5 Inner product spaces

In matrix theory it is common to say that a matrix is *symmetric* if it is equal to its transpose, $\mathsf{S}^T = \mathsf{S}$. This concept does not however transfer meaningfully to the matrix of a linear operator on a vector space unless some extra structure is imposed on that space. For example, let $S : V \to V$ be an operator whose matrix $\mathsf{S} = [S^i{}_j]$ is symmetric with respect to a specific basis. Under a change of basis $e_i = A^j{}_i e'_j$ the transformed matrix is $\mathsf{S}' = \mathsf{ASA}^{-1}$, while for the transpose matrix

$$\mathsf{S}'^T = (\mathsf{ASA}^{-1})^T = (\mathsf{A}^{-1})^T \mathsf{SA}^T.$$

Hence $\mathsf{S}'^T \neq \mathsf{S}'$ in general. We should hardly be surprised by this conclusion for, as commented at the beginning of Chapter 4, the component equation $S^i{}_j = S^j_i$ violates the index conventions of Section 3.6.

Exercise: Show that S' is symmetric if and only if S commutes with $\mathsf{A}^T\mathsf{A}$,

$$\mathsf{SA}^T\mathsf{A} = \mathsf{A}^T\mathsf{AS}.$$

Thus the concept of a 'symmetric operator' is not invariant under general basis transformations, but it is invariant with respect to orthogonal basis transformations, $\mathsf{A}^T = \mathsf{A}^{-1}$.

If V is a complex vector space it is similarly meaningless to talk of an operator $H : V \to V$ as being 'hermitian' if its matrix H with respect to some basis $\{e_i\}$ is *hermitian*, $\mathsf{H} = \mathsf{H}^\dagger$.

Exercise: Show that the hermitian property is not in general basis invariant, but is preserved under *unitary* transformations, $e_i = U^j_i e'_j$ where $\mathsf{U}^{-1} = \mathsf{U}^\dagger$.

In this chapter we shall see that symmetric and hermitian matrices play a different role in vector space theory, in that they represent *inner products* instead of operators [1–3]. Matrices representing inner products are best written with both indices on the subscript level, $\mathsf{G} = \mathsf{G}^T = [g_{ij}]$ and $\mathsf{H} = \mathsf{H}^\dagger = [h_{ij}]$. The requirements of symmetry $g_{ji} = g_{ij}$ and hermiticity $h_{ji} = \overline{h_{ij}}$ are not then at odds with the index conventions.

5.1 Real inner product spaces

Let V be a *real* finite dimensional vector space with $\dim V = n$. A **real inner product**, often referred to simply as an **inner product** when there is no danger of confