example of the latter with $k(x,x') = g(x-x')$, provided the domain of g extends sufficiently far beyond the interval $[a,b]$.

After reviewing some of the basic properties of ODEs $Au = g$ with A given by (7.1), we first specify the class of boundary conditions we which to apply to our solutions $u(x)$. Then we construct the adjoint operator $A^\dagger$ and obtain the conditions under which $A = A^\dagger$, making A Hermitian. We show that many of the ordinary differential equations encountered in physics problems can be cast as eigenvalue equations, $Au = \lambda u$, where A is an operator of the form (7.1) that is also Hermitian. We discuss various properties of the soutions to such equations (i.e., the eigenfunctions of A), before showing various ways of obtaining explicit expressions for them.

7.1 Recap: second-order ODEs

Recall that the second-order ODE $Au = g$ can be turned into a pair of coupled first-order ODEs by introducing an auxiliary function $u'(x)$ (which will just happen to be equal to $\mathrm{d}u/\mathrm{d}x$) and rewriting $Au = g$ as

$$\begin{pmatrix} a_0(x) & a_1(x) + a_2(x)\frac{\mathrm{d}}{\mathrm{d}x} \\ \frac{\mathrm{d}}{\mathrm{d}x} & 0 \end{pmatrix}\begin{pmatrix} u(x) \\ u'(x) \end{pmatrix} = \begin{pmatrix} g(x) \\ u'(x) \end{pmatrix}. \tag{7.2}$$

This is a first-order ODE for the vector $\begin{pmatrix} u & u' \end{pmatrix}^{\mathrm{T}}$.

Once we've specified the value $\begin{pmatrix} u_0 & u'_0 \end{pmatrix}^{\mathrm{T}}$ of the vector $\begin{pmatrix} u(x) & u'(x) \end{pmatrix}^{\mathrm{T}}$ at some initial location $x = x_0$, we can integrate equation (7.2) forwards or backwards along the x axis to construct $u(x)$. Thus there are in general two LI solutions $u_1(x)$ and $u_2(x)$ to the equation (7.2), which are parametrised by the values chosen for $u_0 = u(x_0)$ and $u'_0 = u'(x_0)$.

As we march along the x axis, we might encounter points $x = x_0$ at which either $a_1(x)/a_2(x) \to \infty$ or $a_0(x)/a_2(x) \to \infty$ as $x \to x_0$. The most obvious example is any point x_0 at which $a_2(x)$ vanishes. These are known as **singular points** and are important in classifying the solutions to the ODE. For this course though we can simply ignore them, even if we do encounter one.

The solution to the inhomogeneous equation $Au = g$ is given by $u(x) = \alpha_1 u_1(x) + \alpha_2 u_2(x) + u_{\rm p}(x)$, where $u_1(x)$ and $u_2(x)$ are the LI solutions to the homogeneous problem $Au = 0$ and the **particular integral** $u_{\rm p}(x)$ is *any* solution to $Au = g$.

Exercise: Write down the condition for a pair of functions $u_1(x)$ and $u_2(x)$ to be LI. Use this to show that $u_1(x)$ and $u_2(x)$ are linearly dependent if the **Wronskian** determinant

$$W(u_1, u_2) = \det\begin{pmatrix} u_1 & u_2 \\ u_1' & u_2' \end{pmatrix} \tag{7.3}$$

is identically zero.

Exercise: Suppose we've found one solution, $u_1(x)$, to the homogeneous equation $Au = 0$. The method of **variation of parameters** is a standard way of finding a second LI solution, $u_2(x)$. Let $u_2(x) = h(x)u_1(x)$. Show that this $h(x)$ is given by

$$\frac{{\rm d}h}{{\rm d}x} = \frac{\text{constant}}{u_1^2(x)} \exp\left[-\int_a^x \frac{c_1(x')}{c_2(x')}{\rm d}x'\right], \tag{7.4}$$

and show that the Wronskian $W(u_1, u_2) = u_1^2 h(x)$.

This method of variation of parameters can also be used to find a particular integral $u_{\rm p}(x)$ for the inhomogeneous equation second-order differential equation $Au = g$. A much more powerful way of tackling inhomogeneous problems is by using the method of Green's functions (§9 below).

7.2 The operator and its diet

Here we are going to consider operators of the form

$$A\bullet = \frac{1}{w(x)}\left[c_2(x)\frac{{\rm d}^2\bullet}{{\rm d}x^2} + c_1(x)\frac{{\rm d}\bullet}{{\rm d}x} + c_0(x)\bullet\right], \tag{7.5}$$

in which the $c_i(x)$ are smooth **real** functions $c_i : [a, b] \to \mathbb{R}$ with $c_2(x) \neq 0$ for any $x \in (a, b)$ and $w(x)$ is the weight function that appears in our definition (4.2) of inner product. Notice that this is identical to (7.1) before apart from the $1/w(x)$ factor, which we absorb into A for later convenience.

We restrict the diet of functions $u(x)$ that this operator is allowed to eat. In particular, we impose the condition that these $u(x)$ can be

- any sufficiently smooth function drawn from $L^2_w(a, b)$ that
- satisfy a pair of homogeneous boundary conditions of the form

$$\begin{aligned} \alpha_{11}u(a) + \alpha_{12}u'(a) + \beta_{11}u(b) + \beta_{12}u'(b) &= 0, \\ \alpha_{21}u(a) + \alpha_{22}u'(a) + \beta_{21}u(b) + \beta_{22}u'(b) &= 0, \end{aligned} \tag{7.6}$$

in which the constants $(\alpha_{ij}, \beta_{ij})$ are chosen, by us, beforehand.

An example of such boundary conditions is the pair $u(a) = u(b)$ and $u'(a) = u'(b)$. Another is the pair $u(a) = u(b) = 0$.

These boundary and differentiability constraints mean that our domain of allowed functions $u(x)$ is a restrcited subset of $L^2_w(a, b)$. Nevertheless, these $u(x)$ are dense in the space $L^2_w(a, b)$: even though $L^2_w(a, b)$ contains piecewise continuous functions that A would choke on, we can always approximate such indigestible functions to arbitrary accuracy by functions $u(x)$ that satisfy the constraints above (see §4.2).

7.3 The adjoint operator

Recall from §3.3 that the dual $A^\dagger$ to an operator A is defined by requiring that $\langle v, Au\rangle = \langle A^\dagger v, u\rangle$ for all admissible choices of $u(x)$ and $v(x)$. Notice that A acts only on the function u, which we assume satisies the homogeneous bcs (7.6). The adjoint $A^\dagger$ acts only on v. For now let us keep an open mind about the bcs satisfied the objects v that $A^\dagger$ operates on.

Using the definition (4.2) of inner product, the condition $\langle v, Au\rangle = \langle A^\dagger v, u\rangle$ becomes

$$\int_a^b \mathrm{d}x\, v^\star \left[c_2\frac{\mathrm{d}^2u}{\mathrm{d}x^2} + c_1\frac{\mathrm{d}u}{\mathrm{d}x} + c_0 u\right] = \int_a^b \mathrm{d}x\, w\, \left[A^\dagger v\right]^\star u. \tag{7.7}$$

The LHS is a sum of three terms, two of which involve derivatives of u, whereas on the RHS u appears as only a simple multiplicative factor in the integrand. Let us integrate the LHS by parts to remove the derivatives of u. We have that

$$\begin{aligned}\int_a^b \mathrm{d}x\, v^\star c_2 u'' &= \left[v^\star c_2 u'\right]_a^b - \int_a^b \mathrm{d}x\,(v^\star c_2)' u' \\ &= \left[v^\star c_2 u'\right]_a^b - \left[(v^\star c_2)' u\right]_a^b + \int_a^b \mathrm{d}x\,(v^\star c_2)'' u, \\ \int_a^b \mathrm{d}x\, v^\star c_1 u' &= \left[v^\star c_1 u\right]_a^b - \int_a^b \mathrm{d}x\,(v^\star c_1)' u,\end{aligned} \tag{7.8}$$

so that

$$\text{LHS} = \left[v^\star c_2 u' - (v^\star c_2)' u + v^\star c_1 u\right]_a^b + \int_a^b \mathrm{d}x\, u\left[(c_2 v^\star)'' - (c_1 v^\star)' + c_0 v^\star\right]. \tag{7.9}$$

Comparing this to the RHS, we see that, provided we impose the condition that the boundary term vanish, namely

$$\left[v^\star c_2 u' - (v^\star c_2)' u + v^\star c_1 u\right]_a^b = 0, \tag{7.10}$$

then the adjoint $A^\dagger$ is given by the linear differential operator

$$\begin{aligned}A^\dagger\bullet &= \frac{1}{w(x)}\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2}(c_2\bullet) - \frac{\mathrm{d}}{\mathrm{d}x}(c_1\bullet) + c_0\bullet\right] \\ &= \frac{1}{w(x)}\left[c_2\frac{\mathrm{d}^2}{\mathrm{d}x^2} + (2c_2' - c_1)\frac{\mathrm{d}}{\mathrm{d}x} + (c_2'' - c_1' + c_0)\right]\bullet,\end{aligned} \tag{7.11}$$

The messy condition (7.10) sets the bcs that $v(x)$ must satisfy: the pair of homogeneous bcs (7.6) on $u(x)$ mean that we can eliminate two of the four quantities $u(a)$, $u(b)$, $u'(a)$ and $u'(b)$ from the it. The coefficients in front of the two remaining quantities will have the form

$$\alpha_{i1}v^\star(a) + \alpha_{i2}v'^\star(a) + \beta_{i1}v^\star(b) + \beta_{i2}v'^\star(b). \tag{7.12}$$

for $i = 3, 4$. Setting these coefficients to zero gives two homogenenous BCs for the function $v(x)$. The resulting conditions are the **adjoint boundary conditions**. This is easier to see with a concrete example.

Example: Adjoint of the simplest possible linear differential operator

Take $[a, b] = [0, 1]$ with weight function $w(x) = 1$. Consider the operator $A = \frac{\mathrm{d}}{\mathrm{d}x}$, so that $c_0 = c_2 = 0$, $c_1 = 1$. We restrict this A to functions $u(x)$ that satisfy the bc $u(1) = 0$. Then equation (7.11) gives $A^\dagger = -\frac{\mathrm{d}}{\mathrm{d}x}$, while the boundary term (7.10) becomes $v(1)u^\star(1) - v(0)u^\star(0) = 0$. Therefore the adjoint bc is $v(0) = 0$.

7.4 Hermitian operators

An operator A is Hermitian if (i) $A = A^\dagger$ and (ii) A and $A^\dagger$ admit the same functions. Comparing (7.5) and (7.11) we see that $A = A^\dagger$ if $c_1 = c_2'$ Setting $c_1 = c_2'$, the condition (7.10) becomes

$$[c_2(v^\star u' - (v^\star)' u)]_a^b = 0, \tag{7.13}$$

which must also hold with $u(x)$ and $v(x)$ swapped. Both $u(x)$ and $v(x)$ should also satisfy the homogeneous bcs (7.6), but we are free to choose those as we see fit. In contrast, the condition (7.10) is non-negotiable.

Putting this together, and, making one final rewrite of the operator A, we have the following.

The second-order differential operator

$$A_{\mathrm{SL}}\bullet = \frac{1}{w(x)}\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}\bullet}{\mathrm{d}x}\right) + q(x)\bullet\right], \tag{7.14}$$

is Hermitian provided it is restricted to functions $u(x)$, $v(x)$ that satisfy the boundary condition

$$\left[v^\star p\frac{\mathrm{d}u}{\mathrm{d}x}\right]_a^b = 0. \tag{7.15}$$

Here $p(x)$ and $q(x)$ are smooth, real functions with $p(x) \neq 0$ for $x \in (a, b)$.

The “SL” subscript stands for “**Sturm–Liouville**”, after the pair who in the 1830s worked out most the theory of this kind of Hermitian operator. The simplified boundary condition (7.15) suffices to satisfy (7.13).

Exercise: Show that any differential operator

$$A = a_2(x)\frac{d^2}{\mathrm{d}x^2} + a_1(x)\frac{\mathrm{d}}{\mathrm{d}x} + a_0(x) \tag{7.16}$$

can be written in Sturm–Liouville form (7.14) by choosing

$$\begin{aligned} p(x) &= \exp\left[\int_a^x \frac{a_1(t)}{a_2(t)}\mathrm{d}t\right], \\ w(x) &= \frac{p(x)}{a_2(x)} = \frac{1}{a_2(x)}\exp\left[\int_a^x \frac{a_1(t)}{a_2(t)}\mathrm{d}t\right], \\ q(x) &= a_0(x)w(x). \end{aligned} \tag{7.17}$$

Therefore any operator of the form (7.1) can be taken to be Hermitian *provided the boundary conditions (7.15) are satisfied.* The condition that the $c_i(x)$ are real with $c_2(x) > 0$ ensures that $w(x) > 0$ for $x \in (a, b)$. What must we do to ensure $w(x) > 0$ if $c_2(x) < 0$ for $x \in (a, b)$?

7.5 Eigenfunctions

The eigenvalue equation,

$$A_{\mathrm{SL}} e_n(x) = \lambda_n e_n(x), \tag{7.18}$$

in which A_{SL} is given by (7.14) and the eigenfunctions are required to satisfy the SL boundary conditions (7.15), is known as a **Sturm–Liouville equation**.

NB: Many books omit the $1/w(x)$ factor in the definition of the operators (7.5) and (7.14), only to have the weight function reappear in the Sturm–Liouville equation, which becomes $A_{\mathrm{SL}} e_n = \lambda_n w(x) e_n(x)$.

Some eigenproperties of $A_{\rm SL}$:

(i) The eigenvalues λ_n are real.

(ii) The eigenfunctions $e_i(x)$, $e_j(x)$ corresponding to different eigenvalues λ_i, λ_j are orthogonal:

$$\langle e_i|e_j\rangle \equiv \int_a^b e_i^\star(x) e_j(x) w(x)\,{\rm d}x = 0. \tag{7.19}$$

(iii) The eigenfunctions are a complete set (basis): any function $f \in L^2_w(a,b)$ can be expressed as the generalised Fourier series

$$\begin{aligned} f(x) &= \sum_i a_i e_i(x), \\ \text{with}\quad a_i &= \langle e_i|f\rangle = \int_a^b e_i^\star(x) f(x) w(x)\,{\rm d}x, \end{aligned} \tag{7.20}$$

assuming the eigenfunctions $e_i(x)$ have been normalised so that $\langle e_i|e_i\rangle = 1$. If the weight of tradition means that $\langle e_i|e_i\rangle \neq 1$, we need to divide the RHS of the expression for a_i in (7.20) by $\langle e_i|e_i\rangle$.

By construction A_{SL} satisfies the condition $\langle u|\, A_{\rm SL}\, |v\rangle = \langle v|\, A_{\rm SL}\, |u\rangle^\star$, which means that the proofs of (i) and (ii) are identical to the corresponding proofs for finite-dimensional Hermitian operators in §3.11. The proof of (iii) is far too intricate for us, but you should remember the result and be prepared to invoke it fearlessly.

An aside: The analysis of the properties of the eigenfunctions of $A_{\rm SL}$ is made easier if some additional conditions are imposed, namely:

(i) the interval $[a,b]$ is finite;

(ii) neither $p(x)$ nor $w(x)$ vanish at $x=a$ or $x=b$;

(iii) the bcs are of the restricted form

$$\begin{aligned} \alpha_0 u(a) + \alpha_1 u'(a) &= 0, \\ \beta_0 u(b) + \beta_1 u'(b) &= 0, \end{aligned} \tag{7.21}$$

where at least one of α_0 and α_1 is nonzero and similarly for β_0 and β_1.

If these conditions are imposed then the SL problem is said to be *regular*. Eigenfunctions of regular SL problems can be shown to have lots of nice properties, such as:

(iv) there is only one LI eigenfunction for each eigenvalue: the eigenvalues are singly degenerate.

(v) there is an infinite number of eigenvalues, which, if arranged in order $\lambda_j < \lambda_{j+1}$, have $\lambda_j \to \infty$ as $j \to \infty$ (that is, only a finite number of λ_j are negative).

(vi) the $j^{\rm th}$ eigenfunction has exactly j zeros on $[a,b]$.

Many of these properties happen to be shared by eigenfunctions of particular SL problems that aren't regular.

7.6 Examples

Fourier series A simple but important example of a Sturm–Liouville equation is

$$\frac{{\rm d}^2 e_n}{{\rm d}x^2} = \lambda_n e_n. \tag{7.22}$$

Let us take the domain of the $e_n(x)$ to be $[a,b] = [-\pi,\pi]$ and impose the pair of boundary conditions $e_n(-\pi) = e_n(\pi)$, $e_n'(-\pi) = e_n'(\pi)$. Then the SL condition (7.15) becomes

$$\left[e_m^\star \frac{{\rm d}e_n}{{\rm d}x} \right]_{-\pi}^{\pi} = 0, \tag{7.23}$$

which is satisfied automatically. The ${\rm d}^2/{\rm d}x^2$ operator on the LHS is can be written as the $A_{\rm SL}$ of (7.14) with $p(x) = w(x) = 1$ and $q(x) = 0$. There are two LI solutions to (7.22) for any choice of λ_n: $e_n^+(x) = \exp({\rm i}\sqrt{-\lambda_n}x)$ and $e_n^-(x) = \exp(-{\rm i}\sqrt{-\lambda_n}x)$. To satisfy the boundary conditions $e_n(-\pi) = e_n(\pi)$ and

$e'_n(-\pi) = e'_n(\pi)$ we require that $\sin(\sqrt{-\lambda_n}\pi) = 0$, which restricts the eigenvalues for both $e^+_n(x)$ and $e^-_n(x)$ to $\lambda_n = -n^2$, where $n = 0, 1, 2, 3, \ldots$. This is a generic feature of SL problems:

boundary conditions ⇒ *quantization of eigenvalues.*

The full set of eigenfunctions is then $e^+_0(x) = e^-_0(x) = 1$, $e^+_1(x) = \mathrm{e}^{\mathrm{i}x}$, $e^+_2(x) = \mathrm{e}^{2\mathrm{i}x}$, ..., plus $e^-_1(x) = \mathrm{e}^{-\mathrm{i}x}$, $e^-_2(x) = \mathrm{e}^{-2\mathrm{i}x}$, ..., which we can normalise and combine into the single set

$$e_k(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{\mathrm{i}kx}, \quad k = 0, \pm 1, \pm 2, ..., \tag{7.24}$$

that we encountered in §4.4. We have already seen that these eigenfunctions are orthogonal and complete.

Legendre's differential equation Solving Laplace's equation $\nabla^2 V = 0$ in spherical polar coordinates for axisymmetric systems ($\partial_\phi = 0$) by substituting $V(r, \theta) = R(r)\Theta(\theta)$ and separating variables leads to the following equation for $\Theta(\theta)$:

$$\frac{1}{\sin\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left[\sin\theta\frac{\mathrm{d}}{\mathrm{d}\theta}\right]\Theta = \lambda\Theta, \tag{7.25}$$

where λ is a separation constant. Substituting $x = \cos\theta$, this becomes

$$\left[\frac{\mathrm{d}}{\mathrm{d}x}\left((1-x^2)\frac{\mathrm{d}}{\mathrm{d}x}\right)\right]\Theta = \lambda\Theta, \tag{7.26}$$

which is known as **Legendre's differential equation**. It is a Sturm–Liouville equation (7.18),

$$A_{\mathrm{SL}}\Theta_l = \lambda_l\Theta_l, \tag{7.27}$$

with $w(x) = 1$, $p(x) = 1 - x^2$ and $q(x) = 0$ in A_{SL} (equation 7.14). The natural limits on $x = \cos\theta$ are $a = -1$ and $b = 1$. Notice then that $p(-1) = p(1) = 0$ and so the boundary condition (7.15) is naturally satisified for any $\Theta(x)$. The eigenfunctions of this A_{SL} are **Legendre polynomials**, $\Theta_l(x) = P_l(x)$ with corresponding eigenvalues $\lambda_l = -l(l+1)$, where $l = 0, 1, 2, \ldots$ Section §8.1 explains how to explain explicit expressions for $P_l(x)$ and why the eigenvalues have the form $\lambda_l = -l(l+1)$.

The table below lists some common examples of Sturm–Liouville equations. Notice that in most cases $p(a) = p(b) = 0$, which means that the boundary condition (7.15) is automatically satisfied.

name	equation	$p(x)$	$q(x)$	$w(x)$	a, b	λ_n	$e_n(x)$
Fourier	$u'' = \lambda u$	1	0	1	$-\pi, \pi$	$-n^2$	$\mathrm{e}^{\pm\mathrm{i}nx}$
Legendre	$(1-x^2)u'' - 2xu' = \lambda u$	$1-x^2$	0	1	$-1, 1$	$-n(n+1)$	$P_n(x)$
Assoc. Legendre	$(1-x^2)u'' - 2xu' - \frac{m^2}{1-x^2}u = \lambda u$	$1-x^2$	$-\frac{m^2}{1-x^2}$	1	$-1, 1$	$-n(n+1)$	$P^m_n(x)$
Laguerre	$xu'' + (1-x)u' = \lambda u$	$x\mathrm{e}^{-x}$	0	e^{-x}	$0, \infty$	$-n$	$L_n(x)$
Hermite	$u'' - 2xu' = \lambda u$	e^{-x^2}	0	e^{-x^2}	$-\infty, \infty$	$-2n$	$H_n(x)$
Bessel	$x^2u'' + xu' + (x^2 - \nu^2)u = 0$	x	$-\frac{\nu^2}{x}$	x	(Bessel's special)		$J_\nu(x)$

Table: Some examples of ODEs in Sturm–Liouville (7.14) form, along with their eigenvalues λ_n and eigenfunctions $e_n(x)$, where $n = 0, 1, 2, \ldots$. Explicit expressions for many of the eigenfunctions listed here are obtained in the next section. Bessel's equation is peculiar – see §8.4.

7.7 Construction of eigenfunctions

In the next section we explain how to find explicit expressions for the eigenvalues and eigenfunctions of specific Sturm–Liouville equations by using the following methods.

(i) **Series solution** is the most direct method. Substituting $e_n(x+a) = x^\sigma \sum_{k=0}^\infty c_k(x+a)^k$ into the Sturm–Liouville equation and equating powers of x leads a recurrence relation for the series coefficients c_k. The series converges only for certain values of λ. This gives the spectrum of eigenvalues λ_n and the associated eigenvectors $e_n(x)$.

(ii) **Rodrigues' formula** works in some special cases. In particular, if (i) $q(x) = 0$ and (ii) $s(x) \equiv p(x)/w(x)$ is a quadratic (at most) polynomial with real roots, then the eigenfunctions are given by

$$u_n(x) \propto \frac{1}{K_n w(x)} \frac{\mathrm{d}^n}{\mathrm{d}x^n}\left[ws^n\right], \tag{7.28}$$

provided $u_1 \propto x$ and $w(a)s(a) = w(b)s(b) = 0$. The K_n are normalisation constants. Each such $u_n(x)$ is an n^{th}-order polynomial that satisfies the Sturm–Liouville equation (7.18) with eigenvalue

$$\lambda_n = n\left[\frac{1}{2}(n-1)s'' + K_1 u_1'\right] \tag{7.29}$$

for $n = 0, 1, 2, \ldots$. See DK III§10.2 for more on this method.

(iii) **Generating functions** are by far the most powerful method, but are much less direct. For example, Legendre polynomials can be *defined* by the function

$$G(x,t) = \frac{1}{\sqrt{1-2xt+t^2}} = \sum_{l=0}^\infty P_l(x)t^l, \tag{7.30}$$

in which each $P_l(x)$ is defined to be the coefficient of t^l in the Taylor-series expansion of $(1-2xt+t^2)^{-1/2}$. Legendre's differential equation can actually be derived from this $G(x,t)$, as can various useful recurrence relations among the P_l.

A proper understanding *how* methods (ii) and (iii) work requires familiarity with the material in the S1: *Functions of a complex variable* short option course and more. You don't need to acquire this, but you should be familiar with the most basic properties of the eigenfunctions in the next section. You should also have at least a vague awareness that, just as sines and cosines satisfy various trigonometric identities, there are are similar relations among the eigenfunctions of other differential operators.

8 Example Sturm–Liouville problems

This section explains how to derive explicit expressions for the eigenfunctions and eigenvalues of some of the more frequently encountered Sturm–Liouville equations (7.18). You need only be able to identify the particular ODEs involved and be broadly familiar with the properties of the solutions and how they are obtained.

The methods presented here can easily be applied to other Sturm–Liouville problems.

8.1 Legendre's differential equation

Legendre's differential equation is

$$(1-x^2)u'' - 2xu' = \lambda u. \tag{8.1}$$

This is a Sturm–Liouville equation with $p(x) = 1 - x^2$, $q(x) = 0$ and $w(x) = 1$. The boundaries on x are $a = -1$, $b = 1$. It is a special case of the associated Legendre equation to be discussed later.

Series solution Let us look for solutions of the form

$$u(x) = \sum_{k=0}^{\infty} a_k x^k, \tag{8.2}$$

where the coefficients a_i depend on the choice of the eigenvalue λ in (8.1). Substituting the series (8.2) into the differential equation (8.1) gives

$$(1-x^2)\sum_{k=2}^{\infty} k(k-1)a_k x^{k-2} - 2\sum_{k=0}^{\infty} k a_k x^k = \lambda \sum_{k=0}^{\infty} a_k x^k. \tag{8.3}$$

Changing the index label of the first sum from k to $j = k-2$, this becomes

$$\sum_{j=0}^{\infty}(j+2)(j+1)a_{j+2}x^j - \sum_{k=0}^{\infty} k(k-1)a_k x^k - 2\sum_{k=0}^{\infty} k a_k x^k = \lambda \sum_{k=0}^{\infty} a_k x^k, \tag{8.4}$$

in which all terms involve x raised to the power of some summation index. Notice too that all four sums have the same limits (0 and ∞) on their index. Rewriting all four sums so that they use a common label (k) for their index and gathering together into a single sum, we have that

$$\sum_{k=0}^{\infty} \underbrace{\left\{(k+2)(k+1)a_{k+2} - [k(k+1)+\lambda]a_k\right\}}_{0} x^k = 0. \tag{8.5}$$

The quantity in curly brackets here must be zero because the x^k are LI. So, we have derived a **recurrence relation**,

$$a_{k+2} = \frac{k(k+1)+\lambda}{(k+2)(k+1)} a_k, \tag{8.6}$$

for the coefficients of the series solution (8.2). Notice that $a_{k+2}/a_k \to 1$ as $k \to \infty$, which means that the series will in general diverge as $x \to \pm 1$. The only way of avoiding this is if the series terminates: there must be some value of k for which the coefficient

$$\frac{k(k+1)+\lambda}{(k+2)(k+1)} = 0, \tag{8.7}$$

so that $a_{k+2} = a_{k+4} = \cdots = 0$. This means that the eigenvalues λ must be of the form $\lambda_l = -l(l+1)$, where $l = 0$ or 1 or 2 or 3 or ....

If we choose l to be an even (odd) number, then all of the odd- (even-)numbered a_k must be zero to avoid a divergent series. Therefore, if l is even (odd) the eigenfunctions are even (odd) functions of x. The most natural starting point for the recurrence relation is $(a_0, a_1) = (1, 0)$ if l is even, or $(a_0, a_1) = (0, 1)$ if l is odd. The first few $u_l(x)$ constructed in this way are

$$u_0(x) = 1, \quad u_1(x) = x, \quad u_2(x) = -3x^2 + 1, \quad u_3(x) = -\frac{5}{3}x^3 + x. \tag{8.8}$$

If we normalise these $u_l(x)$ so that $u_l(1) = 1$ we obtain the first few **Legendre polynomials**,

$$P_0(x) = 1, \quad P_1(x) = x, \quad P_2(x) = \frac{1}{2}(3x^2 - 1), \quad P_3(x) = \frac{1}{2}(5x^3 - 3x). \tag{8.9}$$

Note that this weighty historical convention that $P_l(1) = 1$ means that the $P_l(x)$ are not orthonormal. Instead they satsify the orthogonality relation

$$\langle P_l | P_m \rangle \equiv \int_{-1}^{1} P_l^\star(x) P_m(x)\,\mathrm{d}x = \frac{2}{2l+1}\delta_{lm}. \tag{8.10}$$

The $P_l(x)$ are a basis for the space $L_1^2(-1, 1)$ because they are eigenfunctions of a Sturm–Liouville operator (7.14) that has $a = -1$, $b = 1$ and $w(x) = 1$. Therefore we can express any well-behaved $f : [-1, 1] \to \mathbb{C}$ as a **Fourier–Legendre** series,

$$f(x) = \sum_{l=0}^{\infty} a_l P_l(x) \tag{8.11}$$

with coefficients

$$a_l = \frac{2l+1}{2} \int_{-1}^{1} P_l(x)^\star f(x)\,\mathrm{d}x. \tag{8.12}$$

[The * in the integrands of (8.12) and (8.10) is redundant because the P_l are real, but I leave it in to remind you that all this is making use of the inner product (4.2) that appears in the generalised Fourier series (4.8).] We have already encountered this series in §4.3, when we used the Gram–Schmidt procedure to construct an orthonormal basis for this space starting from the list of monomials 1, x, x^2.... The $e_n(x)$ we constructed there happen to be normalised Legendre polynomials.

Rodrigues' formula Legendre's differential equation satisfies the conditions for Rodrigues' formula (equation 7.28) to apply: $q(x) = 0$; $s(x) = p(x)/w(x) = 1 - x^2$ is quadratic with real roots and with $s(-1) = s(1) = 0$; the first $P_1(x) \propto x$. Substituting this $s(x)$ and $w(x) = 1$ into equation (7.28) we obtain Rodrigues' formula for the $P_l(x)$,

$$P_l(x) = \frac{1}{2^l l!} \frac{\mathrm{d}^l}{\mathrm{d}x^l} (x^2 - 1)^l, \tag{8.13}$$

which satisfies Legendre's differential equation (8.1) with eigenvalue $\lambda_l = -l(l + 1)$. The prefactor in this expression is chosen to maintain the convention that $P_l(1) = 1$.

Generating function Finally, we note that $P_l(x)$ can be *defined* in terms of the generating function

$$G(x, t) = \frac{1}{(1 - 2xt + t^2)^{1/2}} = \sum_{l=0}^{\infty} P_l(x) t^l. \tag{8.14}$$

Carrying out a Taylor expansion, we have that

$$\begin{aligned} \left[1 - 2xt + t^2\right]^{-1/2} &= 1 - \frac{1}{2}\left(-2xt + t^2\right) + \frac{1}{2!}\frac{1}{2}\frac{3}{2}\left(-2xt + t^2\right)^2 - \frac{1}{3!}\frac{1}{2}\frac{3}{2}\frac{5}{2}\left(-2xt + t^2\right)^3 + \ldots \\ &= t^0 + xt^1 + \frac{1}{2}(3x^2 - 1)t^2 + \frac{1}{2}(5x^3 - 3x)t^3 + \ldots, \end{aligned} \tag{8.15}$$

in agreement with the first few $P_l(x)$ found above in (8.9). The innocent-looking expression (8.14) also encodes Legendre's differential equation, the normalisation properties of the $P_l(x)$, recurrence relations and much more. See homework for some examples.

8.2 Associated Legendre equation

The associated Legendre equation,

$$(1-x^2)\frac{\mathrm{d}^2u}{\mathrm{d}x^2} - 2x\frac{\mathrm{d}u}{\mathrm{d}x} - \frac{m^2}{1-x^2}u = \lambda u, \tag{8.16}$$

appears when using separation of variables to solve equations that involve the Laplacian ∇^2 in spherical polar coordinates (r,θ,ϕ). The variable $x=\cos\theta$, so that the natural boundaries are $a=-1$, $b=1$. The equation is an example of a Sturm–Liouville problem with $p(x)=1-x^2$, $q(x)=-m^2/(1-x^2)$ and $w(x)=1$. Legendre's differential equation corresponds to the special case $m=0$.

The eigenfunctions that satisfy (8.16) are the **associated Legendre functions**,

$$P_l^m(x) \equiv (1-x^2)^{m/2}\frac{\mathrm{d}^m}{\mathrm{d}x^m}P_l(x), \quad (m \geq 0). \tag{8.17}$$

The eigenvalue corresponding to the eigenfunction $P_l^m(x)$ is $\lambda_l = -l(l+1)$, where $l=0,1,2,...$.

One way of showing this is by differentiating Legendre's equation (8.1),

$$(1-x^2)P_l'' - 2xP_l' = -l(l+1)P_l, \tag{8.18}$$

m times to obtain

$$(1-x^2)u'' - 2x(m+1)u' - m(m+1)u = -l(l+1)u, \tag{8.19}$$

where we have introduced

$$u(x) \equiv \frac{\mathrm{d}^m}{\mathrm{d}x^m}P_l(x). \tag{8.20}$$

If we now rewrite (8.19) in terms of

$$v(x) = (1-x^2)^{m/2}u(x), \tag{8.21}$$

the result, after much manipulation, is

$$(1-x^2)v'' - 2xv' - \frac{m^2}{1-x^2}v = -l(l+1)v, \tag{8.22}$$

which is the associated Legendre equation (8.16). This proves the statement around equation (8.17).

We have so far assumed that $m \geq 0$. Using Rodrigues' formula (8.13) for $P_l(x)$ in (8.17), we have that

$$P_l^m(x) = \frac{1}{2^l l!}(1-x^2)^{m/2}\frac{\mathrm{d}^{l+m}}{\mathrm{d}x^{l+m}}(x^2-1)^l, \tag{8.23}$$

which is used to define $P_l^m(x)$ for both positive and negative m. Clearly $P_l^m(x)=0$ unless $|m|\leq l$.

Exercise: By applying Leibnitz' formula to $(x+1)^l(x-1)^l$, show that

$$P_l^{-m}(x) = (-1)^m\frac{(l-m)!}{(l+m)!}P_l^m(x). \tag{8.24}$$

These $P_l^m(x)$ are used in the definition of **spherical harmonics**, which are a natural basis for two-dimensional functions $f(\theta,\phi)$ defined on the surface of a sphere (see §12.4 later). For reference, we note here that the associated Legendre functions satisfy the orthogonality relation

$$\int_{-1}^{1} P_k^m(x)P_l^m(x)\,\mathrm{d}x = \frac{2}{2l+1}\frac{(l+m)!}{(l-m)!}\delta_{kl}. \tag{8.25}$$

8.3 Hermite's differential equation

Hermite's differential equation,

$$u'' - 2xu' = \lambda u, \tag{8.26}$$

can be reexpressed in Sturm–Liouville form by using equation (7.17) to rewrite it as

$$\mathrm{e}^{x^2}\frac{\mathrm{d}}{\mathrm{d}x}\left[\mathrm{e}^{-x^2}\frac{\mathrm{d}u}{\mathrm{d}x}\right] = \lambda u, \tag{8.27}$$

in which $w(x) = p(x) = \exp(-x^2)$ and $q(x) = 0$. The natural limits to choose in order to make $p(a) = p(b) = 0$ so that the boundary conditions (7.15) hold are $a \to -\infty$, $b \to \infty$.

Series solution Let us find the eigenfunctions and eigenvalues of Hermite's differential equation by determining the coefficients of the series solution $u(x) = \sum_{n=0}^{\infty} a_n x^n$ that satisfies (8.26). Substituting this into (8.26), we have

$$\sum_{n=2}^{\infty} n(n-1)a_n x^{n-2} - 2\sum_{n=1}^{\infty} n a_n x^n - \lambda\sum_{n=0}^{\infty} a_n x^n = 0. \tag{8.28}$$

Introduce a new label $k = n + 2$ in the first sum, and replace the index n by k in the second and third sums. Notice that we can safely change the lower limit of the second sum from $k = 1$ to $k = 0$. Then equation (8.28) becomes

$$\sum_{k=0}^{\infty}(k+2)(k+1)a_{k+2}x^k - 2\sum_{k=0}^{\infty} k a_k x^k - \lambda\sum_{k=0}^{\infty} a_k x^k = 0, \tag{8.29}$$

or, when written in terms of a single sum,

$$\sum_{k=0}^{\infty}\left\{(k+2)(k+1)a_{k+2} - 2k a_k - \lambda a_k\right\}x^k = 0. \tag{8.30}$$

The quantity in curly brackets must be zero because the x^k are LI. Therefore the recurrence relation for the coefficients a_k is

$$a_{k+2} = \frac{2k+\lambda}{(k+2)(k+1)}a_k. \tag{8.31}$$

Recall from (4.1) that we require that $\int_{-\infty}^{\infty} e^{-x^2}|u(x)|^2\,\mathrm{d}x$ be finite. If the recurrence relation (8.31) does not terminate then it is easy to show that for large x the resulting series $u(x) = \sum_k a_k x^k$ increases faster than $\exp(+x^2/2)$: the Maclaurin expansion of the latter has $a_{k+2}/a_k = 1/(k+2)$. So, the recurrence relation (8.31) must terminate if the integral (4.1) is to converge. That means that λ is restricted to values $\lambda_n = -2n$, where $n = 0, 1, 2, \ldots$. By the same argument used earlier for Legendre polynomials, the eigenfunctions are polynomials involving even (odd) powers of x if n is even (odd). Starting from either $(a_0, a_1) = (1, 0)$ (even n) or $(a_0, a_1) = (0, 1)$ (odd n), the recurrence relation gives the first few eigenfunctions $u_n(x)$ as

$$u_0(x) = 1, \quad u_1(x) = x, \quad u_2(x) = -2x^2 + 1, \quad u_3(x) = -\frac{2}{3}x^3 + x. \tag{8.32}$$

If we wanted to, we could normalise these $u_n(x)$ and anoint them as our "standard" eigenfunctions for Hermite's differential equation. Choosing this normalisation would mean throwing away centuries of accumulated wisdom, however.

Generating function Instead, let us follow tradition and introduce the generating function

$$G(x,t) = \mathrm{e}^{x^2}\mathrm{e}^{-(t-x)^2} = \sum_{n=0}^{\infty} H_n(x)\frac{t^n}{n!}, \tag{8.33}$$

which defines the **Hermite polynomials** $H_n(x)$. The factor of $1/n!$ is included to make the H_n satisfy the orthogonality relation

$$\langle H_m|H_n\rangle = \int_{-\infty}^{\infty} \mathrm{e}^{-x^2} H_m^\star(x) H_n(x)\,\mathrm{d}x = 2^n\pi^{1/2} n!\,\delta_{mn}. \tag{8.34}$$

The fact that the $H_n(x)$ are eigenfunctions of (8.26) with eigenvalue $\lambda = -2n$ is encoded in the generating function (8.33), as are various recurrence relations among the $H_n(x)$. For reference, the first few $H_n(x)$ extracted by expanding the generating function as a power series in t are

$$H_0(x) = 1, \quad H_1(x) = 2x, \quad H_2(x) = 4x^2 - 2, \quad H_3(x) = 8x^3 - 12, \tag{8.35}$$

in agreement (apart from normalisation) with the results (8.32) obtained from the recurrence relation.

Rodrigues' formula Another way of defining Hermite polynomials is by the Rodrigues' formula

$$H_n(x) = (-1)^n \mathrm{e}^{x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\mathrm{e}^{-x^2}\right). \tag{8.36}$$

> **Exercise:** Show that (8.36) follows from the generating function (8.33) by differentiating $G(x,t)$ n times with respect to t and then setting $t = 0$.

The completeness property of Hermite polynomials means that they are a basis for the space of functions defined on the real line with weight function e^{-x^2}. Any well-behaved function $f : \mathbb{R} \to \mathbb{C}$ can be expressed as the generalised Fourier series

$$f(x) = \sum_{n=0}^{\infty} a_n H_n(x), \tag{8.37}$$

where the coefficients

$$a_n = \frac{1}{2^n\sqrt{\pi}n!} \int_{-\infty}^{\infty} H_n^\star(x) f(x) \mathrm{e}^{-x^2}\,\mathrm{d}x, \tag{8.38}$$

the prefactor coming from the orthogonality relation (8.34). The $\star$ in the integrand here is redundant, but is left in as a reminder that equations (8.37) and (8.38) are simply special cases of equation (4.8).

8.4 Bessel's equation

Bessel's equation,

$$x^2u'' + xu' + (x^2 - \nu^2)u = 0, \tag{8.39}$$

often appears in problems involving cylindrical polar coordinates (R, ϕ, z), with the variable x being some multiple of the radius R and the constant ν set by the details of the problem. Equation (8.39) can be squeezed into Sturm–Liouville form by introducing a new independent variable $\bar{x} = x/\lambda$, in which case it can be written as

$$\frac{1}{\bar{x}}\left[\frac{\mathrm{d}}{\mathrm{d}\bar{x}}\left(\bar{x}\frac{\mathrm{d}}{\mathrm{d}\bar{x}}\right) - \frac{\nu^2}{\bar{x}}\right] u = -\lambda^2 u, \tag{8.40}$$

so that $w(\bar{x}) = \bar{x}$, $p(\bar{x}) = \bar{x}$, $q(\bar{x}) = -\nu^2/\bar{x}$ and the definition of the independent variable $\bar{x} = x/\lambda$ depends on the eigenvalue $-\lambda^2$. Unlike the other examples in this section, the boundary conditions in applications of Bessel's equation normally depend on the details of the problem (see §13 later for an example).

Series solution For simplicity we consider only the case where $\nu = m \geq 0$, a non-negative integer. Substituting $u(x) = \sum_{n=0}^{\infty} a_n x^n$ into (8.39) gives

$$\sum_{n=2}^{\infty} n(n-1)a_n x^n + \sum_{n=1}^{\infty} n a_n x^n + \sum_{n=0}^{\infty} a_n x^{n+2} - m^2 \sum_{n=0}^{\infty} a_n x^n = 0. \tag{8.41}$$

Writing $k = n$ in the first, second and fourth sums and $k = n + 2$ in the third, this becomes

$$\sum_{k=2}^{\infty} k(k-1)a_k x^k + \sum_{k=1}^{\infty} k a_k x^k + \sum_{k=2}^{\infty} a_{k-2} x^k - m^2 \sum_{k=0}^{\infty} a_k x^k = 0, \tag{8.42}$$

which, gathering together powers of x^k, is

$$-m^2 a_0 + (1-m^2)a_1 x + \sum_{k=2}^{\infty} \left\{(k^2 - m^2)a_k + a_{k-2}\right\} x^k. \tag{8.43}$$

As the x^k are LI, we immediately have the reccurence relation

$$a_k = -\frac{a_{k-2}}{k^2 - m^2}, \tag{8.44}$$

which relates the even coefficients to one another and the odd coefficients to one another. Clearly, if $m = 0$ then $a_1 = 0$, all odd coefficients a_k vanish and we may choose $a_0 \neq 0$ to start the recurrence. If $m = 1$, then $a_0 = 0$, all even coefficients vanish and we may start with $a_1 \neq 0$. More generally, given $m \geq 0$ we need to set $a_0 = \cdots = a_{m-1} = 0$ and can then choose $a_m \neq 0$ to start. The conventional choice is $a_m = 1/(2^m m!)$. The resulting eigenfunctions are the (integer-order) **Bessel functions**,

$$J_m(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!(m+n)!} \left(\frac{x}{2}\right)^{m+2n}. \tag{8.45}$$

Generating function More generally, Bessel functions of integer order can be *defined* through

$$G(x,t) = \exp\left[\frac{1}{2}x\left(t - \frac{1}{t}\right)\right] = \sum_{n=-\infty}^{\infty} J_n(x)t^n. \tag{8.46}$$

Further reading

Almost any book with "mathematical methods" and "physics" in its title will cover the topics we have merely skimmed over in this section. See, for example, RHB§18. Arfken & Weber's *Mathematical Methods for Physicists* provides good overviews of how the various methods used to define special functions (differential equation, generating functions, Rodrigues and more) are related to one another.

The Sturm–Liouville equation is a second-order ODE, which of course has two LI solutions. Here we have focused on finding the "well-behaved" solutions that satisfy certain boundary conditions. (We did not explicitly state our assumptions about the boundary conditions for the solution to Bessel's equation, but the form of the series we assumed was an implicit boundary condition.) The books mentioned above give more details on how to find the second, LI solutions that are less well behaved, but sometimes useful.

9 Inhomogeneous ODEs: Green's functions

Suppose that we want to solve the inhomogenous differential equation

$$A_x u(x) = f(x), \tag{9.1}$$

where A_x is a second-order linear differential operator of the form $A_{\rm SL}$ from equation (7.14) and the subscript x merely emphasises we're going to consider it doing its thing to functions of x. We impose the condition that the solution $u(x) \in L^2_w(a,b)$ should satisfy the pair of homogeneous boundary conditions (7.6).

If we can find a function $G(x,y)$ for which

$$A_x G(x,y) = \frac{1}{w(x)}\delta(x-y) \tag{9.2}$$

that satisfies these boundary conditions, then we would have that

$$\begin{aligned} A_x\left[\int_a^b G(x,y)f(y)\,w(y){\rm d}y\right] &= \int_a^b A_x G(x,y)f(y)\,w(y){\rm d}y \\ &= \int_a^b \frac{1}{w(x)}\delta(x-y)f(y)\,w(y){\rm d}y = f(x). \end{aligned} \tag{9.3}$$

That is, the solution to (9.1) would be simply

$$u(x) = \int_a^b G(x,y)f(y)\,w(y){\rm d}y. \tag{9.4}$$

The function $G(x,y)$ that satisfies (9.2) and the appropriate boundary conditions is the **Green's function** for the problem. In terms of this $G(x,y)$, the "inverse" of A_x, assuming the given boundary conditions, is therefore the operator

$$A_x^{-1}\bullet = \int_a^b {\rm d}y\, w(y)\, G(x,y)\bullet\,. \tag{9.5}$$

9.1 Properties of Green's functions

Here are the most important properties satisfied by Green's functions $G(x,y)$:

(0) Boundary conditions matter! $G(x,y)$ depends on both the operator A_x and on the choice made for the pair of boundary conditions (7.6).

(1) As we are considering Hermitian operators A_x, then $\langle G(x,y), A_x G(x,y')\rangle = \langle G(x,y'), A_x G(x,y)\rangle^\star$. Writing out both sides as integrals and using the definition (9.2) results in $G^\star(y',y) = G(y,y')$. That is,

$$G(y,x) = G^\star(x,y). \tag{9.6}$$

(2) $A_x G(x,y) = 0$, except at $x=y$.

(2) $G(x,y)$ must be a continuous function of x,

(3) but its first derivative with respect to x is discontinuous at $x=y$: substituting the form (7.14) for $A_{\rm SL}$ into (9.2) and integrating from $x=y-\epsilon$ to $x=y+\epsilon$, then taking the limit $\epsilon\to 0$, we have that

$$\left.\frac{\partial G(x,y)}{\partial x}\right|_{x=y+\epsilon} - \left.\frac{\partial G(x,y)}{\partial x}\right|_{x=y-\epsilon} = \frac{1}{p(y)}. \tag{9.7}$$

Let $e_n(x)$ be the $n^{\rm th}$ normalised eigenfunction of A_x, with corresponding eigenvalue λ_n. Since these $e_n(x)$ are complete, we can expand $G(x,y)$ for fixed y as $G(x,y) = \sum_n a_n(y)e_n(x)$. Substituting this $G(x,y)$ into (9.2) and taking the inner product with $e_m(x)$, we find that $a_m(y)\lambda_m = e^\star_m(y)$. That is,

$$G(x,y) = \sum_n \frac{e_n(x)e^\star_n(y)}{\lambda_n}. \tag{9.8}$$

9.2 Examples of constructing Green's functions

Having established these properties, the procedure for constructing $G(x,y)$ is straightforward, at least in principle:

(1) Ignoring the boundary conditions, find two LI solutions $u_1(x)$ and $u_2(x)$ to the homogeneous equation $A_x u_i(x) = 0$.

(2) Treating y as a constant, $G(x,y)$ must be one linear combination of $u_1(x)$ and $u_2(x)$ when $x < y$, and another when $x > y$. So, write

$$G(x,y) = \begin{cases} B_1(y)u_1(x) + B_2(y)u_2(x), & \text{when } x < y, \\ C_1(y)u_1(x) + C_2(y)u_2(x), & \text{when } x > y. \end{cases} \tag{9.9}$$

(3) The boundary conditions on $G(x,y)$ give two equations among the four unknown functions $B_i(y)$ and $C_i(y)$. Continuity of G and the discontinuity of its first derivative provide another two. Solve to find B_i and C_i.

Forced harmonic oscillator The displacement $x(t)$ of the oscillator is described by the ODE

$$\ddot{x} + x = f(t), \tag{9.10}$$

where $f(t)$ is the forcing. Find an expression for $x(t)$ in terms of $f(t)$ given that $x(0) = \dot{x}(0) = 0$.

We have $A_t = \frac{\mathrm{d}^2}{\mathrm{d}t^2} + 1$, which is of the form A_{SL} with $p(x) = q(x) = w(x) = 1$. The solutions to the homogeneous equation $A_t x = 0$ are $x = \cos t$ and $x = \sin t$. Superposing these on either side of $t = t'$,

$$G(t,t') = \begin{cases} A(t')\cos t + B(t')\sin t, & \text{if } t < t' \\ C(t')\cos t + D(t')\sin t, & \text{if } t > t'. \end{cases} \tag{9.11}$$

The boundary condition $G(0,t') = \dot{G}(0,t') = 0$ means that $A = B = 0$. So,

$$G(t,t') = \begin{cases} 0, & \text{if } t < t' \\ C(t')\cos t + D(t')\sin t, & \text{if } t > t'. \end{cases} \tag{9.12}$$

At $t = t'$ the continuity of G and discontinuity of $\dot{G}$ require that

$$\begin{aligned} C(t)\cos t + D(t)\sin t &= 0, \\ -C(t)\sin t + D(t)\cos t &= 1, \end{aligned} \tag{9.13}$$

respectively. So, $C(t') = -\sin t'$ and $D(t') = \cos t'$. Then Green's function is, finally,

$$G(t,t') = \begin{cases} 0, & \text{if } t < t' \\ -\sin t'\cos t + \cos t'\sin t = \sin(t-t'), & \text{if } t > t', \end{cases} \tag{9.14}$$

and the solution to (9.10) subject to the initial conditions $x(0) = \dot{x}(0) = 0$ is

$$x(t) = \int_0^t \mathrm{d}t' \, \sin(t-t') f(t'). \tag{9.15}$$

Spherical charge distribution The potential $V(r)$ due to a spherical distribution of charge $\rho(r)$ satisfies Poisson's equation,

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}V}{\mathrm{d}r}\right) = -\tfrac{1}{\epsilon_0}\rho(r). \tag{9.16}$$

Find $V(r)$ subject to the condition that $\mathrm{d}V/\mathrm{d}r \to 0$ as $r \to 0$ and $V \to 0$ as $r \to \infty$.

This A_r is of Sturm–Liouville form with $w(r) = p(r) = r^2$ and $q(r) = 0$. The two LI solutions to the homogeneous equation $A_r V = 0$ are $V(r) = 1$ and $V(r) = 1/r$. Superposing,

$$G(r, r') = \begin{cases} B_+(r') + C_+(r')/r, & \text{if } r < r', \\ B_-(r') + C_-(r')/r, & \text{if } r > r'. \end{cases} \tag{9.17}$$

The boundary conditions imply that $B_- = C_+ = 0$, so

$$G(r, r') = \begin{cases} B(r'), & \text{if } r < r', \\ C(r')/r, & \text{if } r > r'. \end{cases} \tag{9.18}$$

Continuity of $G(r, r')$ at $r = r'$ requires that $B(r) = C(r)/r$, while the discontinuity of $\partial G/\partial r$ there requires $-C(r)/r^2 = 1/p(r) = 1/r^2$. So, $C(r') = -1$, $B(r') = -1/r'$ and the Green's function is

$$G(r, r') = -\begin{cases} 1/r', & \text{if } r < r', \\ 1/r, & \text{if } r > r'. \end{cases} \tag{9.19}$$

Using (15.4) and remembering that $w(r') = r'^2$, the potential satisfying these boundary conditions is therefore

$$V(r) = \tfrac{1}{\epsilon_0}\left[\frac{1}{r}\int_0^r \rho(r') r'^2 \, \mathrm{d}r' + \int_r^\infty \rho(r') r' \, \mathrm{d}r'\right]. \tag{9.20}$$

Maths Methods Week 4–: PDEs

http://www-thphys.physics.ox.ac.uk/people/JohnMagorrian/mm
john.magorrian@physics.ox.ac.uk

10 PDEs: introduction

Most undergraduate physics problems boil down to solving differential equations, often partial differential equations. For example, Laplace's equation,

$$\nabla^2\psi = 0, \tag{10.1}$$

occurs in electrostatics, steady-state hydrodynamics and heat flow. The time-dependent *diffusion equation*,

$$\frac{\partial\psi}{\partial t} = \nabla\cdot(D\nabla\psi), \tag{10.2}$$

occurs in problems involving heat flow, nuclear reactions, approximations of random walks, among others. In many problems the diffusion coefficient $D(\psi,\mathbf{r})$ is constant and (10.2) becomes the *heat equation*

$$\frac{\partial\psi}{\partial t} = D\nabla^2\psi. \tag{10.3}$$

Using separation of variables $\psi(\mathbf{r},t)=\phi(\mathbf{r})T(t)$ in the familiar wave equation,

$$c^2\nabla^2\psi - \frac{\partial^2\psi}{\partial t^2} = 0, \tag{10.4}$$

leads to one sign choice in the Helmholtz equation,

$$\nabla^2\phi \pm k^2\phi = 0. \tag{10.5}$$

Finally, Schrödinger's equation,

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V(\mathbf{r})\psi = \mathrm{i}\hbar\frac{\partial\psi}{\partial t}, \tag{10.6}$$

is familiar from quantum mechanics.

These equations are all *linear, second-order homogenous* PDEs: they can all be written in the form

$$\hat{D}\psi = 0, \tag{10.7}$$

where $\hat{D}$ is a linear second-order differential operator in $(\mathbf{x},t)$. Linear homogenous equations have the convenient property that we can superpose solutions: if $\psi_1(\mathbf{x},t)$ and $\psi_2(\mathbf{x},t)$ are two solutions, then so too is $\alpha_1\psi_1+\alpha_2\psi_2$.

The next few sections will give some examples that illustrate how to solve such PDEs in practice. The general procedure is:

(0) Identify the boundary conditions (physics);
(0') Decide on a suitable coordinate system for the problem;
(1) Obtain the general solution for the PDE to be solved;
(2) Find the specific solution that satisfies the boundary conditions.

In these lectures we ignore steps (0) and (0') and concentrate on steps (1) and (2).

Equations of the form $\hat{D}=g$ in which $g\neq 0$ are *inhomogenous*. An example of an inhomogenous equation is Poisson's equation, $\nabla^2\psi = -\rho/\epsilon_0$. Such equations can be solved using the method of Green's functions for PDEs (§15 below).

11 The method of characteristics for PDEs[⋆]

Before going on to the detailed examples, note that there are many important PDEs that are not second-order linear equations. For example, the continuity equation,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \tag{11.1}$$

is a linear first-order equation if we happen to know $\mathbf{u}(\mathbf{r}, t)$, but nonlinear if neither ρ nor $\mathbf{u}$ are known. The momentum equation

$$\rho \left(\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} \right) = -\nabla p - \rho \nabla \Phi, \tag{11.2}$$

is an obviously non-linear first-order inhomogenous PDE for $\mathbf{u}(\mathbf{r}, t)$. We can gain insight into such first-order equations by examining their *characteristic curves*. The method of characteristics is extremely important in understanding the mechanics of fluids. It also offers a way of classifying second-order linear PDEs, with implications for the type of boundary conditions needed to specify the solution uniquely.

11.1 Characteristics for first-order quasilinear PDEs

Consider the first-order inhomogeneous quasilinear PDE for $u(t, x)$,

$$a(t, x, u) \frac{\partial u}{\partial t} + b(t, x, u) \frac{\partial u}{\partial x} = g(t, x, u). \tag{11.3}$$

This is linear in the derivatives of u, but not necessarily in u itself. There is a simple geometrical way of looking at this equation. Introduce rectilinear coordinates (t, x, u) and consider surfaces of constant

$$f(t, x, u) \equiv U(t, x) - u. \tag{11.4}$$

In this space the vector

$$\mathbf{n} \equiv \nabla f = (\partial_t f, \partial_x f, \partial_u f) = (\partial_t U, \partial_x U, -1) \tag{11.5}$$

is perpendicular to such surfaces. The PDE (11.3) can be written as $\mathbf{c} \cdot \mathbf{n} = 0$, where $\mathbf{c} = (a, b, g)$, and its solution as $u = U(t, x)$, which corresponds to the surface $f = 0$. We can remain on this solution surface by making a displacement $\mathrm{d}\mathbf{s} = (\mathrm{d}t, \mathrm{d}x, \mathrm{d}u)$ that is parallel to $\mathbf{c}$: the components $(\mathrm{d}t, \mathrm{d}x, \mathrm{d}u)$ should then vary together as

$$\frac{\mathrm{d}t}{a} = \frac{\mathrm{d}x}{b} = \frac{\mathrm{d}u}{g}. \tag{11.6}$$

This ODE is known as the **characteristic equation** for the PDE (11.3). Integrating it yields the **characteristic curves** (or simply the **characteristics**) for the PDE. By starting at, say $t = 0$, with different values of x, we can follow the characteristic curves to (attempt to) build up the full solution $u(t, x)$ from the initial values $u(t = 0, x)$.

Example: The function $u(x, y)$ satisfies the PDE

$$y \frac{\partial u}{\partial x} - x \frac{\partial u}{\partial y} = 0. \tag{11.7}$$

Find $u(x, y)$ subject to the boundary condition $u(x, 0) = \sin(\pi x / a)$ for $0 < x < a$.

The characteristic equation is

$$\frac{\mathrm{d}x}{y} = -\frac{\mathrm{d}y}{x} = \frac{\mathrm{d}u}{0}. \tag{11.8}$$

[⋆] Bonus material

This says that u is constant along the curve $x\mathrm{d}x = -y\mathrm{d}y$, which integrates to $x^2 + y^2 = c^2$, where c is a constant. Following these characteristic curves starting from the given $u(x, 0) = \sin(\pi x/a)$, the solution is therefore $u(x, y) = \sin(\pi r/a)$ for $r^2 \equiv x^2 + y^2 < a^2$. Outside that circle the solution is undetermined, as no boundary conditions have been given there.

Example: One dimensional pressureless fluid in gravitational field Euler's equation for the velocity $u(t, x)$ of a one-dimensional, pressureless fluid in a gravitational field g is

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = g. \tag{11.9}$$

The characteristic equation is

$$\mathrm{d}t = \frac{\mathrm{d}x}{u} = \frac{\mathrm{d}u}{g}, \tag{11.10}$$

which tells us that along the trajectory $\frac{\mathrm{d}x}{\mathrm{d}t} = u$ the velocity u varies as $\frac{\mathrm{d}u}{\mathrm{d}t} = g$. For simplicity, let us remove the gravitational field by setting $g = 0$. Then the general solution is given implicitly by $u(t, x) = f(x - ut)$, where f is an arbitrary function of one variable. If our initial conditions (the function $u = f(x)$, evaluated at $t = 0$), result in characteristic curves that cross (meaning $u(t, x)$ becomes multivalued) then we have a **shock**, signifying that the PDE (11.9) is inadequate and we need to return to examine the physics of the system. Nevertheless, the characteristic equations (11.10) are key in understanding how information propagates away from the shock.

Exercise: Sketch the characteristic curves for the contrived initial condition $u(t = 0, x) = -\sin x$. Show that they cross at $t = \frac{\pi}{2}$.

11.2 Classification of second-order PDEs

The same idea can be generalised to the case of n coupled first-order PDEs for n functions $u_k(x, y)$, $k = 1, ..., n$. Consider the n equations ($i = 1, ..., n$)

$$\sum_{k=1}^{n}\left[X_{ik}\frac{\partial u_k}{\partial x} + Y_{ik}\frac{\partial u_k}{\partial y}\right] = H_i, \tag{11.11}$$

in which X_{ik}, Y_{ik} and H_{ik} are functions of x, y, the u_k and their first derivatives. We can't hope to make the contents of each square bracket above vanish individually; the best we can do is to extract n independent coupled ODEs for the n functions u_k. To try to find such coupled ODEs, let's look for linear combinations of the equations (11.11) in which the LHS turns into a sum of total derivatives,

$$\mathrm{d}u_k = \frac{\partial u_k}{\partial x}\mathrm{d}x + \frac{\partial u_k}{\partial y}\mathrm{d}y, \tag{11.12}$$

of the u_k. That is, we want to be able to find $\mathrm{d}s_i$ for which

$$\sum_{i=1}^{n}\mathrm{d}s_i X_{ik} = L_k\mathrm{d}x \quad \text{and} \quad \sum_{i=1}^{n}\mathrm{d}s_i Y_{ik} = L_k\mathrm{d}y \tag{11.13}$$

simultaneously for some L_k. Then, multiplying (11.11) by $\mathrm{d}s_i$ and summing over i, we'd have

$$\sum_{k=1}^{n} L_k\mathrm{d}u_k = \sum_{i=1}^{n} H_i\mathrm{d}s_i. \tag{11.14}$$

If we could find n independent such equations (i.e., n independent choices of $\mathrm{d}s_i$ and corresponding L_k) then we'd have n simultaneous ODEs for the n functions u_k. From (11.13) the condition for this is that there be

n independent solutions $\mathrm{d}s_i$ to the equation $\sum_i \mathrm{d}s_i(X_{ik}\mathrm{d}y - Y_{ik}\mathrm{d}x) = 0$. A necessary and sufficient condition for this to be true is that

$$\det(X_{ik}\mathrm{d}y - Y_{ik}\mathrm{d}x) = 0. \tag{11.15}$$

With this in hand, we can now turn to using the method of characteristics to examine the second-order PDE

$$A\frac{\partial^2\Phi}{\partial x^2} + 2B\frac{\partial^2\Phi}{\partial x\partial y} + C\frac{\partial^2\Phi}{\partial y^2} = D \tag{11.16}$$

for the function $\Phi(x, y)$ in which A, B, C are functions of (x, y) and D may depend on $\partial\Phi/\partial x$ and $\partial\Phi\partial y$ as well. Writing

$$u_1 = \frac{\partial\Phi}{\partial x}, \quad u_2 = \frac{\partial\Phi}{\partial y}, \tag{11.17}$$

this reduces to the coupled first-order PDEs

$$\frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y} = 0, \quad A\frac{\partial u_1}{\partial x} + B\left(\frac{\partial u_1}{\partial y} + \frac{\partial u_2}{\partial x}\right) + C\frac{\partial u_2}{\partial y} = D. \tag{11.18}$$

Plugging $X_{12} = -Y_{11} = 1$ and $Y_{21} = A$, $Y_{21} = X_{22} = B$, $Y_{22} = D$ into (11.15), the condition for existence of nontrivial $\mathrm{d}s_i$ is

$$\det\begin{pmatrix} \mathrm{d}x & \mathrm{d}y \\ A\mathrm{d}y - B\mathrm{d}x & B\mathrm{d}y - C\mathrm{d}x \end{pmatrix} = 0, \tag{11.19}$$

which when expanded out gives the pair of characteristic curves

$$\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)_\pm = \frac{B \pm \sqrt{B^2 - AC}}{A}. \tag{11.20}$$

So, the second-order PDE (11.16) can be classified according to the sign of $B^2 - AC$:

(1) If $B^2 - AC > 0$ then two real characteristics exist and the equation is *hyperbolic*. The prototypical example is the wave equation,

$$\frac{\partial^2\Phi}{\partial t^2} - c^2\frac{\partial^2\Phi}{\partial x^2} = 0, \tag{11.21}$$

which has characteristics $x = \text{const} \pm ct$. The full solution $\Phi(x, t)$ can be found by marching along characteristic curves given appropriate initial conditions, such as the value of Φ along some open curve in the (t, x) plane and its derivative normal to the curve. This curve should not be a characteristic.

(2) If $B^2 - AC = 0$ then there is a single real characteristic. The equation is *parabolic*. The prototypical example is the diffusion equation,

$$\frac{\partial\Phi}{\partial t} = D\frac{\partial^2\Phi}{\partial x^2}. \tag{11.22}$$

Although the equation cannot be solved by following characteristic curves, the solution can still be built up by oozing away from given initial conditions: either the value of Φ along an open curve or its normal derivative.

(3) If $B^2 - AC < 0$ then there are no real characteristics and the equation is called *elliptic*. The prototypical example is Laplace's equation,

$$\frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} = 0. \tag{11.23}$$

This cannot be solved by marching off from an open initial condition curve in the (x, y) plane. But its solution is unique if we specify either Φ or its normal derivative $\partial\Phi/\partial n$ along a *closed* curve.

This classification assumes that the sign of $B^2 - AC$ does not change with (x, y).

Some terminology you might encounter: a *Dirichlet* boundary condition gives the value of Φ along a curve in the (x, y) plane, while a *Neumann* boundary condition gives its derivative normal to the curve. A *Cauchy* boundary condition gives both.

Exercise: For an elliptic second-order linear PDE, explain why we may give *either* Neumann or Dirichlet boundary conditions along the closed boundary curve, but not both simultaneously. [Hint: break the boundary curve into two segments.]

12 Laplace's equation by separation of variables

The derivations presented in the next few sections are slow and (mostly) careful, proceeding only in baby steps. They'll probably put you to sleep. Nevertheless, they might prove useful if there are details of the procedure you don't understand.

12.1 Example: Laplace's equation in Cartesian co-ordinates

The steady-state temperature distribution $T(x,y)$ within a semi-infinite metal sheet of width L satisfies Laplace's equation, $\nabla^2 T = 0$, with boundary conditions

(i) the temperature at the edges $T(0,y) = T(L,y) = 0$,
(ii) $T(x,y) \to 0$ as $y \to \infty$, and
(iii) $T(x,0) = T_0$.

What is $T(x,y)$ within the plate?

General solution We try a solution of the form

$$T(x,y) = X(x)Y(y), \tag{12.1}$$

in which case Laplace's equation becomes

$$Y\frac{\mathrm{d}^2X}{\mathrm{d}x^2} + X\frac{\mathrm{d}^2Y}{\mathrm{d}y^2} = 0. \tag{12.2}$$

Dividing both sides by XY gives

$$\frac{1}{X}\frac{\mathrm{d}^2X}{\mathrm{d}x^2} + \frac{1}{Y}\frac{\mathrm{d}^2Y}{\mathrm{d}y^2} = 0, \tag{12.3}$$

in which the first term depends only on x, and the second only on y. Therefore we must have

$$\frac{1}{X}\frac{\mathrm{d}^2X}{\mathrm{d}x^2} = -\frac{1}{Y}\frac{\mathrm{d}^2Y}{\mathrm{d}y^2} = -k^2, \tag{12.4}$$

where $-k^2$ is (in general) some complex constant. We could write this **separation constant** as $+k^2$ or even just k, but choosing $-k^2$ simplifies the following.

We are left with two ODEs (both eigenvalue equations)

$$\begin{aligned} \frac{\mathrm{d}^2X}{\mathrm{d}x^2} &= -k^2X, \\ \frac{\mathrm{d}^2Y}{\mathrm{d}y^2} &= +k^2Y, \end{aligned} \tag{12.5}$$

for which the (k-dependent) general solutions are

$$\begin{aligned} X_k(x) &= \begin{cases} A_0 + B_0x, & \text{if } k=0, \\ A_k\cos kx + B_k\sin kx, & \text{otherwise,} \end{cases} \\ Y_k(y) &= \begin{cases} C_0 + D_0y, & \text{if } k=0, \\ C_k\mathrm{e}^{ky} + D_k\mathrm{e}^{-ky}, & \text{otherwise,} \end{cases} \end{aligned} \tag{12.6}$$

where A_k, B_k, C_k and D_k are some (possibly complex) constants.

Laplace's equation is linear and homogeneous. Therefore a more general solution to $\nabla^2 T = 0$ is

$$T(x,y) = \sum_k X_k(x)Y_k(y) = [A_0 + B_0x][C_0 + D_0y] + \sum_{k\neq 0}[A_k\cos kx + B_k\sin kx][C_k\mathrm{e}^{ky} + D_k\mathrm{e}^{-ky}]. \tag{12.7}$$

Comments:

(i) If we had chosen $+K^2$ instead of $-k^2$ as our separation constant then $X_K(x)$ would be a sum of exponentials and $Y_K(y)$ would involve sines and cosines. Both sets of solutions are equivalent because $k = \pm iK$. Looking at the bcs, we guess that T will decay exponentially as $y \to \infty$ and has to vanish at $x = 0$ and $x = L$, suggesting trigonometric series in x. Therefore to keep the subsequent algebra as simple as possible, we label our constant $-k^2$.

(ii) The separation constant is $-k^2$. This means that if we include a term with, say $k = 2$ in (12.7) we do not need the term with $k = -2$: any A_{-2} can be absorbed into A_2, any C_{-2} into D_2 etc. We therefore assume that for any $X_k Y_k \neq 0$ in the series (12.7) the corresponding $X_{-k} Y_{-k} = 0$: in other words, each possible value of $-k^2$ appears at most once.

(iii) The $X_k(x)$ and the $Y_k(y)$ are then LI. That is, the only solution to $\sum_k c_k X_k(x) = 0$ or $\sum_k c_k Y_k(y) = 0$ is if all $c_k = 0$.

Application of boundary conditions First, let us use the condition that the temperature at the edges $T(0, y) = T(L, y) = 0$. Setting $x = 0$ in equation (12.7) gives

$$0 = \sum_k X_k(0) Y_k(y) = \sum_k A_k Y_k(y), \tag{12.8}$$

which, since the $Y_k(y)$ are LI, is satisfied only if all $A_k = 0$. For $x = L$ we have that

$$0 = \sum_k X_k(L) Y_k(y) = \sum_k B_k \sin(kL) Y_k(y), \tag{12.9}$$

which, by the same reasoning, requires that $B_0 = C_0 = D_0 = 0$ and that $B_k \sin(kL) = 0$ for $k \neq 0$. We satisfy the latter condition by imposing $k = n\pi/L$, where n is an integer. (If we were to choose any $k \neq n\pi/L$ we would have to set the corresponding $B_k = 0$ in order to satisfy the bc, so we have simply that $X_k(x) = 0$ when $k \neq n\pi/L$.) As our separation constant in (12.4) is actually k^2, we need only include $n > 0$.

Substituting these results into (12.7), the solution subject to the bcs on $x = 0$ and $x = L$ is

$$\begin{aligned} T(x, y) &= \sum_{n=1}^{\infty} B_n \sin\left(\frac{n\pi x}{L}\right) \left[C_n \mathrm{e}^{n\pi y/L} + D_n \mathrm{e}^{-n\pi y/L}\right] \\ &= \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x}{L}\right) \left[C_n \mathrm{e}^{n\pi y/L} + D_n \mathrm{e}^{-n\pi y/L}\right], \end{aligned} \tag{12.10}$$

where in the last line we have absorbed the constant B_n into C_n and D_n. We will show later that this is in fact the most general solution to $\nabla^2 T = 0$ subject to the condition that T vanishes on the edges of the sheet, $T(0, y) = T(L, y) = 0$.

Next we use the bc that $T(0, y) \to 0$ as $y \to \infty$. From (12.10),

$$\begin{aligned} 0 &= \lim_{y\to\infty} \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x}{L}\right) \left[C_n \mathrm{e}^{n\pi y/L} + D_n \mathrm{e}^{-n\pi y/L}\right] \\ &= \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x}{L}\right) \lim_{y\to\infty} \left[C_n \mathrm{e}^{n\pi y/L} + D_n \mathrm{e}^{-n\pi y/L}\right] \\ &= \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x}{L}\right) \lim_{y\to\infty} \left[C_n \mathrm{e}^{n\pi y/L}\right], \end{aligned} \tag{12.11}$$

giving all $C_n = 0$, because the functions $\sin(n\pi x/L)$ are LI.

Finally, we apply the condition that $T = T_0$ on the $y = 0$ end. Using (12.10) again,

$$T_0 = T(x, 0) = \sum_{n=1}^{\infty} D_n \sin\left(\frac{n\pi x}{L}\right). \tag{12.12}$$

In geometrical terms, this says that the coefficients $\{D_n\}$ are the co-ordinates of the function $T = T_0$ expressed in terms of the orthogonal basis $\sin(n\pi x/L)$. To find, say, D_m we simply take the scalar product of (12.12) with $\sin(m\pi x/L)$:

$$\begin{aligned}\int_0^L \sin\left(\frac{m\pi x}{L}\right) T_0\,\mathrm{d}x &= \sum_{n=1}^{\infty} D_n \int_0^L \sin\left(\frac{m\pi x}{L}\right)\sin\left(\frac{n\pi x}{L}\right)\,\mathrm{d}x \\ -\frac{T_0L}{m\pi}\left[\cos\left(\frac{m\pi x}{L}\right)\right]_0^L &= \sum_{n=1}^{\infty} D_n \frac{L}{2}\delta_{mn} \\ \frac{T_0L}{m\pi}\left[1-(-1)^m\right] &= \frac{L}{2}D_m \\ D_m &= \begin{cases}\frac{4T_0}{\pi m} & m \text{ odd,}\\ 0 & m \text{ even.}\end{cases}\end{aligned} \tag{12.13}$$

Substituting this into (12.10), our final solution subject to all the bcs is

$$T(x,y) = \sum_{m=1}^{\infty} \frac{4T_0}{\pi(2m-1)} \sin\left[\frac{(2m-1)\pi x}{L}\right] \exp\left[-\frac{(2m-1)\pi y}{L}\right], \tag{12.14}$$

in which $n = 2m-1$.

Alternative method Just in case you're curious, here's one way of showing that (12.10) is indeed the most general solution to our problem. You could use the following method as an alternative to separation of variables when solving problems, but I don't recommend it: see the comments below for why.

The problem: we want to find $T(x,y)$ in the sheet $0 < x < L$, $0 \le y < \infty$. Using the property of completeness of Fourier series, we can take a horizontal slice, $y = \text{const}$, across the sheet and write the temperature profile along the slice as

$$T(x,y) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n(y)\cos\left(\frac{n\pi x}{L}\right) + \sum_{n=1}^{\infty} b_n(y)\sin\left(\frac{n\pi x}{L}\right), \tag{12.15}$$

in which the coefficients a_n and b_n depend on which horizontal slice we're looking at. The values of a_n and b_n at different heights are related through Laplace's equation $\nabla^2 T = 0$. Substituting this $T(x,y)$ into Laplace's equation we obtain

$$0 = \nabla^2 T = \sum_{n=1}^{\infty}\left[\frac{\mathrm{d}^2 a_n}{\mathrm{d}y^2} - \frac{n^2\pi^2}{L^2}a_n\right]\cos\left(\frac{n\pi x}{L}\right) + \sum_{n=1}^{\infty}\left[\frac{\mathrm{d}^2 b_n}{\mathrm{d}y^2} - \frac{n^2\pi^2}{L^2}b_n\right]\sin\left(\frac{n\pi x}{L}\right) = 0. \tag{12.16}$$

The contents of each of the square brackets must vanish because the sines and cosines are LI. So we have

$$\frac{\mathrm{d}^2 a_n}{\mathrm{d}y^2} - \frac{n^2\pi^2}{L^2}a_n = 0, \tag{12.17}$$

which means $a_n(y) = F_n\exp(n\pi y/L) + G_n\exp(-n\pi y/L)$, where F_n and G_n are constants of integration. Similarly, $b_n(y) = H_n\exp(n\pi y/L) + I_n\exp(-n\pi y/L)$. Therefore our general solution to $\nabla^2 T = 0$ for $0 < x < L$ is

$$\begin{aligned}T(x,y) = \text{constant} &+ \sum_{n=1}^{\infty}\left[F_n\exp\left(\frac{n\pi y}{L}\right) + G_n\exp\left(-\frac{n\pi y}{L}\right)\right]\cos\left(\frac{n\pi x}{L}\right) \\ &+ \sum_{n=1}^{\infty}\left[H_n\exp\left(\frac{n\pi y}{L}\right) + I_n\exp\left(-\frac{n\pi y}{L}\right)\right]\sin\left(\frac{n\pi x}{L}\right).\end{aligned} \tag{12.18}$$

Applying the bc $T = 0$ at $x = 0$ and $x = L$ – remember that the constant and the exponentials form an LI set – gives eq (12.10).

Comments

1. Recall that the usual Fourier-series expansion assumes periodicity (i.e., that the function is defined along the circumference of a circle). Therefore we would have been in trouble had our bc required that $T(x,0) \neq T(x,L)$. Take a look at §21.2 of RHB to see one way of circumventing this.
2. To avoid this problem we could try to be clever and expand the temperature along our constant-y slice as, e.g., a Legendre series:

$$T(x,y) = \sum_{l=0}^{\infty} a_l(y) P_l\left(\frac{x+L}{2L}\right), \tag{12.19}$$

where the argument $(x+L)/2L$ to the Legendre polynomial P_l lies in the range $[-1,1]$. It is far from easy, however, to solve for $a_l(y)$ when this $T(x,y)$ is substituted into Laplace's equation. In constrast, the method of separation of variables (usually) reduces the problem to familiar ODEs that we know how to solve.

12.2 Laplace's equation in plane polar co-ordinates

In polar co-ordinates (R,ϕ) Laplace's equation is

$$\frac{1}{R}\frac{\partial}{\partial R}\left[R\frac{\partial V}{\partial R}\right] + \frac{1}{R^2}\frac{\partial^2 V}{\partial \phi^2} = 0. \tag{12.20}$$

Substituting a trial solution of the form $V(R,\phi) = V_R(R)V_\phi(\phi)$ into (12.20) gives

$$\begin{aligned}
&\frac{V_\phi}{R}\frac{\mathrm{d}}{\mathrm{d}R}\left[R\frac{\mathrm{d}V_R}{\mathrm{d}R}\right] + V_R\frac{1}{R^2}\frac{\mathrm{d}^2V_\phi}{\mathrm{d}\phi^2} = 0\\
\text{(Multiply by } R^2/V_RV_\phi\text{)}\quad &\frac{R}{V_R}\frac{\mathrm{d}}{\mathrm{d}R}\left[R\frac{\mathrm{d}V_R}{\mathrm{d}R}\right] + \frac{1}{V_\phi}\frac{\mathrm{d}^2V_\phi}{\mathrm{d}\phi^2} = 0\\
&\frac{R}{V_R}\frac{\mathrm{d}}{\mathrm{d}R}\left[R\frac{\mathrm{d}V_R}{\mathrm{d}R}\right] = -\frac{1}{V_\phi}\frac{\mathrm{d}^2V_\phi}{\mathrm{d}\phi^2} = m^2,
\end{aligned} \tag{12.21}$$

where m^2 is a separation constant. The equation for $V_\phi(\phi)$,

$$\frac{\mathrm{d}^2V_\phi}{\mathrm{d}\phi^2} = -m^2V_\phi, \tag{12.22}$$

has solutions

$$V_\phi^{(m)}(\phi) = \begin{cases} A_0 + B_0\phi, & \text{if } m = 0,\\ A_m\cos m\phi + B_m\sin m\phi, & \text{if } m \neq 0.\end{cases} \tag{12.23}$$

We require that $V_\phi(\phi) = V_\phi(\phi + 2\pi)$ because ϕ is an angular co-ordinate. Looking at (12.23), this means that if $m=0$ then we must have $B_0 = 0$, whereas if $m \neq 0$ then m must be an integer. As the separation constant is m^2, we need only consider one sign of m. Therefore, without loss of generality, we have that

$$V_\phi^{(m)}(\phi) = A_m\cos m\phi + B_m\sin m\phi \tag{12.24}$$

for $m = 0, 1, 2, \ldots$.

By inspection, the solutions to the radial equation,

$$R\frac{\mathrm{d}}{\mathrm{d}R}R\frac{\mathrm{d}V_R}{\mathrm{d}R} = m^2V_R, \tag{12.25}$$

are given by

$$V_R^{(m)}(R)=\begin{cases}C_0+D_0\log R, & \text{if } m=0,\\ C_mR^m+D_mR^{-m}, & \text{if } m>0.\end{cases}\tag{12.26}$$

Therefore, in plane polar co-ordinates the general solution $V(R,\phi)$ to Laplace's equation, $\nabla^2V=0$, is

$$\begin{aligned}V(R,\phi)&=\sum_{m=0}^{\infty}V_R^{(m)}(R)V_\phi^{(m)}(\phi)\\ &=C_0+D_0\log R+\sum_{m=1}^{\infty}\left[C_mR^m+D_mR^{-m}\right]\left[A_m\cos m\phi+B_m\sin m\phi\right].\end{aligned}\tag{12.27}$$

Example: earthed rod in uniform electric field

An infinite rod of radius a is earthed and placed in a uniform electric field $(E_x,E_y,E_z)=(E,0,0)$, with the axis of the rod coincident with the Oz axis. The boundary conditions on the unknown potential $V(R,\phi)$ are
(i) $V(R,\phi)\to -ER\cos\phi$ as $R\to\infty$ and
(ii) $V(R,\phi)=0$ on $R=a$.

Taking the general solution (12.27) and applying the first bc, we have that

$$C_0+D_0\log R+\sum_{m=1}^{\infty}C_mR^m[A_m\cos m\phi+B_m\sin m\phi]\to -ER\cos\phi,\tag{12.28}$$

in which we've dropped any term in V that decreases as $R\to\infty$. There are two ways to find the coefficients A_m, B_m etc. One is to rearrange (12.28) as a sum of LI basis functions (i.e., $\sin n\phi$ and $\cos n\phi$) and then to argue that the coefficient multiplying each basis function must be zero. The more powerful alternative is to project (12.28) along each of the basis functions to pick off the coefficients one by one. Taking the scalar product of both sides of (12.28) with $\sin n\phi$ for $n=1,2,3,\ldots$, gives

$$\begin{aligned}\sum_{m=1}^{\infty}C_mR^mB_m\int_0^{2\pi}\sin n\phi\sin m\phi\,\mathrm{d}\phi\to 0,\quad\text{or}\\ \sum_{m=1}^{\infty}C_mR^mB_m\pi\delta_{nm}\to 0,\end{aligned}\tag{12.29}$$

which can only be satisfied if all $C_mB_m=0$. Next take the scalar product of (12.28) with $\cos n\phi$ for $n=0,1,2,\ldots$. For $n=0$ we have that

$$C_0+D_0\log R\to 0,\tag{12.30}$$

so that $D_0=0$. For $n\geq 1$,

$$\begin{aligned}\sum_{m=1}^{\infty}C_mR^mA_m\int_0^{2\pi}\cos n\phi\cos m\phi\to -ER\int_0^{2\pi}\cos n\phi\cos\phi\,\mathrm{d}\phi,\quad\text{or}\\ \sum_{m=1}^{\infty}C_mR^mA_m\pi\delta_{nm}\to -ER\pi\delta_{n1},\end{aligned}\tag{12.31}$$

so that all $C_mA_m=0$, except for $C_1A_1=-E$. Substituting these results into (12.27), our potential has become

$$V(R,\phi)=C_0-ER\cos\phi+\sum_{m=1}^{\infty}D_mA_mR^{-m}\cos m\phi+\sum_{m=1}^{\infty}D_mR^{-m}B_m\sin m\phi\tag{12.32}$$

Applying the bc $V(a,\phi)=0$, we have that

$$0=V(a,\phi)=C_0-Ea\cos\phi+\sum_{m=1}^{\infty}D_mA_ma^{-m}\cos m\phi+\sum_{m=1}^{\infty}D_ma^{-m}B_m\sin m\phi. \tag{12.33}$$

Taking the scalar product of (12.33) with $\cos n\phi$ tells us that $C_0=0$ (from the $n=0$ case) and that all $D_mA_m=0$ except for $D_1A_1=Ea^2$. Similarly, the scalar product with $\sin n\phi$ gives all $D_mB_m=0$. Our solution to Laplace's equation subject to both bcs is therefore

$$V(R,\phi)=-E\left(R-\frac{a^2}{R}\right)\cos\phi. \tag{12.34}$$

12.3 Laplace's equation in spherical polar co-ordinates

Finally, what is the most general $V(r,\theta,\phi)$ that solves Laplace's equation in spherical polar co-ordinates,

$$\nabla^2V=\frac{1}{r^2}\frac{\partial}{\partial r}\left[r^2\frac{\partial V}{\partial r}\right]+\frac{1}{r^2\sin^2\theta}\frac{\partial^2V}{\partial\phi^2}+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left[\sin\theta\frac{\partial V}{\partial\theta}\right]=0? \tag{12.35}$$

As usual, we first try a solution of the form $V=V_r(r)V_\theta(\theta)V_\phi(\phi)$. Substituting this trial solution into Laplace's equation (12.35), multiplying by $r^2\sin^2\theta/V_rV_\theta V_\phi$ and rearranging, we find that

$$\underbrace{\frac{\sin^2\theta}{V_r}\frac{\mathrm{d}}{\mathrm{d}r}\left[r^2\frac{\mathrm{d}V_r}{\mathrm{d}r}\right]+\frac{\sin\theta}{V_\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left[\sin\theta\frac{\mathrm{d}V_\theta}{\mathrm{d}\theta}\right]}_{\text{function of }(r,\theta)\text{ only}}=\underbrace{-\frac{1}{V_\phi}\frac{\mathrm{d}^2V_\phi}{\mathrm{d}\phi^2}}_{\text{fn }\phi\text{ only}}=m^2, \tag{12.36}$$

where m^2 is a separation constant. By the same reasoning we used earlier in the plane-polar case, we impose $V_\phi(\phi+2\pi)=V_\phi(\phi)$, which means that the general solution to $V_\phi(\phi)$ is

$$V_\phi(\phi)=A\cos m\phi+B\sin m\phi, \tag{12.37}$$

where $m=0,1,2,3,\dots$ is a non-negative integer.

To find $V_r(r)$ and $V_\theta(\theta)$, divide (12.36) by $\sin^2\theta$ and rearrange slightly to obtain

$$\underbrace{-\frac{1}{V_r}\frac{\mathrm{d}}{\mathrm{d}r}\left[r^2\frac{\mathrm{d}V_r}{\mathrm{d}r}\right]}_{\text{only }r}=\underbrace{\frac{1}{V_\theta\sin\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left[\sin\theta\frac{\mathrm{d}V_\theta}{\mathrm{d}\theta}\right]-\frac{m^2}{\sin^2\theta}}_{\text{only }\theta}=-l(l+1), \tag{12.38}$$

in which $l(l+1)$ is an inspired choice of separation constant. The equation for V_θ is therefore

$$\begin{gathered}\frac{1}{\sin\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left[\sin\theta\frac{\mathrm{d}V_\theta}{\mathrm{d}\theta}\right]-\frac{m^2}{\sin^2\theta}V_\theta=-l(l+1)V_\theta,\quad\text{or}\\ \left[\frac{\mathrm{d}}{\mathrm{d}x}\left((1-x^2)\frac{\mathrm{d}}{\mathrm{d}x}\right)-\frac{m^2}{1-x^2}\right]V_\theta=-l(l+1)V_\theta,\end{gathered} \tag{12.39}$$

where in the last line we've substituted $x=\cos\theta$, so that $\frac{\mathrm{d}}{\mathrm{d}\theta}=-\sin\theta\frac{\mathrm{d}}{\mathrm{d}x}$. This equation for V_θ is the **associated Legendre equation** of §8.2. The eigenfunctions (i.e., the V_θ) are **associated Legendre functions**, $P_l^m(\cos\theta)$, and exist only for $l=0,1,\dots,$ and $|m|\le l$.

Finally, the V_r equation becomes

$$\frac{\mathrm{d}}{\mathrm{d}r}\left[r^2\frac{\mathrm{d}V_r}{dr}\right]=l(l+1)V_r, \tag{12.40}$$

which, by inspection, has a general solution

$$V_r(r) = Cr^l + Dr^{-(l+1)}. \tag{12.41}$$

Putting this together, we have that

$$V_r V_\theta V_\phi = \left[Cr^l + Dr^{-(l+1)}\right] P_l^m(\cos\theta)[A\cos m\phi + B\sin m\phi] \tag{12.42}$$

is one solution to Laplace's equation for $l = 0, 1, 2, \ldots$ and $m = 0, 1, 2, \ldots, l$. Because Laplace's equation is linear and homogeneous, we can superpose solutions and write

$$V(r,\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=0}^{l}\left[C_{lm}r^l + D_{lm}r^{-(l+1)}\right] P_l^m(\cos\theta)[A_{lm}\cos m\phi + B_{lm}\sin m\phi], \tag{12.43}$$

in which the constants of integration A, B, C, D depend on the choice of l and m.

> **Exercise:** Show that (12.43) is the most general solution to Laplace's equation. (Hint: first use the property of completeness of SL eigenfunctions to show that any $V(r = \text{const}, \theta, \phi)$ can be expressed as $\sum_{lm} P_l^m(\cos\theta)[F_{lm}(r)\cos m\phi + G_{lm}(r)\sin m\phi]$. Then substitute this expression into Laplace's equation to obtain ODEs for $F_{lm}(r)$ and $G_{lm}(r)$.)

In problems with axisymmetry, $V = V(r,\theta)$, we have that all $B_{lm} = 0$ and the only A_{lm} that are non zero are those with $m = 0$. Then the solution to Laplace's equation becomes

$$V(R,\theta) = \sum_{l=0}^{\infty}\left[C_l r^l + D_l r^{-(l+1)}\right] P_l(\cos\theta) \tag{12.44}$$

because $P_l^0(x) = P_l(x)$, the l^{th}-order Legendre polynomial (see §8.1).

Example: Earthed sphere in a uniform electric field

An earthed sphere is placed in a uniform electric field $\mathbf{E} = E\hat{\mathbf{z}}$. There are no charges outside the sphere, so the electric potential V satisfies Laplace's equation, $\nabla^2 V = 0$, with bcs (i) $V \to -Er\cos\theta$ as $r \to \infty$ and (ii) $V(r,\theta) = 0$ on the surface of the sphere $r = a$.

The first bc can also be written $V \to -ErP_1(\cos\theta)$ as $r \to \infty$. From the general solution (12.44) we than have that, as $r \to \infty$,

$$\sum_{l=0}^{\infty} C_l r^l P_l(\cos\theta) \to -ErP_1(\cos\theta). \tag{12.45}$$

To find the coefficients C_i, take the scalar product[†] of both sides with $P_m(\cos\theta)$: that is, multiply both sides by $P_m(\cos\theta)$ and integrate $\mathrm{d}(\cos\theta)$. The result is

$$\begin{aligned}\sum_{l=0}^{\infty} C_l r^l \int_{-1}^{1} P_m(\cos\theta)P_l(\cos\theta)\,\mathrm{d}(\cos\theta) &\to -Er\int_{-1}^{1} P_m(\cos\theta)P_1(\cos\theta)\,\mathrm{d}(\cos\theta),\\ \sum_{l=0}^{\infty} C_l r^l \frac{2}{2m+1}\delta_{lm} &\to -Er\frac{2}{2m+1}\delta_{m1},\end{aligned} \tag{12.46}$$

where in the last line we have used the orthogonality relation for the P_l, namely,

$$\int_{-1}^{1} P_l(x)P_m(x)\,\mathrm{d}x = \frac{2}{2l+1}\delta_{lm}. \tag{12.47}$$

[†] An alternative method is of course to exploit the linear independence of the P_l.

Therefore all $C_l = 0$, except for $C_1 = -E$.

Similarly, applying the second bc, $V(r,\theta) = 0$ on $r = a$, gives

$$0 = \underbrace{(-Ea + D_1 a^{-2})P_1(\cos\theta)}_{l=1} + \sum_{l\neq 1} D_l a^{-(l+1)} P_l(\cos\theta). \tag{12.48}$$

As the P_l are linearly independent, we must have that $D_1 = a^3 E$ and all other $D_l = 0$. Our final solution for the potential outside the sphere is

$$V(r,\theta) = -Er\left(1 - \frac{a^3}{r^3}\right) P_1(\cos\theta). \tag{12.49}$$

12.4 Spherical harmonics

Spherical harmonics are defined for $l = 0, 1, 2, \ldots$ and $|m| \le l$ via

$$Y_{lm}(\theta,\phi) = \begin{cases} \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) \mathrm{e}^{\mathrm{i}m\phi}, & m \ge 0, \\ (-1)^{|m|} Y^\star_{l|m|}(\theta,\phi), & m < 0. \end{cases} \tag{12.50}$$

Notice that they're linear combinations of the angular part,

$$P_l^m(\cos\theta)[A\cos m\phi + B\sin m\phi], \tag{12.51}$$

of the solutions (12.42) to Laplace's equation. Therefore another way of writing the general solution to Laplace's equation $\nabla^2 V = 0$, eq. (12.43), in spherical polar co-ordinates is

$$V(r,\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} \left[C_{lm} r^l + D_{lm} r^{-(l+1)}\right] Y_{lm}(\theta,\phi). \tag{12.52}$$

Examples

For reference, here are the first few spherical harmonics:

$$Y_{00} = \frac{1}{\sqrt{4\pi}}; \quad Y_{1,-1} = \sqrt{\frac{3}{8\pi}} \sin\theta \mathrm{e}^{-\mathrm{i}\phi}; \quad Y_{10} = \sqrt{\frac{3}{4\pi}}\cos\theta; \quad Y_{11} = -\sqrt{\frac{3}{8\pi}}\sin\theta \mathrm{e}^{\mathrm{i}\phi}. \tag{12.53}$$

Important properties

By construction, the $Y_{lm}(\theta,\phi)$ are **eigenfunctions** both of $r^2\nabla^2$ (i.e., the angular terms in ∇^2) and of $\frac{\partial}{\partial\phi}$:

$$\begin{aligned} r^2\nabla^2 Y_{lm} &= -l(l+1)Y_{lm}, \\ \frac{\partial}{\partial\phi} Y_{lm} &= mY_{lm}. \end{aligned} \tag{12.54}$$

They are **orthonormal**:

$$\int Y^\star_{l'm'}(\theta,\phi) Y_{lm}(\theta,\phi) \sin\theta \mathrm{d}\theta \mathrm{d}\phi = \delta_{l'l}\delta_{m'm}. \tag{12.55}$$

Exercise: use equation (8.25) to show that (i) the Y_{lm} are orthogonal and (ii) normalized.

The $Y_{lm}(\theta,\phi)$ are also **complete**: any well-behaved function $f(\theta,\phi)$ defined on the surface of a sphere can be expressed as

$$f(\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} c_{lm} Y_{lm}(\theta,\phi), \tag{12.56}$$

where the coefficients,

$$c_{lm} = \int Y^\star_{lm}(\theta,\phi) f(\theta,\phi) \sin\theta \mathrm{d}\theta \mathrm{d}\phi, \tag{12.57}$$

are the projections of f onto each Y_{lm}.

13 Helmholtz/wave equation: vibrations of a circular drum

Here is another example that involves two separation constants. The vertical displacement $u(R,\phi,t)$ of the surface of a circular drum satisfies the wave equation

$$\nabla^2 u - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = 0, \tag{13.1}$$

with the boundary condition that $u=0$ on the edge $R=a$ of the drum. What are the normal modes of the drum?

Let us separate variables in two steps. First we write $u(R,\phi,t)=U(R,\phi)T(t)$. Substituting this into the wave equation and separating variables, we obtain

$$\begin{aligned}\frac{\mathrm{d}^2T}{\mathrm{d}t^2} &= -\omega^2 T,\\ \nabla^2 U + \frac{\omega^2}{c^2}U &= 0,\end{aligned} \tag{13.2}$$

using ω^2 as the separation constant. The time-dependent factor in our assumed $u(R,\phi,t)$ is clearly $T(t)\propto \mathrm{e}^{\pm\mathrm{i}\omega t}$, while the spatial part $U(R,\phi)$ is given by the solution of the Helmholtz equation subject to the boundary condition that $U=0$ for $R=a$.

Making the further substitution $U(R,\phi)=U_R(R)U_\phi(\phi)$ to separate variables in the Helmholtz equation gives

$$\frac{U_\phi}{R}\frac{\mathrm{d}}{\mathrm{d}R}\left(R\frac{\mathrm{d}U_R}{\mathrm{d}R}\right) + \frac{U_R}{R^2}\frac{\mathrm{d}^2U_\phi}{\mathrm{d}\phi^2} + \frac{\omega^2}{c^2}U_RU_\phi = 0. \tag{13.3}$$

Multiplying by R^2/U_RU_ϕ, gives

$$\frac{R}{U_R}\frac{\mathrm{d}}{\mathrm{d}R}\left(R\frac{\mathrm{d}U_R}{\mathrm{d}R}\right) + \frac{1}{U_\phi}\frac{\mathrm{d}^2U_\phi}{\mathrm{d}\phi^2} + \frac{\omega^2}{c^2}R^2 = 0, \tag{13.4}$$

which separates into the two equations

$$\begin{aligned}\frac{\mathrm{d}^2U_\phi}{\mathrm{d}\phi^2} + m^2U_\phi &= 0,\\ R\frac{\mathrm{d}}{\mathrm{d}R}\left(R\frac{\mathrm{d}U_R}{\mathrm{d}R}\right) + \left(\frac{\omega^2}{c^2}R^2 - m^2\right)U_R &= 0,\end{aligned} \tag{13.5}$$

where m^2 is another separation constant. The first of these has solution $U_\phi \propto \mathrm{e}^{\mathrm{i}m\phi}$. The periodicity condition $U_\phi(\phi+2\pi)=U_\phi(\phi)$ restricts m to be an integer: $m=0,\pm1,\pm2,\ldots$.

> **Exercise:** The separation constant in these equations is m^2, not m. Why do we nevertheless need to include both signs of m in our general solution for $U_\phi(\phi)$?

The second equation is Bessel's equation (8.39). Letting $x\equiv\omega R/c$, it becomes

$$x\frac{\mathrm{d}}{\mathrm{d}x}\left(x\frac{\mathrm{d}U_R}{\mathrm{d}x}\right) + (x^2-m^2)U_R = 0. \tag{13.6}$$

The solutions to this equation that are well-behaved at the origin $x=0$ are the Bessel functions $U_R(x)=J_m(x)$.

General solution By the linearity and homogeneity of the wave equation, we can superpose solutions with different separation constants m and ω to obtain the general solution

$$u(R,\phi,t) = \sum_{m=-\infty}^{\infty}\sum_{\omega} A_{\omega,m}J_m\left(\frac{\omega R}{c}\right)\mathrm{e}^{\mathrm{i}m\phi}\mathrm{e}^{\pm\mathrm{i}\omega t}, \tag{13.7}$$

where the coefficients $A_{\omega,m}$ are set by the boundary conditions of the problem.

Boundary conditions The treatment of the boundary condition in most problems involving Bessel functions differs slightly from the previous problems we have tackled. In the present problem we have that $u(R=a,\phi,t)=0$, which means that

$$0=\sum_{m=-\infty}^{\infty}\sum_{\omega}A_{\omega,m}J_m\left(\frac{\omega a}{c}\right)\mathrm{e}^{\mathrm{i}m\phi}\mathrm{e}^{\pm\mathrm{i}\omega t}. \tag{13.8}$$

The linear independence of the $e^{\mathrm{i}m\phi}$ and $\mathrm{e}^{\pm\mathrm{i}\omega t}$ factors means that this can hold only if $J_m(\omega a/c)=0$. The general solution subject to the boundary condition that $u=0$ on the edge $R=a$ is therefore

$$u(R,\phi,t)=\sum_{n=1}^{\infty}\sum_{m=-\infty}^{\infty}J_m\left(\frac{\omega_{mn}R}{c}\right)\mathrm{e}^{\mathrm{i}m\phi}\left(A^+_{mn}\mathrm{e}^{\mathrm{i}\omega_{mn}t}+A^-_{mn}\mathrm{e}^{-\mathrm{i}\omega_{mn}t}\right) \tag{13.9}$$

where ω_{mn} is the n^{th} solution to $J_m(\omega_{mn}a/c)=0$. That is $\omega_{mn}=c\alpha_{mn}/a$, where the coefficients α_{mn} enumerate the zeros $J_m(\alpha_{mn})=0$ of the m^{th} Bessel function.

> **Exercise:** Explain why any choice of initial displacement $U_0(R,\phi)$ for which $U_0(a,\phi)=0$ can be represented by the series (13.9) with a suitable choice of coefficients A^+_{mn} and A^-_{mn}. Find expressions for these coefficients in terms of $U_0(R,\phi)$.

14 Cool down: heat equation

The temperature $\Theta(x,t)$ of a semi-infinite wall $(x>0)$ satisfies the heat equation

$$\frac{\partial\Theta}{\partial t}=D\frac{\partial^2\Theta}{\partial x^2}, \tag{14.1}$$

where D is a constant, subject to the boundary condition that the temperature at the surface $x=0$ varies with time as $\Theta(x=0,t)=T_0+T_1\cos\omega t$. How does $\Theta(x,t)$ vary inside the wall?

To answer this, let us try a solution of the form $\Theta(x,t)=X(x)T(t)$. Substituting this into the heat equation gives

$$X\frac{\mathrm{d}T}{\mathrm{d}t}=DT\frac{\mathrm{d}^2X}{\mathrm{d}x^2}. \tag{14.2}$$

Dividing by DTX results in

$$\frac{1}{DT}\frac{\mathrm{d}T}{\mathrm{d}t}=\frac{1}{X}\frac{\mathrm{d}^2X}{\mathrm{d}x^2}, \tag{14.3}$$

in which the LHS depends only on t, not x, and the RHS depends only on x, not t. The only way this can happen is if they are both equal to some constant, $-\lambda^2$, where λ may be complex. Therefore the PDE (14.1) separates into the two ODEs,

$$\begin{aligned}\frac{\mathrm{d}T}{\mathrm{d}t}&=-(\lambda^2D)T,\\ \frac{\mathrm{d}^2X}{\mathrm{d}x^2}&=-\lambda^2X.\end{aligned} \tag{14.4}$$

The general solutions to these are $T_\lambda(t)\propto\exp(-\lambda^2Dt)$ and $X(x)=A_\lambda\mathrm{e}^{\mathrm{i}\lambda x}$.

> **Comment:** Notice that we have slightly simplified the expression for these solutions by writing the separation constant as $-\lambda^2$ instead of, say, λ. For any choice of separation constant $-\lambda^2$ there are *two* possible values of λ: one of these gives a solution $X(x)=A_{+\lambda}\mathrm{e}^{+\mathrm{i}\lambda x}$, the other to $X(x)=A_{-\lambda}\mathrm{e}^{-\mathrm{i}\lambda x}$.

Absorbing the constant of proportionality in this $T_\lambda(t)$ into A_λ, the (λ-dependent) solution is

$$\Theta_\lambda(x,t) = A_\lambda \mathrm{e}^{\mathrm{i}\lambda x - \lambda^2 Dt}. \tag{14.5}$$

The heat equation is linear and homogeneous, so we can superpose solutions, giving the more general solution

$$\Theta(x,t) = \sum_\lambda A_\lambda \exp(\mathrm{i}\lambda x - \lambda^2 Dt). \tag{14.6}$$

Now we turn to boundary conditions. The condition that $\Theta(x=0,t) = T_0 + T_1 \cos\omega t$ means that

$$\sum_\lambda A_\lambda \mathrm{e}^{-\lambda^2 Dt} = T_0 + \frac{T_1}{2}\left(\mathrm{e}^{\mathrm{i}\omega t} + \mathrm{e}^{-\mathrm{i}\omega t}\right) \tag{14.7}$$

The RHS is clearly periodic. Imposing this same periodicity on the LHS implies that $-\lambda^2 D = \mathrm{i}n\omega$, where $n \in \mathbb{Z}$. Therefore

$$\lambda_n = \pm(1-\mathrm{i})\sqrt{\frac{n\omega}{2D}}. \tag{14.8}$$

We rely on an implicit boundary condition to choose the correct sign in this expression: we expect $\Theta(x) \not\to \infty$ as $x \to \infty$. This means that the real part of $\mathrm{i}\lambda$ must be less than 0, so that

$$\mathrm{i}\lambda_n = \begin{cases} -(1+\mathrm{i})\sqrt{\frac{|n|\omega}{2D}}, & \text{if } n > 0, \\ -(1-\mathrm{i})\sqrt{\frac{|n|\omega}{2D}}, & \text{if } n < 0. \end{cases} \tag{14.9}$$

Now equation (14.7) becomes

$$\sum_{n=-\infty}^{\infty} A_n \mathrm{e}^{\mathrm{i}n\omega t} = T_0 + \frac{T_1}{2}\left(\mathrm{e}^{\mathrm{i}\omega t} + \mathrm{e}^{-\mathrm{i}\omega t}\right). \tag{14.10}$$

Exploiting the linear independence of the $\mathrm{e}^{\mathrm{i}n\omega t}$ for different n, we have that $A_0 = T_0$ and $A_1 = A_{-1} = T_1/2$, with all other $A_n = 0$. The full solution (14.6) is

$$\begin{aligned} \Theta(x,t) &= T_0 + \frac{T_1}{2}\left[\exp\left(-\frac{1+\mathrm{i}}{\delta}x + \mathrm{i}\omega t\right) + \exp\left(-\frac{1-\mathrm{i}}{\delta}x - \mathrm{i}\omega t\right)\right] \\ &= T_0 + T_1 \mathrm{e}^{-x/\delta}\cos\left(\omega t - \frac{x}{\delta}\right), \end{aligned} \tag{14.11}$$

where the skin depth $\delta \equiv \sqrt{2D/\omega}$.

15 Inhomogeneous PDEs: Green's functions[⋆]

By analogy with §9, suppose that we want to solve the linear, second-order inhomogeneous PDE

$$\hat{D}\psi = f, \tag{15.1}$$

for $\psi(\mathbf{x})$ given $f(\mathbf{x})$. We'll assume that the operator $\hat{D}$ is of the form

$$\hat{D} = \frac{1}{w(\mathbf{x})}\left[\nabla\cdot(p(\mathbf{x})\nabla) + q(\mathbf{x})\right]. \tag{15.2}$$

and that the boundary conditions are such that either ψ (Dirichlet) or its normal derivative (Neumann) vanish on the boundary of the region under consideration.

> **Exercise:** Use Gauss' theorem plus the identity $\nabla\cdot(u\nabla v) = u\nabla^2 v + (\nabla u)\cdot(\nabla v)$ to show that $\left\langle u, \hat{D}v\right\rangle = \left\langle v, \hat{D}u\right\rangle^\star$ when $u(\mathbf{x})$ and $v(\mathbf{x})$ satisfy these boundary conditions.

If we can find a function $G(\mathbf{x}_1, \mathbf{x}_2)$ that (i) satisfies the same boundary conditions as our solution $\psi(\mathbf{x}_1)$ and (ii) for which

$$\hat{D}_1 G(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{w(\mathbf{x}_1)}\delta(\mathbf{x}_1 - \mathbf{x}_2), \tag{15.3}$$

then the solution to (15.1) is given by

$$\psi(\mathbf{x}_1) = \int_V \mathrm{d}^3\mathbf{x}_2\, G(\mathbf{x}_1, \mathbf{x}_2) f(\mathbf{x}_2). \tag{15.4}$$

In (15.3) the operator $\hat{D}_1$ is given by the expression (15.2) for $\hat{D}$ with $\mathbf{x}$ replaced by $\mathbf{x}_1$.

The properties we derived in §9.1 for one-dimensional Green's functions all generalise in a straightforward way to the multi-dimensional case.

15.1 Examples

How to find $G(\mathbf{x}_1, \mathbf{x}_2)$? In §9.2 we constructed Green's functions for various one-dimensional problems by joining up linearly independent solutions to the homogeneous equation $A_x u = 0$. This procedure is less useful in multidimensional problems.

Example: Laplace by inspection Sometimes Green's functions can nevertheless be found by inspection. For example, consider $\hat{D} = \nabla^2$ in everyday three-dimensional space. In Cartesian coordinates the natural weight function $w(\mathbf{x}) = 1$, and then our $\hat{D} = \nabla^2$ is of the form (15.2) with $p(\mathbf{x}) = 1$ and $q(\mathbf{x}) = 0$. G has to satisfy

$$\nabla_1^2 G(\mathbf{x}_1, \mathbf{x}_2) = \delta(\mathbf{x}_1 - \mathbf{x}_2). \tag{15.5}$$

We assume Dirichlet bcs, so that $G \to 0$ at infinity. Integrating over a volume V that includes the point source at $\mathbf{x} = \mathbf{x}_2$ and applying Gauss's divergence theorem, we find that

$$\begin{aligned} \int_V \nabla_1\cdot\nabla_1 G(\mathbf{x}_1, \mathbf{x}_2)\mathrm{d}^3\mathbf{x}_1 &= \int_V \delta(\mathbf{x}_1 - \mathbf{x}_2)\mathrm{d}^3\mathbf{x}_1 \\ \Rightarrow \quad \int_V \nabla_1 G(\mathbf{x}_1, \mathbf{x}_2)\cdot\mathrm{d}^2\mathbf{S}_1 &= 1. \end{aligned} \tag{15.6}$$

If we take V to be a sphere of radius $r_{12} = |\mathbf{x}_1 - \mathbf{x}_2|$ centred on the point $\mathbf{x}_2$, this is satisfied by taking the integrand in the LHS to be

$$\frac{\partial}{\partial r_{12}} G(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{4\pi r_{12}^2}. \tag{15.7}$$

[⋆] Bonus material: examples of the application of the Dirac delta and Fourier transforms

Therefore, assuming Dirichlet bcs, the Green's function for the operator $\hat{D} = \nabla^2$ is $G = -1/4\pi r_{12}$, or

$$G(\mathbf{x}_1, \mathbf{x}_2) = -\frac{1}{4\pi} \frac{1}{|\mathbf{x}_1 - \mathbf{x}_2|}. \tag{15.8}$$

The obvious physical application is Poisson's equation, $\nabla^2 \psi = -\rho/\epsilon_0$. The solution that satisfies $\psi \to 0$ as $r \to \infty$ is therefore

$$\begin{aligned} \psi(\mathbf{x}_1) &= \int_V \mathrm{d}^3\mathbf{x}_2 G(\mathbf{x}_1, \mathbf{x}_2) \left(-\frac{\rho(\mathbf{x}_2)}{\epsilon_0} \right) \\ &= \frac{1}{4\pi\epsilon_0} \int_V \frac{\rho(\mathbf{x}_2) \mathrm{d}^3\mathbf{x}_2}{|\mathbf{x}_1 - \mathbf{x}_2|}. \end{aligned} \tag{15.9}$$

Example: Laplace by Fourier Notice that this $\hat{D}$ and its boundary conditions are translationally invariant. That means that we might expect to be able to find $G(\mathbf{x}_1, \mathbf{x}_2) = G(\mathbf{x}_1 - \mathbf{x}_2)$ using integral transform methods, such as Fourier transforms. Taking advantage of the translational symmetry of ∇^2, the problem of finding $G(\mathbf{x}_1, \mathbf{x}_2)$ for Laplace's equation reduces to finding the single-variable function $G(\mathbf{x})$ that satisfies

$$\nabla^2 G(\mathbf{x}) = \delta(\mathbf{x}), \tag{15.10}$$

subject to the boundary condition that $G(\mathbf{x})$ vanishes at infinity. Fourier transforming this, we have that

$$-(k_x^2 + k_y^2 + k_z^2) \tilde{G}(\mathbf{k}) = \frac{1}{(2\pi)^3}, \tag{15.11}$$

so that

$$G(\mathbf{x}) = -\frac{1}{(2\pi)^3} \int \frac{\mathrm{d}^3\mathbf{k}}{\mathbf{k}^2} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}. \tag{15.12}$$

Changing to polar coordinates in $\mathbf{k}$ space, this becomes

$$\begin{aligned} G(\mathbf{x}) &= -\frac{1}{(2\pi)^2} \int_0^\infty \mathrm{d}k\, k^2 \int_0^\pi \mathrm{d}\theta\, \sin\theta \frac{1}{k^2} \mathrm{e}^{\mathrm{i}k|\mathbf{x}|\cos\theta} \\ &= -\frac{1}{(2\pi)^2} \int_0^\infty \mathrm{d}k\, \frac{2\sin k|\mathbf{x}|}{k|\mathbf{x}|} \\ &= -\frac{1}{2\pi^2 |\mathbf{x}|} \int_0^\infty \mathrm{d}z\, \frac{\sin kz}{z} = -\frac{1}{4\pi|\mathbf{x}|}. \end{aligned} \tag{15.13}$$

This is the response to a unit of "charge" placed at $\mathbf{x} = 0$. By translational invariance, the response at $\mathbf{x} = \mathbf{x}_1$ to a unit charge placed at $\mathbf{x}_2$ is $G(\mathbf{x}_1, \mathbf{x}_2) = G(\mathbf{x}_1 - \mathbf{x}_2)$, in agreement with (15.8).

Example: Electromagnetic waves In Lorenz gauge, the electromagnetic potentials $\phi(\mathbf{x}, t)$ and $\mathbf{A}(\mathbf{x}, t)$ are related to the charge and current densities ρ and $\mathbf{j}$ by the PDE

$$-\left[-\frac{1}{c^2} \frac{\partial^2}{\partial t^2} + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right] \begin{pmatrix} \phi \\ \mathbf{A} \end{pmatrix} = \begin{pmatrix} \rho/\epsilon_0 \\ \mu_0 \mathbf{j} \end{pmatrix}. \tag{15.14}$$

Notice that the operator on the left-hand side is unchanged under (Lorentz) translations. We assume vacuum boundary conditions. The Green's function $G(\mathbf{x}, t)$ should then satisfy

$$-\left[-\frac{1}{c^2} \frac{\partial^2}{\partial t^2} + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right] G(\mathbf{x}, t) = \delta(\mathbf{x})\delta(t). \tag{15.15}$$

Defining the Fourier transform[†]

$$\tilde{G}(\mathbf{k}, z) \equiv \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}^3\mathbf{x}\, \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{x}} \frac{1}{(2\pi)^{1/2}} \int \mathrm{d}t\, \mathrm{e}^{+\mathrm{i}zt} G(\mathbf{x}, t), \tag{15.16}$$

[†] Note that the sign used in the $t \to z$ transform differs from our usual definition. Conventions...

equation (15.15) becomes

$$-\left[\frac{z^2}{c^2}-\mathbf{k}^2\right]\tilde{G}(\mathbf{k},z)=\frac{1}{(2\pi)^2}. \tag{15.17}$$

Rearranging,

$$\tilde{G}(\mathbf{k},z)=\frac{c^2}{c^2\mathbf{k}^2-z^2}, \tag{15.18}$$

for which the inverse Fourier transform is

$$G(\mathbf{x},t)=\frac{1}{(2\pi)^4}\int\frac{c^2}{c^2\mathbf{k}^2-z^2}\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{x}-zt)}\mathrm{d}^3\mathbf{k}\mathrm{d}z. \tag{15.19}$$

Following the same procedure used in our solution to Laplace's equation, we change to polar coordinates in $\mathbf{k}$, obtaining

$$G(\mathbf{x},t)=\frac{c^2}{4\pi^3|\mathbf{x}|}\int_0^\infty\sin(k|\mathbf{x}|)\left(\int_{-\infty}^{\infty}\frac{k}{c^2k^2-z^2}\mathrm{e}^{-\mathrm{i}zt}\,\mathrm{d}z\right)\mathrm{d}k. \tag{15.20}$$

The integrand in brackets has poles at $z=\pm ck$. We can deform our integration contour along the real line to skirt above (z takes on a small positive imaginary part) or below ($\mathrm{Im}\,z<0$) these poles. Let's skirt above and call the resulting Green's function G^{ret}. Then, when $t<0$ we need to close the integration contour in the upper half plane so that the $\mathrm{e}^{-\mathrm{i}zt}$ factor is well behaved. This leaves no poles within the contour and so

$$G^{\mathrm{ret}}(\mathbf{x},t)=0,\quad t<0. \tag{15.21}$$

For $t>0$ we need to close the contour in the lower half plane, which means it encloses the two poles. The z integral becomes

$$\int\frac{k\mathrm{e}^{-\mathrm{i}zt}\,\mathrm{d}z}{c^2k^2-z^2}=-\frac{\mathrm{i}\pi}{c}\left[\mathrm{e}^{\mathrm{i}ckt}-\mathrm{e}^{-\mathrm{i}ckt}\right], \tag{15.22}$$

leading to

$$G^{\mathrm{ret}}(\mathbf{x},t)=\frac{c}{4\pi|\mathbf{x}|}\left[\delta(|\mathbf{x}|-ct)-\delta(|\mathbf{x}|+ct)\right],\quad t>0. \tag{15.23}$$

The condition on the sign of t means that the second Dirac delta cannot contribute and so our complete Green's function for the case in which we deform our integration contour to go above the two poles in the complex plane is given by

$$G^{\mathrm{ret}}(\mathbf{x},t)=\frac{c}{4\pi|\mathbf{x}|}\times\begin{cases}0, & t<0\\ \delta(|\mathbf{x}|-ct), & t>0.\end{cases} \tag{15.24}$$

This is known as the **retarded** Green's function for the electromagnetic wave equation. It corresponds to a spherical shell of light emanating at $t=0$ from $\mathbf{x}=0$, whose radius increases as $|\mathbf{x}|=ct$. Its amplitude varies with radius as $1/|\mathbf{x}|$, as one would expect for a potential.

Deforming our integration contour to go under the poles at $z=\pm ck$ yields the **advanced** Green's function

$$G^{\mathrm{adv}}(\mathbf{x},t)=\frac{c}{4\pi|\mathbf{x}|}\times\begin{cases}\delta(|\mathbf{x}|-ct), & t<0\\ 0, & t>0.\end{cases} \tag{15.25}$$

Appendix: back to linear algebra

http://www-thphys.physics.ox.ac.uk/people/JohnMagorrian/mm

john.magorrian@physics.ox.ac.uk

Appendix A: Tensor algebra[⋆]

Consider an n-dimensional *real* vector space $\mathcal{V}$. Let $\mathbf{e}_1, ..., \mathbf{e}_n$ be a basis for this space and introduce the linear maps $\mathbf{e}_\star^1, ..., \mathbf{e}_\star^n$ defined by $\mathbf{e}_\star^i \mathbf{e}_j \equiv \delta^i_j$.

> **Exercise:** Without assuming the existence of an inner product on $\mathcal{V}$, show that $\mathcal{V}$ and its dual $\mathcal{V}^\star$ are isomorphic and that the $\mathbf{e}_i^\star$ are a basis for $\mathcal{V}^\star$.

Any vectors $\mathbf{v} \in \mathcal{V}$ and $\mathbf{v}^\star \in \mathcal{V}^\star$ can be written as $\mathbf{v} = \sum_i a^i \mathbf{e}_i$ and $\mathbf{v}^\star = \sum_i a_i \mathbf{e}_\star^i$ respectively. Notice that each dummy index appears twice in these expressions: once upstairs (e.g., a^i) and once downstairs ($\mathbf{e}_i$): downstairs indices label basis vectors for $\mathcal{V}$ and the *components* of vectors from $\mathcal{V}^\star$; upstairs indices label basis vectors for $\mathcal{V}^\star$ and the *components* of vectors from $\mathcal{V}$.

A.1 Covectors eat vectors; vectors eat covectors

We have introduced covectors as linear maps from vectors to scalars. What about the dual $(\mathcal{V}^\star)^\star$ to the space of covectors $\mathcal{V}^\star$? As you might guess, it turns out that this $(\mathcal{V}^\star)^\star$ is essentially $\mathcal{V}$.

More precisely, there is a natural isomorphism between $\mathcal{V}$ and $(\mathcal{V}^\star)^\star$. For any $\mathbf{v} \in \mathcal{V}$, introduce a mapping $f(\mathbf{v}) : \mathcal{V}^\star \to \mathcal{F}$ by

$$f(\mathbf{v})(\omega) = \omega(\mathbf{v}). \tag{A.1}$$

This $f(\mathbf{v})(\omega)$ is a linear map from $\omega \in \mathcal{V}^\star$ to $\mathcal{F}$. It is therefore a member of the dual space to $\mathcal{V}^\star$: that is, $f(\mathbf{v}) \in (\mathcal{V}^\star)^\star$. Viewed as a function of $\mathbf{v}$, it is also clear that $f : \mathcal{V} \to (\mathcal{V}^\star)^\star$ is linear. It is also injective: for any nonzero $\mathbf{v} = \sum_i v_i \mathbf{e}_i \in \mathcal{V}$ we have $f(\mathbf{v})(\mathbf{e}_\star^i) = \mathbf{e}_\star^i(\mathbf{v}) = v_i$. Now $\mathcal{V}$, $\mathcal{V}^\star$ and $(\mathcal{V}^\star)^\star$ all have the same dimension. So this $f : \mathcal{V} \to (\mathcal{V}^\star)^\star$ is an isomorphism.

Just as elements of $\mathcal{V}^\star$ are linear maps from $\mathcal{V}$ to scalars, elements of $(\mathcal{V}^\star)^\star$ are linear maps from $\mathcal{V}^\star$ to $\mathcal{F}$. As $(\mathcal{V}^\star)^\star$ and $\mathcal{V}$ are isomorphic, we can treat elements of $\mathcal{V}$ as being linear maps from $\mathcal{V}^\star$ to scalars.

A.2 Definition of tensors

A **covariant tensor of rank** r or, simply a **covariant** r**-tensor** on $\mathcal{V}$ is a scalar-valued multilinear function of r elements of $\mathcal{V}$:

$$T : \underbrace{\mathcal{V} \times \cdots \times \mathcal{V}}_{r \text{ copies}} \to \mathcal{F}. \tag{A.2}$$

Some examples: scalars are covariant 0-tensors; members of the dual space $\mathcal{V}^\star$ are covariant 1-tensors; the scalar product of two real vectors is a covariant 2-tensor; the determinant of n real vectors from an n-dimensional vector space $\mathcal{V}$ is a covariant n-tensor.

If T is unchanged under exchange of two of its arguments – that is, $T(..., \mathbf{v}_i, ..., \mathbf{v}_j, ...) = T(..., \mathbf{v}_j, ..., \mathbf{v}_i, ...)$ – then T is **symmetric** with respect to that pair of arguments. Similarly, if swapping two arguments changes the sign of T then T is **antisymmetric** with respect to those two arguments. The determinant is an example of a tensor that is **completely antisymmetric** with respect to exchanges of any pair of its arguments.

[⋆] Bonus material

The set of all covariant r-tensors on $\mathcal{V}$, written $\mathcal{T}^r(\mathcal{V})$, is itself is a vector space under the usual operations of pointwise addition and scalar multiplication: for any $T, T' \in \mathcal{T}^r(\mathcal{V})$ the linear combination

$$(\alpha T + \alpha' T')(\mathbf{v}_1, ..., \mathbf{v}_k) = \alpha T(\mathbf{v}_1, ..., \mathbf{v}_k) + \alpha' T'(\mathbf{v}_1, ..., \mathbf{v}_k), \tag{A.3}$$

is itself another covariant r-tensor, with the sum and products on the right being the usual scalar operations. We construct a basis below.

A **contravariant s-tensor** is a scalar-valued multilinear function of s elements of $\mathcal{V}^\star$:

$$T : \underbrace{\mathcal{V}^\star \times \cdots \times \mathcal{V}^\star}_{s \text{ copies}} \to \mathcal{F}. \tag{A.4}$$

A vector $\mathbf{v} \in \mathcal{V}$ is a contravariant 1-tensor. The set of all contravariant s-tensors on $(\mathcal{V})$, written $\mathcal{T}_s(\mathcal{V})$ is itself a vector space under the operations of pointwise addition and scalar multiplication.

More generally, a **mixed $\binom{s}{r}$-tensor** is a multilinear map

$$T : \underbrace{\mathcal{V}^\star \times \cdots \times \mathcal{V}^\star}_{s \text{ copies}} \times \underbrace{\mathcal{V} \times \cdots \times \mathcal{V}}_{r \text{ copies}} \to \mathcal{F} \tag{A.5}$$

of s dual vectors and r vectors to scalars. Linear operators $A : \mathcal{V} \to \mathcal{V}$ are examples of type $\binom{1}{1}$ tensors: when fed a vector they emit another vector, which when fed with a covector emits a scalar. Covariant r-tensors are type $\binom{0}{r}$ tensors, while contravariant s-tensors are type $\binom{s}{0}$. Again, for given r and s the set of all $\binom{s}{r}$-tensors forms a vector space.

To construct a basis for these vector spaces we first need to introduce a way of constructing higher-rank tensors from vectors and covectors.

A.3 Tensor products

Tensor (or outer) product of two covariant 1-tensors Consider covariant 1-tensors $S, T \in \mathcal{V}^\star$ and define their tensor product to be the map $S \otimes T : \mathcal{V} \times \mathcal{V} \to \mathcal{F}$ given by

$$S \otimes T(\mathbf{a}, \mathbf{b}) = S(\mathbf{a})T(\mathbf{b}), \tag{A.6}$$

the product on the right being just ordinary multiplication of scalars. Linearity of S and T guarantees that $S \otimes T$ is a bilinear function of $\mathbf{a}$ and $\mathbf{b}$. Therefore it is a covariant 2-tensor.

> **Exercise:** Show that the tensor product so defined is associative. That is, $R \otimes (S \otimes T) = (R \otimes S) \otimes T$ for covariant 1-tensors R, S and T.

Generalisation Let $S \in \mathcal{T}^k$ and $T \in \mathcal{T}^l$. Then define their tensor product to be the map

$$S \otimes T : \underbrace{\mathcal{V} \times \cdots \times \mathcal{V}}_{k+l \text{ copies}} \to \mathcal{F} \tag{A.7}$$

given by

$$S \otimes T(\mathbf{v}_1, ..., \mathbf{v}_{k+l}) = S(\mathbf{v}_1, ..., \mathbf{v}_k)T(\mathbf{v}_{k+1}, ..., \mathbf{v}_{k+l}). \tag{A.8}$$

Multilinearity of S and T means that $S \otimes T$ is multilinear too. Therefore it is a covariant $(k + l)$-tensor.

> **Exercise:** Show that the tensor product defined by (A.8) is bilinear and associative: i.e., that $S \otimes T$ depends linearly on S and T and that $R \otimes (S \otimes T) = (R \otimes S) \otimes T$.

Exterior products of contravariant tensors and, more generally, of mixed-rank tensors are constructed in the same way. For example, the exterior product of a covector $S \in \mathcal{V}^\star$ and $\mathbf{v} \in \mathcal{V}$ is defined as

$$S \otimes \mathbf{v}(\mathbf{a}, B) = S(\mathbf{a})\mathbf{v}(B), \tag{A.9}$$

for any $\mathbf{a} \in \mathcal{V}$ and $B \in \mathcal{V}^\star$. This is a $\binom{1}{1}$-tensor.

A.4 Bases

Let $\mathbf{e}_\star^i$ be any basis for $\mathcal{V}^\star$. Then a basis for the space $\mathcal{T}^r(\mathcal{V})$ of covariant r-tensors is

$$\mathbf{e}_\star^{i_1} \otimes \cdots \otimes \mathbf{e}_\star^{i_r}, \qquad 1 \le i_1, ..., i_r \le n. \tag{A.10}$$

Therefore $\mathcal{T}^r(\mathcal{V})$ is an n^r-dimensional vector space.

Proof: Each of the objects (A.10) is clearly a covariant r-tensor. We need to show that that they span $\mathcal{T}^r(\mathcal{V})$ and are LI. First we show that they span the space. Take any $T \in \mathcal{T}^r(\mathcal{V})$ and consider the linear combination $\sum_{i_1 \ldots i_r} T_{i_1 \ldots i_r} \mathbf{e}_\star^{i_1} \otimes \cdots \otimes \mathbf{e}_\star^{i_r}$, with coefficients

$$T_{i_1 \ldots i_r} \equiv T(\mathbf{e}_{i_1}, ..., \mathbf{e}_{i_r}). \tag{A.11}$$

This linear combination of tensors is itself also a tensor. Applied to any set of basis vectors $(\mathbf{e}_{j_1}, ..., \mathbf{e}_{j_r})$ we have that

$$\begin{aligned}
\sum_{i_1=1}^{n} \cdots \sum_{i_r=1}^{n} T_{i_1 \ldots i_r} \mathbf{e}_\star^{i_1} \otimes \cdots \otimes \mathbf{e}_\star^{i_r}(\mathbf{e}_{j_1}, ..., \mathbf{e}_{j_r}) &= \sum_{i_1=1}^{n} \cdots \sum_{i_r=1}^{n} T_{i_1 \ldots i_r} \mathbf{e}_\star^{i_1}(\mathbf{e}_{j_1}) \otimes \cdots \otimes \mathbf{e}_\star^{i_r}(\mathbf{e}_{j_r}) \\
&= \sum_{i_1=1}^{n} \cdots \sum_{i_r=1}^{n} T_{i_1 \ldots i_r} \delta_{j_1}^{i_1} \cdots \delta_{j_r}^{i_r} \\
&= T_{j_1 \ldots j_r} \\
&= T(\mathbf{e}_{j_1}, ..., \mathbf{e}_{j_r}).
\end{aligned} \tag{A.12}$$

By multilinearity, any tensor is determined by its action on the basis vectors. Therefore we have shown that any $T \in \mathcal{T}^r(\mathcal{V})$ can be expressed as $T = \sum_{i_1 \ldots i_r} T_{i_1 \ldots i_r} \mathbf{e}_\star^{i_1} \otimes \cdots \otimes \mathbf{e}_\star^{i_r}$ with components $T_{i_1 \ldots i_r}$ given by (A.11): the tensors (A.10) span $\mathcal{T}^r(\mathcal{V})$.

To show that the tensors (A.10) are LI, suppose that there is some linear combination of them that equals zero, $\sum_{i_1 \ldots i_r} T_{i_1 \ldots i_r} \mathbf{e}_\star^{i_1} \otimes \cdots \otimes \mathbf{e}_\star^{i_r} = 0$. Applying this equation to the tensor $\mathbf{e}_{j_1} \otimes \cdots \otimes \mathbf{e}_{j_s}$ yields $T_{j_1 \ldots j_r} = 0$ for any choice of $(j_1, ..., j_r)$. Therefore the only solution to $\sum_{i_1 \ldots i_r} T_{i_1 \ldots i_r} \mathbf{e}_\star^{i_1} \otimes \cdots \otimes \mathbf{e}_\star^{i_r} = 0$ is when all $T_{i_1 \ldots i_r} = 0$: the tensors (A.10) are LI.

Similarly, a basis for the space $\mathcal{T}_s(\mathcal{V})$ of contravariant s-tensors is

$$\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_s}, \qquad 1 \le i_1, ..., i_s \le n, \tag{A.13}$$

and, more generally, a basis for the space of rank $\binom{s}{r}$ mixed tensors is

$$\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_s} \otimes \mathbf{e}_\star^{j_1} \otimes \cdots \otimes \mathbf{e}_\star^{j_r} \qquad 1 \le i_1, ..., i_s, j_1, ..., j_r \le n. \tag{A.14}$$

Therefore the dimension of the vector space of $\binom{s}{r}$-tensors is n^{r+s}.

A.5 Inner products: the metric tensor

All of our discussion of tensors to this point has been mere algebraic bookkeeping. In applications physics usually supplies an inner product on $\mathcal{V}$, which allows us to define lengths of vectors and the projection of one vector onto another. From a bookkeeping point of view, this inner product identifies a special isomorpism between vectors and covectors.

Recall that for real vector spaces the scalar product is a symmetric bilinear mapping from pairs of vectors to scalars. That is, it is a symmetric covariant 2-tensor, which is known as the **metric tensor** or simply "the **metric**". Given any basis $\mathbf{e}_1, ..., \mathbf{e}_n$ for $\mathcal{V}$ and corresponding $\mathbf{e}_\star^i$ for $\mathcal{V}^\star$ the metric tensor can be written

$$g = \sum_{ij} g_{ij} \mathbf{e}_\star^i \otimes \mathbf{e}_\star^j \tag{A.15}$$

with components

$$g_{ij} = g(\mathbf{e}_i, \mathbf{e}_j). \tag{A.16}$$

The inner product of the vectors **a** and **b** then is $\mathbf{a} \cdot \mathbf{b} = \langle \mathbf{a}, \mathbf{b} \rangle = g(\mathbf{a}, \mathbf{b})$.

In normal 3-space, the mathematical definition of a metric includes the condition that the metric is **positive definite**: $g(\mathbf{u}, \mathbf{u}) > 0$ for all $\mathbf{u} \neq 0$, equivalent to condition (2.4) in our definition of the inner product. Then a basis $\mathbf{e}_1, ..., \mathbf{e}_n$ is **orthonormal** iff $g(\mathbf{e}_i, \mathbf{e}_j) = \delta_{ij}$.

When applied to spacetime (relativity) we need to relax this to the weaker condition that the metric be merely **definite**: that is, if $g(\mathbf{u}, \mathbf{v}) = 0$ for all **v** implies that $\mathbf{u} = 0$, but $g(\mathbf{u}, \mathbf{u})$ can be any sign (or even zero). Such a g is sometimes called a pseudometric, but we broaden our definition of the word "metric" to encompass this important case. The orthonormality condition is then $g(\mathbf{e}_i, \mathbf{e}_j) = \pm\delta_{ij}$. For example, in (ct, x, y, z) Minkowski space the orthonormality condition is

$$g(\mathbf{e}_i, \mathbf{e}_j) = \eta_{ij} \equiv \mathrm{diag}(1, -1, -1, -1) \tag{A.17}$$

and a vector **u** is called timelike if $g(\mathbf{u}, \mathbf{u}) > 0$, lightlike if $g(\mathbf{u}, \mathbf{u}) = 0$ and spacelike if $g(\mathbf{u}, \mathbf{u}) < 0$.

Metric duality: raising and lowering indices Given **v**, then the object $g(\mathbf{v}, \bullet)$ is a covector. Expand $\mathbf{v} = \sum_k v^k \mathbf{e}_k$ and feed an arbitrary vector **b** in the empty slot $\bullet$. Then

$$g(\mathbf{v}, \bullet)\mathbf{b} = \sum_{ij} g_{ij} \mathbf{e}^i_\star \otimes \mathbf{e}^j_\star (\sum_k v^k \mathbf{e}_k, \mathbf{b}) = \sum_{ijk} g_{ij} v^k \mathbf{e}^i_\star(\mathbf{e}_k) \otimes \mathbf{e}^j_\star(\mathbf{b}) = \sum_{ij} g_{ij} v^i \mathbf{e}^j_\star(\mathbf{b}) \tag{A.18}$$

That is, $g(\mathbf{v}, \bullet)$ is a covector whose j^{th} component in the $\mathbf{e}^1_\star, ..., \mathbf{e}^n_\star$ basis is $\sum_i g_{ij} v^i$.

Having this covector $g(\mathbf{v}, \bullet) = \sum_{ij} g_{ij} v^i \mathbf{e}^j_\star$, can we extract our original **v** from it? Yes, we can:

> **Exercise:** Consider a contravariant 2-tensor $\bar{g} = \sum_{ij} \bar{g}^{ij} \mathbf{e}_i \otimes \mathbf{e}_j$. Explain how $\bar{g}$ turns covectors into vectors. By filling in the first slot of $\bar{g}$ with the covector $\sum_{ij} g_{ij} v^i \mathbf{e}^j_\star(\bullet)$, show that our original vector $\mathbf{v} = \sum_i v^i \mathbf{e}_i$ is returned if we choose the components $\bar{g}^{ij} = (g^{-1})_{ij}$, where $(g^{-1})_{ij}$ are the elements of the inverse matrix of g_{ij}.

To summarise, the metric (i.e., the scalar product) induces a natural pairing between vectors and covectors, with $\mathbf{e}_j$ mapped to $\sum_i g_{ij} \mathbf{e}^\star_i$ and $\mathbf{e}^\star_j$ to $\sum_i g^{ij} \mathbf{e}_i$. Therefore

$$\begin{aligned} \mathbf{v} = \sum_i v^i \mathbf{e}_i \in \mathcal{V} \quad &\leftrightarrow \quad \mathbf{v}^\star = \sum_{ij} g_{ij} v^i \mathbf{e}^j_\star \in \mathcal{V}^\star, \\ \mathbf{v}^\star = \sum_i v^\star_i \mathbf{e}^i_\star \in \mathcal{V}^\star \quad &\leftrightarrow \quad \mathbf{v} = \sum_{ij} g^{ij} v^\star_i \mathbf{e}_j \in \mathcal{V}. \end{aligned} \tag{A.19}$$

Here g^{ij} is the inverse of the metric tensor, $g^{ij} \equiv (g^{-1})_{ij}$.

More generally, any tensor T that wants to be fed a vector in one of its input slots can converted into one that accepts a covector: just attach a $\bar{g} = g^{-1}$ over the input slot to intercept the input covector and turn it into a vector before feeding to the original T. Conversely, if T wants a covector, we can nevertheless feed it vectors by using g to turn vectors into covectors before feeding to T. So, using the metric we can turn a type-$\binom{m}{n}$ tensor into another type $\binom{r}{s}$ one, as long as $m + n = r + s$. This procedure is usually called "raising" or "lowering" and is easier to follow if we use index notation.

A.6 Index notation

We have seen that a general rank $\binom{s}{r}$ tensor can be expressed as

$$T = \sum_{i_1} \cdots \sum_{i_s} \sum_{j_1} \cdots \sum_{j_r} T^{i_1 \cdots i_s}_{j_1 \cdots j_r} \mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_s} \otimes \mathbf{e}^{j_1}_\star \otimes \cdots \otimes \mathbf{e}^{j_r}_\star, \tag{A.20}$$

where the n^{r+s} *components* are the numbers

$$T^{i_1\cdots i_s}_{j_1\cdots j_r} = T(\mathbf{e}^{i_1}_\star, ..., \mathbf{e}^{i_s}_\star, \mathbf{e}_{j_1}, ..., \mathbf{e}_{j_r}). \tag{A.21}$$

obtained by applying T to the basis vectors. Just as we often loosely refer to a vector $\mathbf{v} = \sum_i v^i \mathbf{e}_i$ by its components $(v^1, ..., v^n)$, you will usually see the components $T^{i_1\cdots i_s}_{j_1\cdots j_r}$ of the tensor (A.20) being used as shorthand for the tensor itself, the basis vectors being suppressed.

The distinction between tensors of different type is made solely by the number and position of the indices. So A^i is a (contravariant) vector. The corresponding (covariant) (co)vector is written A_i, where $A_i = g_{ij}A^j$: any index that appears once downstairs and once upstairs is to be summed over (the Einstein summation convention); the same symbol is used for a vector and the corresponding covector.

The scalar product of the vectors A and B is $g_{ij}A^iB^j = A^iB_i = g^{ij}A_iB_j = A_iB^i$. Their outer product can be expressed as A^iB^j (a contravariant 2-tensor) or A_iB_j (a covariant 2-tensor) and so on.

Given, e.g., a covariant 2-tensor T_{ij}, we can turn the first slot from a vector-eating machine to a covector-eating machine by raising the first index: $T^i{}_j = g^{ik}T_{kj}$: notice that we leave a space underneath the raised index i to indicate that it labels the first input slot of the original $T : \mathcal{V} \times \mathcal{V} \to \mathcal{F}$ and that j labels the second. If the original T_{ij} were symmetric then we could safely omit this space.

A.7 Transformation properties

Tensors do not care about which basis we choose to talk about them. It is up to us to ensure that anything we say about the components of a tensor in a particular basis transforms in a way that respects this indifference.

Suppose that we introduce a new basis (remember, summation convention)

$$\mathbf{e}'_i = (\Lambda^{-1})^i{}_j \mathbf{e}_j, \tag{A.22}$$

so that

$$\mathbf{e}'^i_\star = \Lambda^i{}_j \mathbf{e}^j_\star. \tag{A.23}$$

Inverting both of these equations and plugging directly into (A.20) it follows that the components of the tensor in the new basis are given by

$$T'^{i_1\cdots i_s}_{j_1\cdots j_r} = \Lambda^{i_1}{}_{k_1} \cdots \Lambda^{i_s}{}_{k_s} (\Lambda^{-1})^{l_1}{}_{j_1} \cdots (\Lambda)^{l_q}{}_{j_r} T^{k_1\cdots k_s}_{l_1\cdots l_r}. \tag{A.24}$$

Any multilinear object that transforms in this way is automatically a type-$\binom{s}{r}$ tensor.

In particular, the components of a contravariant vector transform as

$$A'^i = \Lambda^i{}_j A^j, \tag{A.25}$$

and those of a covariant vector as

$$B'_i = (\Lambda^{-1})^j{}_i B_j. \tag{A.26}$$

> **Exercise:** Show that the condition for the transformation $\Lambda^i{}_j$ to preserve the Minkowski scalar product (A.17) is $\eta_{ij}\Lambda^i{}_k\Lambda^j{}_l = \eta_{kl}$. Any $\Lambda^i{}_j$ that satisfies this condition is a **Lorentz transformation**. A familiar example of such a transformation is
>
> $$\Lambda = \begin{pmatrix} \gamma & -\beta\gamma & & \\ -\beta\gamma & \gamma & & \\ & & 1 & \\ & & & 1 \end{pmatrix}, \tag{A.27}$$
>
> with $\beta = v/c$ and $\gamma = 1/\sqrt{1-\beta^2}$.

A.8 Vector and tensor fields

In classical physics (e.g., electromagnetism, relativity) the vector space $\mathcal{V}$ is usually a tangent space.

Example: Tangent vectors (sphere) Take a point $\mathbf{r}$ on the surface of a sphere. At the point $\mathbf{r}$ attach a plane whose normal coincides with that of the surface of the sphere there. Any vector tangent to the surface of the sphere at $\mathbf{r}$ is represented by an arrow that starts at $\mathbf{r}$ and is confined to this plane. The collection of all such arrows is known as the **tangent space** at $\mathbf{r}$. It is a two-dimensional vector space. Each point $\mathbf{r}$ on the sphere has its own two-dimensional tangent space.

Tangent vectors (general) Consider smooth functions $f : \mathcal{M} \to \mathcal{F}$ defined on some n-dimensional space $\mathcal{M}$. Choose a point P in $\mathcal{M}$ and introduce coordinates $(x_1, ..., x_n)$ whose origin is at P. The set of directional derivatives evaluated at P,

$$\left(a^1\partial_1 + \cdots + a^n\partial_n\right), \tag{A.28}$$

forms a n-dimensional vector space, known as the **tangent space** at P. The partial derivative operators $\partial_i \equiv \frac{\partial}{\partial x^i}\big|_P$ are a basis for this space and, in terms of this basis, the a^i are the coordinates. The basis for the dual space is $\mathrm{d}x^i$: we have that $\mathrm{d}x^i(\partial_j f) = \delta_{ij} f$ for any smooth function f.

Transformation laws Given a new set of coordinates $(x'^1, ..., x'^n)$ for the tangent space at P, equation (A.23) becomes

$$\mathrm{d}x'^i = \Lambda^i{}_j\,\mathrm{d}x^j = \frac{\partial x'^i}{\partial x^j}\,\mathrm{d}x^j. \tag{A.29}$$

Therefore the transformation matrix between the x^i and the x'^i is $\Lambda^i{}_j = \frac{\partial x'^i}{\partial x^j}$ and the components (A.24) of a general tensor transform as

$$T'^{i_1\cdots i_s}_{j_1\cdots j_r} = \left(\frac{\partial x'^{i_1}}{\partial x^{k_1}}\right)\cdots\left(\frac{\partial x'^{i_s}}{\partial x^{k_s}}\right)\left(\frac{\partial x^{l_1}}{\partial x'^{j_1}}\right)\cdots\left(\frac{\partial x^{l_s}}{\partial x'^{j_s}}\right)T^{k_1\cdots k_s}_{l_1\cdots l_r}. \tag{A.30}$$

In particular,

$$\begin{aligned} A'^i &= \left(\frac{\partial x'^i}{\partial x^j}\right)A^j, \\ A'_i &= \left(\frac{\partial x^j}{\partial x'^i}\right)A_j. \end{aligned} \tag{A.31}$$

The gradient of a scalar field ϕ has components $\partial_i\phi$. It is covariant (check how it transforms under changes of coordinates). It lives in the tangent space to the point $\mathbf{x}$ at which we evaluate the derivatives $\partial_i\phi$.

Defining gradients of vector fields is more subtle, as it requires some way of connecting the tangent space at points $\mathbf{x}$ with the tangent space (or cotangent space) at neighbouring points $\mathbf{x} + \Delta\mathbf{x}$. This can be done using the *affine connection*, but that is a topic for another course.

Appendix: probability

http://www-thphys.physics.ox.ac.uk/people/JohnMagorrian/mm
john.magorrian@physics.ox.ac.uk

Appendix B: Basic ideas of probability theory

You already have an intuitive notion of what probability theory is all about. It can be reduced to some simple ideas about sets and mappings between sets, which were identified by Kolmogorov (1933).

B.1 Sample space and events

Suppose that we are doing an experiment, perhaps multiple times. Every time we do the experiment (that is, for each *trial*) we obtain some result (or *outcome* or *sample*). The space of all possible outcomes is known as the **sample space**, denoted Ω. Some examples:

- If our experiment involves making a single toss of a coin, the sample space Ω is the set of the possible outcomes, "heads" or "tails". Then $\Omega = \{\text{heads}, \text{tails}\}$.
- For a single throw of a die the sample space $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- If we throw a pair of non-identical dice (e.g., one is red, the other is green) the sample space contains the 36 possible outcomes $\{(1,1), (1,2), ..., (1,6), (2,1), ..., (2,6), ..., (6,6)\}$, where the first number i in each pair (i, j) is the result shown on the red die, the second j on the green one.
- If the dice are indistinguishable then we can't tell the difference between, say, $(1,2)$ and $(2,1)$: in this case there are only 21 possible outcomes, which we could represent using ordered pairs (i, j) with $i = 1, .., 6$ and $j = 1, ..., i$. Alternatively, we could recognise that the dice are physically distinct (even though we can't get close enough to tell them apart) and take Ω to be the same 36 outcomes as in the previous case.
- A coach measures how long it takes a runner to complete a lap of a 400m track. The sample space consists of all non-negative real numbers, corresponding to the length of time (in seconds) the runner takes. (We assume that our coach can measure times to abitrary precision.)

We are often not interested in the precise details of the outcome of each trial. For example, we might care only that our runner completes a lap in less than 60 seconds, or that the ball lands in a red slot in a spin of roulette. To deal such cases we define an **event** $\mathcal{F}$ as a subset of the sample space Ω. If the outcome of the trial is a member of this subset, we say that the corresponding event **occurs**. For example, running a sub-minute lap corresponds to the event $[0, 60)$. Throwing an even number on the die means that the event $\{2, 4, 6\}$ occurs. The outcome of each trial can correspond to many events, however: if we throw a six on our die then there are 32 distinct events that occur (namely, $\{6\}$, $\{2, 4, 6\}$ and all 30 other subsets of $\Omega = \{1, ..., 6\}$ that include the outcome 6). A runner completing a 400m lap in 42.something seconds would be quite an event, corresponding to the subset $[42, 43)$ of the real line. The more pedestrian sub-minute event $[0, 60)$ would also occur in this case.

So, events are represented by subsets of Ω, not by the elements of Ω themselves. The sample space Ω itself is an event, known as the *certain event*, while the empty set $\varnothing$ is the *impossible event*. Following the usual set notation, given two events A and B we can define the following new events:

- $A \cup B$ contains all elements of Ω that are in A or B (union, A *or* B);
- $A \cap B$ contains all elements that are in both A and B (intersection, A *and* B);
- $A \backslash B$ is the *difference*, containing all elements of A that are not in B (A *but not* B);
- $A^{\mathrm{c}} \equiv \Omega \backslash A$ is the *complement* of A (*not* A).

Let us write $\mathcal{F}$ for the collection of all possible events. If Ω is countable[⋆] then this $\mathcal{F}$ is just the set of all subsets of Ω. If Ω is uncountable (e.g., if it is some interval the real numbers) then this notion of "set of all sets" leads to a morass. Nevertheless, we can still just assume that some good choice of $\mathcal{F}$ exists in such cases, without needing to know the details.

In case you're wondering, here is the secret recipe. We require only that $\mathcal{F}$ be closed under countable unions, specifically:

- $\varnothing \in \mathcal{F}$;
- if $A_1, A_2, ... \in \mathcal{F}$ then $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$;
- if $A \in \mathcal{F}$ then $A^{\mathrm{c}} \in \mathcal{F}$.

Any collection $\mathcal{F}$ of subsets of Ω that satisfies these conditions is called a σ-field. If Ω is countable then the set of all subsets of Ω is automatically a σ-field.

Examples of σ-fields: $\mathcal{F} = \{\varnothing, \Omega\}$. If A is any subset of Ω then $\mathcal{F} = \{\varnothing, A, A^{\mathrm{c}}, \Omega\}$.

For our purposes this is probably best kept a secret: you really don't need to know it.

B.2 Probability measure

That's enough about sets of subsets of Ω. Where does probability come in to this? A **probability measure** $\mathbb{P}$ on $(\Omega, \mathcal{F})$ is a mapping $\mathbb{P} : \mathcal{F} \to [0,1]$ that assigns a real number between 0 and 1 to each event in $\mathcal{F}$, subject to the following conditions:

(i) $\mathbb{P}(\varnothing) = 0$, $\mathbb{P}(\Omega) = 1$;

(ii) if $A_1, A_2, ...$ is a collection of disjoint members of $\mathcal{F}$ (i.e., with $A_i \cap A_j = \varnothing$ for all $i \neq j$), then

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i). \tag{B.1}$$

There is nothing surprising here. The probability assigned to the impossible event should be zero and that assigned to the certain event should be 1. The probabilities assigned to distinct, non-overlapping events between these two extremes should add up in the obvious way. The triplet $(\Omega, \mathcal{F}, \mathbb{P})$ is called a **probability space**.

That's the absolute basics of probability theory: it's all about mappings between **events** and the real numbers between 0 and 1 inclusive. Physics enters into the construction of the mapping $\mathbb{P}$. The rest is maths. For example, the probability that a coin lands heads up depends on the mass distribution of the coin, how it's thrown and the surrounding environment (air conditions, details of the surface on which it lands, and so on). Once we've decided to assign $\mathbb{P}(\text{heads}) = \frac{1}{2}$ (which might be from experiment, detailed calculation or just plain indifference), the maths takes over.

Exercise: By drawing Venn diagrams (how many?), or otherwise, verify the following distributive laws for events A, B and C:

$$\begin{aligned} A \cap (B \cup C) =& (A \cap B) \cup (A \cap C), \\ A \cup (B \cap C) =& (A \cup B) \cap (A \cup C). \end{aligned} \tag{B.2}$$

B.3 Conditional probability

Let us suppose that the event B occurs. Because events are just sets, then some other event A occurs if and only if the event $A \cap B$ occurs. So the probability that A occurs given that B occurs must be proportional to $\mathbb{P}(A \cap B)$. To find the constant of proportionality consider condition (i) of §B.2: the probability of the

[⋆] a "countable" set is one that can be mapped one-to-one to the natural numbers $\{0, 1, 2, 3, 4, ...\}$. Any subset of the integers is countable. The set of all rational numbers is (perhaps surprisingly?) countable, as is the set of all ordered pairs (i_1, i_2) of integers. The reals are not countable.

certain event Ω given that B occurs must equal 1. Therefore, assuming $\mathbb{P}(B) > 0$, we define the **conditional probabilty** of the event A given B as

$$\mathbb{P}(A|B) \equiv \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}. \tag{B.3}$$

Swapping A and B and using the symmetry of the $\cap$ operator we have the familiar

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\,\mathbb{P}(B) = \mathbb{P}(B|A)\,\mathbb{P}(A). \tag{B.4}$$

B.4 Independence

Two events, A and B, are **independent** if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B). \tag{B.5}$$

That is, if $\mathbb{P}(A|B) = \mathbb{P}(A)$ and $\mathbb{P}(B|A) = \mathbb{P}(B)$. Notice that independence does *not* mean that $A \cap B = \varnothing$: independence is a statement about *probabilities* of events, not about the nature of the events themselves. For example, suppose we draw a playing card from a full deck. Let A be the event of obtaing a 4, B obtaining a club. The probabilities $\mathbb{P}(A)$ and $\mathbb{P}(B)$ are independent according to (B.5), even though the events A and B overlap (both events would occur if we drew the 4 of clubs). On the other hand, if we started from a deck that did not have the full complement of either 4s or clubs then the probabilities would change and these two events would no longer be independent.

Three events A, B, C are independent if each pair is independent and also $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\,\mathbb{P}(B)\,\mathbb{P}(C)$. More generally, a collection of events $\{A_1, A_2, ..., A_n\}$ is independent if, for any subset $J \subseteq \{1, 2, .., n\}$ the probabilities satisfy

$$\mathbb{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbb{P}(A_j). \tag{B.6}$$

B.5 Bayes

Suppose we split Ω into a collection of disjoint events $B_1, B_2, ...$ for which $\cup_i B_i = \Omega$ and $B_i \cap B_j = \varnothing$ when $i \neq j$. The collection $\{B_1, B_2, ...\}$ is then said to **partition** Ω: every element of Ω belongs to precisely one of the B_i. For any event A we can construct a collection of new events $(A \cap B_1)$, $(A \cap B_2)$, ... Clearly, these new events are disjoint, with $(A \cap B_i) \cap (A \cap B_j) = \varnothing$ for $i \neq j$, and the union of all of them, $\cup_i (A \cap B_i)$, is just A. Using condition (ii) of §B.2, we have that

$$\begin{aligned} \mathbb{P}(A) = \mathbb{P}\left(\bigcup_i (A \cap B_i)\right) &= \sum_i \mathbb{P}(A \cap B_i) \\ &= \sum_i \mathbb{P}(A|B_i)\,\mathbb{P}(B_i), \end{aligned} \tag{B.7}$$

the last equality following from (B.4). Suppose that we know that A occurs and want to determine the probability of each B_j. Using (B.4) followed by (B.7), we obtain **Bayes' formula**,

$$\mathbb{P}(B_j|A) = \frac{\mathbb{P}(A|B_j)\,\mathbb{P}(B_j)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_j)\,\mathbb{P}(B_j)}{\sum_i \mathbb{P}(A|B_i)\,\mathbb{P}(B_i)}, \tag{B.8}$$

the single most important result in applications of probability theory.

An aside: stepping way back, *all* probabilities are conditional. In §B.2 whenever we assign probability measure $\mathbb{P}$ to the collection of events $\mathcal{F}$, we invariably do so under some assumptions (e.g., that our coin is unbiased, that our die is fair). Moreover, when defining the sample space we invariably restrict ourselves to certain idealised conditions, which are often unstated. For example, in the coin-tossing example we usually exclude freak possibilities, such as the coin landing on its edge, or Magnus the Magpie swooping in and stealing it before it lands. In the mathematical universe of our $(\Omega, \mathcal{F}, \mathbb{P})$ probability space this outside world of bizarre possibilities does not exist. But if we're feeling energetic enough there is nothing to stop us from extending our $(\Omega, \mathcal{F}, \mathbb{P})$ to include these possibilities.

B.6 Random variables

The formal definition of a **random variable** X is a mapping $X : \Omega \to \mathbb{R}$ that assigns a real number to each possible outcome of our experiment: i.e., for each element ω of the sample space Ω it returns some $X(\omega)$.

In practice, random variables usually correspond to quantities that we can (in principle) measure. **Discrete random variables** only return certain discrete values. **Continuous random variables** vary continuously in some interval. We'll use uppercase letters (e.g., X) to denote random variables and lowercase letters (x) to denote particular values taken by the function. The random variable X itself is neither random nor variable – it's just a mapping – but the values x are.

> Here are some technical points that you should ignore, but might find interesting. The mapping X must be so-called $\mathcal{F}$-measurable. That is, for any $x_{\max} \in \mathbb{R}$ we must have $\{\omega \in \Omega : X(\omega) < x_{\max}\} \subset \mathcal{F}$: that is, the set of all outcomes that produce results less than or equal to $x_{\max}$ must belong to the field $\mathcal{F}$ of allowed events. This condition ensures that the cumulative probability distribution (see equation (B.9) below) is well defined. [Recall that a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ involves three objects: the sample space Ω, the field $\mathcal{F}$ of events and the mapping $\mathbb{P} : \mathcal{F} \to [0, 1]$. Clearly the latter determines the distribution of possible values of the function X. But $\mathbb{P}$ only accepts events that belong to $\mathcal{F}$.]
>
> X is a **discrete random variable** if the set $\{X(\Omega) : \omega \subset \Omega\}$ of allowed values of X is countable.
>
> X is a **continuous random variable** if its cumulative DF can be written as (B.24) below for some nonnegative function $f_X(x)$.

Examples of discrete random variables:

(i) The score obtained when we throw a single die. The set of possible values is $\{1, 2, ..., 6\}$.
(ii) The total score obtained when throwing a pair of dice. Possible values are drawn from $\{2, 3, ..., 12\}$
(iii) The balance of your bank account at the end of the month.
(iv) The number of alpha particles emitted by a block of radium during the next minute.
(v) The score obtained when we throw a single dart at a dartboard. Possible values are 0 (if we miss the board altogether), 1, 2, ..., 20, 22, ..., 57, 60.

Examples of continuous random variables:

(i) The distance (in meters) that a dart lands from the centre of the board.
(ii) The x and y coordinates of where the dart lands are each random variables.
(iii) The time interval between successive alpha decays from the block of radium.
(iv) The weight of a randomly-chosen bank manager.

Comments:

(i) If the sample space Ω is countable then all random variables are necessarily discrete. On the other hand, if Ω isn't countable then random variables can be discrete or continuous (or neither).
(ii) Any real-valued function of one or more discrete random variables is itself a discrete random variable. Similarly, any smooth, real-valued function of one or more continuous random variables is another continuous random variable.

B.7 Cumulative distribution function

Any random variable X, whether discrete or continuous (or neither), is completely described by its **cumulative distribution function** (or CDF), $F_X(x)$, which measures the total probability mass associated with all values of X less than some threshold x. That is,

$$F_X(x) \equiv \mathbb{P}(X \leq x). \tag{B.9}$$

Some points to note:

(i) $F_X(x) \to 0$ as $x \to -\infty$ and $F_X(x) \to 1$ as $x \to \infty$.

(ii) $F_X(x)$ never decreases; it has no maxima.

(iii) If X is a discrete random variable then $F_X(x)$ is a staircase function: constant everywhere except for discrete upward jumps at each possible value of X (see Figure B-1 below).

(iv) If X is a continuous random variable then $F_X(x)$ is a smoothly varying function of x.

(v) The probability of measuring X in the interval $(a, b]$ is simply

$$\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a). \tag{B.10}$$

[Confirmation: write the event $\{X : a < X \leq b\}$ as the difference $B\backslash A$ of the events $B = \{X : X \leq b\}$ and $A = \{X : X \leq a\}$. We have that $B = (B\backslash A) \cup A$ and that the events $(B\backslash A)$ and A are disjoint: $(B\backslash A) \cup A = \varnothing$. From the definition of $\mathbb{P}$ in §B.2 it follows then that $\mathbb{P}(B) = \mathbb{P}(B\backslash A) + \mathbb{P}(A)$.]

B.8 Discrete random variables

Let X be a discrete random variable and let $I = \{x_1, x_2, ...\}$ be the set of discrete possible results that X can return. The **probability mass function** of X is defined as

$$\mathbb{P}(X = x) \equiv \mathbb{P}(\mathcal{E}(x)), \tag{B.11}$$

where $\mathcal{E}(x) = \{X(\omega) = x : \omega \in \Omega\}$ is the event composed of the union of all outcomes for which X takes the value x. Notice that this overloads the meaning of the $\mathbb{P}$ symbol. Fundamentally $\mathbb{P}$ stands for the mapping between events and $[0, 1]$, but any random variable depends on that same mapping (because the set of all possible ways in which X can take the value x is itself a well-defined event) and so the same symbol $\mathbb{P}$ is used for both. Sometimes we'll write $p_X(x)$ or just $p(x)$ as shorthand for this $\mathbb{P}(X = x)$.

This probability mass function is related to the CDF of X via

$$F_X(x) \equiv \mathbb{P}(X < x) = \sum_{\substack{x' < x \\ x \in I}} \mathbb{P}(X = x'). \tag{B.12}$$

If we know the CDF then we know the probability mass function, and vice versa. The CDF is a discontinuous, staircase-like function, which jumps upwards at each value of x_i (see, e.g., Figure B-1 below).

You've probably already encountered the following three ways of locating the "centre" of the probability mass function $\mathbb{P}(X = x)$. The **mode** is the value(s) of x at which $\mathbb{P}(X = x)$ reaches its maximum. The **median** is the value of x for which $\mathbb{P}(X < x) = \frac{1}{2}$: half of the "mass" lies to the left of the median, half to the right. Most important, however, is the **mean** or **expectation** value, defined as

$$\mathbb{E}[X] \equiv \sum_{x \in I} x\, \mathbb{P}(X = x). \tag{B.13}$$

It gives the location of the centre of (probability) mass of the distrbution. Some comments:

- This $\mathbb{E}[X]$ is not a random variable: it's just a number.
- Alternative ways of denoting $\mathbb{E}[X]$ include $\overline{X}$ and $\langle X \rangle$.

- Note that neither the mean nor the median need belong to the set of possible values $I = \{x_i\}$ that X is allowed to take.

The k^{th} **moment** of X (k a positive integer) is obtained by taking $g(x) = x^k$ in the example above:

$$\mathbb{E}[X^k] = \overline{X^k} = \sum_{x \in I} x^k \, \mathbb{P}(X = x). \tag{B.14}$$

The zeroth moment is unity, while the first moment is equal to the mean. The k^{th} **central moment** of X is the k^{th} moment of X after subtracting off this mean:

$$\mathbb{E}[(X - \bar{X})^k] = \overline{(X - \bar{X})^k} = \sum_{x \in I} (x - \bar{X})^k \, \mathbb{P}(X = x). \tag{B.15}$$

The first-order central moment is zero. The second central moment is the **variance**,

$$\mathrm{var}[X] \equiv \mathbb{E}\left[(X - \bar{X})^2\right] = \mathbb{E}[X^2] - \bar{X}^2, \tag{B.16}$$

which is the single most useful and convenient way of quantifying how concentrated the distribution is its mean value. Recalling that the expectation value $\mathbb{E}[X]$ is the location of the centre of (probability) mass of X, it follows that the variance $\mathrm{var}[X]$ is its moment of inertia. The characteristic width of the distribution is then

$$\mathrm{std}[X] = \sqrt{\mathrm{var}[X]}, \tag{B.17}$$

which is known as the **standard deviation** of X.

[There are alternative ways of measuring the width of a distribution, such as the locations $(x_{1/4}, x_{3/4})$ of the first and third quartiles defined in terms of the cumulative distribution function through $\mathbb{P}(X \le x_{1/4}) = \frac{1}{4}$ and $\mathbb{P}(X \le x_{3/4}) = \frac{3}{4}$. They are sometimes more robust in practice, but harder to reason about than the variance.]

Exercise: For real numbers a and b, show that

$$\begin{aligned} \mathbb{E}[aX + b] &= a\mathbb{E}[X] + b, \\ \mathrm{var}[aX + b] &= a^2 \, \mathrm{var}[X]. \end{aligned} \tag{B.18}$$

B.9 Some important discrete distributions

Bernoulli distribution The mother of all discrete distributions is the simple coin flip. This is the prototypical example of a *Bernoulli trial*, the name given to the situation in which there are just two possible outcomes: heads/tails, success/failure, left/right, fight/flee, truth/dare. Let us assign $X = +1$ if the coin lands heads up, $X = 0$ for tails. If p is the probability of heads, then $\mathbb{E}[X] = \bar{X} = p$ and $\mathrm{var}[X] = \mathbb{E}[(X - \bar{x})^2] = p(1 - p)$.

Exercise: There is nothing magical about the values 0 and 1 for X. Suppose that X takes the value $+1$ with probability p and -1 with probability $1 - p$. What are $\mathbb{E}[X]$ and $\mathrm{var}[X]$ then?

Binomial distribution Suppose we carry out n independent Bernoulli trails. Then our sample space Ω consists of all 2^n sequences $(X_1, X_2, ..., X_n)$ in which each X_i takes the value $+1$ with probability p, 0 with probability $(1 - p)$. By (B.5) the probability of any sequence $(X_1, X_2, ..., X_n) \in \Omega$ is simply

$$\mathbb{P}(X_1, X_2, ..., X_n) = \mathbb{P}(X_1) \, \mathbb{P}(X_2) \cdots \mathbb{P}(X_n) \tag{B.19}$$

Let K be a new random variable that counts the number of times $X_i = +1$ occurs. From (B.19) it is clear that, given a particular sequence $(X_1, ..., X_n)$ for which $K = k$, the probability associated with that

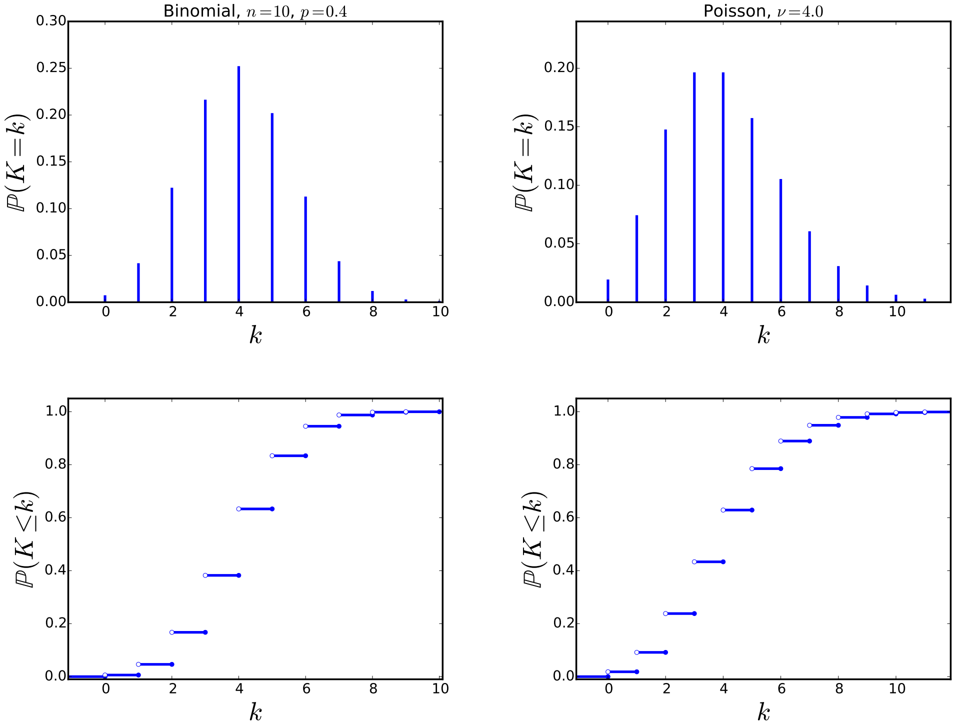


Figure B-1. Probability mass functions (top) and cumulative distribution functions (bottom) of a binomial distribution (B.20) having $n = 10$, $p = 0.4$ (left) and a Poisson distribution (B.23) having $\nu = 4$ (right). The Possion distribution is defined for $k \to \infty$, but the plots are restricted to $k < 12$.

particular sequence is $p^k(1-p)^{n-k}$. But there are *many* sequences for which $K = k$: in fact, there are $\binom{n}{k} = n!/k!(n-k)!$ of them and each one occurs with the same probability, $p^k(1-p)^{n-k}$. Therefore

$$\mathbb{P}(K = k) = \binom{n}{k} p^k (1-p)^{n-k}. \tag{B.20}$$

Multinomial distribution Keeping the number of throws n fixed, consider what happens when we increase the number of sides of our "coin" to $m > 2$. For example, if we replaced our coin by a die we would have $m = 6$. Then the result of each throw, $X_1, ..., X_n$, is one of m values, which occur with probabilities $p_1, ..., p_m$, satisfying $p_1 + \cdots + p_m = 1$. The probability of any particular sequence of throws is still given by (B.19).

Let $K_1, ..., K_m$ be random variables that count the number of times that each of our m possible per-throw outcomes occurs. So, K_6 would be the number times we rolled a six in our die example. We can work out the probability of any given sequence $(x_1, ..., x_n)$ by counting the number of times each outcome occurs among the $\{x_i\}$: if outcome 1 appears k_1 times, outcome 2 occurs k_2 times, etc, then the probability of that sequence is $p_1^{k+1} p_2^{k_2} \cdots p_m^{k_m}$. But of the m^n possible sequences there are a total of $n!/k_1! \cdots k_m!$ that share the property that oucome 1 occurs k_1 times, outcome 2 occurs k_2 times etc, each occuring with the same probability. Therefore

$$\mathbb{P}(K_1 = k_1, ..., K_m = k_m) = \frac{n!}{k_1! \cdots k_m!} p_1^{k_1} \cdots p_m^{k_m}, \tag{B.21}$$

which is the *multinomial distribution*. The binomial distribution is the special case $m = 2$.

Poisson distribution There is one more very important discrete distribution. Consider the following model of radioactive decay. Split the time interval $[0, T]$ into n subintervals each of length Δt, so that

$T = n\Delta t$. Let p be the probability of an event (e.g., a radioactive decay event) occuring during one of those subintervals. We expect this probability to be propotional to Δt. So let $p = \lambda\Delta t$, where λ is known as the *rate parameter*. If we make Δt small enough there will be at most one event per subinterval and the probability of having k events among these n is just (binomial)

$$\mathbb{P}(K=k) = \binom{n}{k} p^k (1-p)^{n-k}. \tag{B.22}$$

Now let $n \to \infty$ holdiing T fixed, so that $\Delta t = \frac{T}{n} \to 0$. Then

$$\begin{aligned}\mathbb{P}(K=k) &= \lim_{n\to\infty} \frac{n!}{k!(n-k)!} \left[\lambda\Delta t\right]^k \left[1-\lambda\Delta t\right]^{n-k} \\ &= \lim_{n\to\infty} \frac{n!}{k!(n-k)!} \left[\frac{\lambda T}{n}\right]^k \left[1-\frac{\lambda T}{n}\right]^{n-k} \\ &= \tfrac{1}{k!} \left[\lambda T\right]^k \mathrm{e}^{-\lambda T} \\ &= \tfrac{1}{k!} \nu^k \mathrm{e}^{-\nu},\end{aligned} \tag{B.23}$$

where in the last line we have introduced $\nu \equiv \lambda T$.

> **Exercise:** Show that the expectation value and the variance of the Poisson distribution are both equal to ν. Going back to our radioactive decay model, what does this imply about the spread in the number of counts received when the time interval T becomes large?

B.10 Continuous random variables

By definition, a continuous random variable is one whose CDF (B.9) can be expressed as the integral

$$F_X(x) \equiv \mathbb{P}(X \le x) = \int_{-\infty}^{x} f_X(x')\,\mathrm{d}x' \tag{B.24}$$

for some non-negative function $f_X(x)$. This $f_X(x)$ is known as the **probability density function** or PDF of X. Clearly, it can be obtained from the CDF just by differentiating:

$$f_X(x) = \frac{\mathrm{d}}{\mathrm{d}x} F_X(x). \tag{B.25}$$

Using (B.10), the probability (or "probability mass") of finding X in the range $[a,b]$ is then

$$\begin{aligned}\mathbb{P}(a < X \le b) = F_X(b) - F_X(a) &= \int_0^b f_X(x)\,\mathrm{d}x - \int_0^a f_X(x)\,\mathrm{d}x \\ &= \int_a^b f_X(x)\,\mathrm{d}x,\end{aligned} \tag{B.26}$$

and the probability of finding X in some small interval $[x, x+\mathrm{d}x]$ of width $\mathrm{d}x$ around x is $f_X(x)\,\mathrm{d}x$.

Note: the PDF $f_X(x)$ is not a probability: it's a probability *density*. In particular, the probability that a continuous random variable X takes on any specific value x is precisely zero, because choosing a single point implies taking $\mathrm{d}x \to 0$.

Change of variables Any smooth function g of the random variable X is itself another continuous random variable. This is a powerful idea, but it raises the following question: if X has pdf $f_X(x)$, how do we find the pdf $f_Y(y)$ of the new random variable $Y = g(X)$?

To answer this, suppose that this g is a strictly increasing function, so that it has a single-valued inverse function g^{-1}: if $y = g(x)$ then $x = g^{-1}(y)$. Now look at the CDF of y:

$$F_Y(y) \equiv \mathbb{P}(Y \leq y) = \mathbb{P}(g(X) \leq y) = \mathbb{P}(X \leq g^{-1}(y)), \tag{B.27}$$

the last equality following because g is a strictly increasing function. Following (B.25) and differentiating with respect to y, we find that the PDF of y is given by

$$f_Y(y) = f_X(g^{-1}(y))\frac{\mathrm{d}}{\mathrm{d}y}g^{-1}(y), \tag{B.28}$$

which is a bit of a mess: this is a case where you should remember the method, not the result.

Exercise: As a slight variation on this method, consider the event associated with X lying in the interval $[x, x + \mathrm{d}x]$, which has probability $f_X(x)\,\mathrm{d}x$. The same event corresponds to y lying in the interval $[y, y + \mathrm{d}y]$, where $y = g(x)$ and $\mathrm{d}y = g'(x)\mathrm{d}x$. Explain why we must have

$$f_X(x)\,\mathrm{d}x = f_Y(y)\,\mathrm{d}y. \tag{B.29}$$

Hence obtain (B.28).

Expectation, variance, moments The expectation of a continuous random variable X is defined to be

$$\mathbb{E}[X] = \bar{X} \equiv \int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x, \tag{B.30}$$

which is essentially the same as the discrete case (B.13) with $\mathbb{P}(X = x)$ replaced by $f_X(x)\,\mathrm{d}x$ and the sum turned into an integral. Although it is not as easy to show, the expectation of any other random variable $Y = g(X)$ is simply

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\,\mathrm{d}x. \tag{B.31}$$

In particular, the variance of X is (once again) the expectation value of $(X - \bar{X})^2$, which is simply

$$\mathrm{var}[X] \equiv \mathbb{E}[(X - \bar{X})^2] = \int_{-\infty}^{\infty} (x - \bar{X})^2 f_X(x)\mathrm{d}x. \tag{B.32}$$

The k^{th} moment of X is defined to be

$$\mathbb{E}[X^k] = \int_{-\infty}^{\infty} x^k f_X(x)\,\mathrm{d}x. \tag{B.33}$$

So, as for discrete variables, the variance is the second moment of the deviation, $(X - \bar{X})$, of X from its centre of mass $\bar{X} = \mathbb{E}[X]$.

Exercise: Let X and Y be continous random variables and let a and b be real numbers. Show that

$$\begin{aligned} \mathbb{E}[aX + b] &= a\mathbb{E}[X] + b, \\ \mathrm{var}[aX + b] &= a^2\,\mathrm{var}[X]. \end{aligned} \tag{B.34}$$

B.11 Some important continuous distributions

Here are the PDFs of some important continuous distributions.

Uniform distribution between $x = a$ and $x = b$:

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{if } a < x < b, \\ 0 & \text{otherwise.} \end{cases} \tag{B.35}$$

It has mean $\frac{1}{2}(b+a)$ and variance $\frac{1}{12}(b-a)^2$.

Exponential with rate parameter λ:

$$f(x) = \begin{cases} \lambda \mathrm{e}^{-\lambda x}, & \text{if } x > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{B.36}$$

Its mean is λ^{-1}, variance λ^{-2}.

> **Exercise:** Consider the radioactive-decay model we used to motivate the Poisson distribution in §B.9. Show that PDF of the time interval $t > 0$ between successive decay events is $f(t) = \lambda \mathrm{e}^{-\lambda t}$.

Normal or **Gaussian** with mean μ and variance σ^2:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]. \tag{B.37}$$

Sometimes this is written as $N(\mu, \sigma)$.

Cauchy

$$f(x) = \frac{1}{\pi(1+x)^2}. \tag{B.38}$$

B.12 Joint distributions

We can generalise the notion of random variable to random tuples (pairs), random triples, random vectors and so on. For example,

(i) In darts one usually throws three darts per round. The scores (s_1, s_2, s_3) obtained in a round comprise a discrete random triple.

(ii) The (x, y) coordinates of a single dart's landing point describe a continuous two-dimensional random vector.

For clarity we'll restrict our attention to the case of just two continuous random variables, X and Y; the corresponding results for discrete variables and the generalisation to three or more variables should be obvious.

The joint cumulative distribution function of (X, Y) is defined as the probability of the event consisting of all $X \le x$ and $Y \le y$:

$$F_{X,Y}(x, y) = \mathbb{P}\left((X \le x) \cap (Y \le y)\right). \tag{B.39}$$

It is clear that this $F_{X,Y} \to 1$ as both $x, y \to \infty$, and that $F_{X,Y} \to 0$ as either x or $y \to -\infty$. It tells us everything about how X and Y are distributed.

The joint CDF completely determines the probabilities of all events involving X and Y. The condition for X and Y to be continuous random variables is that we can express $F_{X,Y}$ as

$$F_{X,Y}(x, y) = \int_{-\infty}^{x} \mathrm{d}x' \int_{-\infty}^{y} \mathrm{d}y' f_{X,Y}(x', y') \tag{B.40}$$

for some non-negative function $f_{X,Y}(x,y)$. This $f_{X,Y}(x,y)$ is known as the **joint probability density** of (X,Y) and can be obtained by differentiating the joint CDF:

$$f_{X,Y}(x,y)=\frac{\partial^2}{\partial x\partial y}F_{X,Y}(x,y). \tag{B.41}$$

From (B.40) the probability mass associated with the square $\{x<X<x+\Delta x\}\cap\{y<Y<y+\Delta y\}$ is then

$$\begin{aligned}&(F_{X,Y}(x+\Delta x,y+\Delta y)-F_{X,Y}(x,y+\Delta y))-(F_{X,Y}(x,y)-F_{X,Y}(x+\Delta x,y))\\&=\int_x^{x+\Delta x}\mathrm{d}x'\int_y^{y+\Delta y}\mathrm{d}y'\,f_{X,Y}(x',y').\end{aligned} \tag{B.42}$$

Letting Δx and Δy tend towards zero, the probabilty mass associated with a small element $\mathrm{d}x\mathrm{d}y$ at (x,y) is simply $f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y$.

Change of variables Suppose that we have another pair of random variables $U=U(X,Y)$, $V=V(X,Y)$ that are functions of (X,Y). What is the joint PDF of (U,V)? Following (B.29) this new PDF $f_{U,V}(u,v)$ can be obtained from $f_{X,Y}(x,y)$ by requiring that probability masses are independent of the "coordinates" (x,y) or (u,v) that we use to label them. Consider a small rectangle of side $\mathrm{d}x\times\mathrm{d}y$ at (x,y). In (u,v) space this will be a parallelogram of area $|\frac{\partial u}{\partial x}\frac{\partial v}{\partial y}-\frac{\partial v}{\partial x}\frac{\partial u}{\partial y}|\mathrm{d}x\mathrm{d}y$. We must have then that

$$f_{U,V}(u,v)\left|\tfrac{\partial u}{\partial x}\tfrac{\partial v}{\partial y}-\tfrac{\partial v}{\partial x}\tfrac{\partial u}{\partial y}\right|\mathrm{d}x\mathrm{d}y=f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y. \tag{B.43}$$

That is,

$$f_{U,V}(u,v)=f_{X,Y}(x,y)\left|\frac{\partial(u,v)}{\partial(x,y)}\right|^{-1}, \tag{B.44}$$

where

$$\left|\frac{\partial(u,v)}{\partial(x,y)}\right|=\det\begin{pmatrix}\frac{\partial u}{\partial x}&\frac{\partial v}{\partial x}\\\frac{\partial u}{\partial y}&\frac{\partial v}{\partial y}\end{pmatrix} \tag{B.45}$$

is the determinant of the Jacobian of the $(x,y)\to(u,v)$ transformation.

> **Exercise:** The generalisation to three random variables should be obvious. Just to check: suppose that the random variables (X,Y,Z) and (R,Θ,Φ) are related via $X=R\sin\Theta\cos\Phi$, $Y=R\sin\Theta\sin\Phi$, $Z=R\cos\Theta$. Show that $f_{R,\Theta,\Phi}(r,\theta,\phi)=f_{X,Y,Z}(x,y,z)\,r^2\sin\theta$.

Marginal PDFs Having a joint CDF $F_{X,Y}(x,y)$ we can immediately obtain the CDFs of the individual variables X and Y simply by taking appropriate limits:

$$\begin{aligned}F_X(y)&=\lim_{y\to\infty}F_{X,Y}(x,y),\\F_Y(y)&=\lim_{x\to\infty}F_{X,Y}(x,y).\end{aligned} \tag{B.46}$$

The PDFs of X and Y are then

$$\begin{aligned}f_X(x)&=\int_{-\infty}^{\infty}f_{X,Y}(x,y)\,\mathrm{d}y,\\f_Y(y)&=\int_{-\infty}^{\infty}f_{X,Y}(x,y)\,\mathrm{d}x,\end{aligned} \tag{B.47}$$

which are sometimes referred to the the **marginal PDFs** of $f_{X,Y}(x,y)$ or $F_{X,Y}(x,y)$: Integrating over one (or more) of the variables in a joint PDF to eliminate that variable from the PDF is known as "marginalising" that variable. (This jargon makes sense if you think of taking a page containing a 2d table of numbers (our joint PDF/PMF), summing up the rows and writing the result in the margin of the page.)

Independence We have already stated the condition (B.5) for events to be independent. The condition for random variables X and Y to be independent is that the joint CDF factorises into

$$F_{X,Y} = F_X(x)F_Y(y). \tag{B.48}$$

Exercise: Show that this is equivalent to

$$f_{X,Y}(x,y) = f_X(x)f_Y(y), \tag{B.49}$$

if X and Y are continuous random variables, and to

$$\mathbb{P}((X=x)\cap(Y=y)) = \mathbb{P}(X=x)\,\mathbb{P}(Y=y) \tag{B.50}$$

if they are discrete.

Exercise: Show that

$$\mathrm{var}[X+Y] = \mathrm{var}[X] + \mathrm{var}[Y], \quad \text{if } \mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] \tag{B.51}$$

The variables X and Y are **uncorrelated** if $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$. This last exercise shows that the variance of the sum of a pair of uncorrelated variables is equal to the sum of the variances. A sufficient (but not necessary) condition for X and Y to be uncorrelated is that they be independent.

Covariance and correlation The covariance of the random variables X and Y is defined as

$$\begin{aligned}\mathrm{cov}(X,Y) &= \mathbb{E}[(X-\bar{X})(Y-\bar{Y})]\\ &= \int_{-\infty}^{\infty} \mathrm{d}x \int_{-\infty}^{\infty} \mathrm{d}y\,(x-\bar{X})(y-\bar{Y})f_{X,Y}(x,y),\end{aligned} \tag{B.52}$$

the last equality applying if X and Y are continuous. So, while the variance of a single random variable measures its (scalar) moment of inertia, the covariance of two random variables is the inertia tensor of the joint PDF.

The correlation is a rescaled version of the covariance:

$$\rho(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{var}[X]\,\mathrm{var}[Y]}}. \tag{B.53}$$

B.13 Laws of large numbers

An effective, practical explanation of the “probability of an event” is the fraction of times the event occurs in a large number of trials. The expectation or average value of a random variable is often explained in a similar way by considering the average value over multiple trials. How does this fit into the grand framework of events, probability measure and so on that we’ve just seen? Recall that we made no mention of repeated trials in §B.2 when we introduced the definition of probability measure in §B.2 nor in the definition of expectation in equations (B.13) or (B.30).

Here is (part of) the answer to that question. Suppose that we’re interested in some physical quantity, which is represented by the random variable X. We’ve devised an experiment to measure this X. We run this experiment multiple times. Let X_1 be the random variable denoting the result of the first run, X_2 the result of the second, and so on. We assume that our experimental skills are so good that these X_i are independent and identically distributed. Consider the partial sum,

$$S_n = X_1 + \cdots + X_n, \tag{B.54}$$

of the first n such results. The expectation value of the sample average $\frac{1}{n}S_n$ is then

$$\begin{aligned}\mathbb{E}\left[\tfrac{1}{n}S_n\right] &= \frac{1}{n}\left(\mathbb{E}[X_1]+\cdots+\mathbb{E}[X_n]\right)\\ &= \frac{1}{n}\left(\mathbb{E}[X]+\cdots+\mathbb{E}[X]\right) = \mathbb{E}[X] = \bar{X},\end{aligned} \tag{B.55}$$

the first equality following by the standard properties of the expectation, the second from our assumption that the X_i are independent and identically distributed. The variance of the sample average is

$$\begin{aligned}\mathrm{var}\left[\tfrac{1}{n}S_n\right] &= \frac{1}{n^2}\,\mathrm{var}\left[X_1+\cdots+X_n\right]\\ &= \frac{1}{n^2}\left(\mathrm{var}[X_1]+\cdots+\mathrm{var}[X_n]\right)\\ &= \tfrac{1}{n}\,\mathrm{var}[X]\end{aligned} \tag{B.56}$$

using the properties (B.18) and (B.51) of the variance together with the assumption that the X_i are independent. So, in the limit of a large number $n\to\infty$ of measurements, the distribution of the sample average becomes more and more tightly concentrated on $\mathbb{E}[X]$, with a characteristic width (standard deviation) that shrinks as $1/\sqrt{n}$. This is an example of a "law of large numbers"

Comments:

- This particular law of large numbers requires only that the X_i be independent with the same means and variances; apart from that it does not actually require that their PMFs/PDFs be identical.
- There are variants that show more explicitly that the probability mass associated with any nonzero value of $|\frac{1}{n}S_n - \bar{X}|$ tends to 0 as $n\to\infty$. Some of these require that the X_i are identically distributed, but do not require the variances to exist.

Exercise: You may have encountered the following empirical fact. Experiments – even if repeated many times – do not always converge on the correct result. Here is one way of attempting to model that. Suppose that the random variables X_i are related to the true value x_{true} of the physical quantity by $X_i = x_{\text{true}} + \Delta + b$, where Δ is a random measurement error and b is a constant bias. How is the distribution of sample average of these X_i related to b, Δ and the underlying x_{true}?

We have just found seen that the expectation value of the sample average $\frac{1}{n}S_n$ tends to $\mathbb{E}[X]$ and that its variance is $\frac{1}{n}\,\mathrm{var}[X]$. We can do better: the **central limit theorem** tells us that, under certain reasonable conditions, the PDF of $\frac{1}{n}S_n$ tends towards a normal distribution with that mean and variance. We postpone the proof until we've covered Fourier transforms.

B.14 Further reading

See also your *Statistics and Data Analysis* course, the first few lectures of which give plenty of examples involving "pure" probability. Lots of books cover this material: my treatment follows Grimmet & Welsh's *Probability: an introduction*. RHB§30 is another good starting point, not least because it gives plenty of examples. Beware that the meaning of the term "**distribution function**" of a random variable depends on who you talk to. In maths books on probability theory it refers to the *cumulative* distribution function. Among physicists, the "distribution function" is usually taken to mean either the probability mass function (in the case of a discrete variable) or the probability density function (for a continuous variable).

You may have noticed that we've successfully avoided stating what the map $\mathbb{P}$ actually "means". The most obvious definition is perhaps that $\mathbb{P}$ stands for the frequency of occurence of an event in the limit of many trials. But what if someone says "I'm 90% certain that it's going to rain tomorrow": how do you hold them to account? Or less fluffily, what if a cosmologist claims that "with 68% confidence, the Hubble constant lies in the range 67.0 to 68.3 km/s/Mpc"? A more general way of looking at probability is to treat events as statements ("it will rain tomorrow") and to use $\mathbb{P}$ to quantify your degree of belief in each statement. See Chapters 1 and 2 and Appendix A of Jaynes' *Probability Theory: The Logic of Science*.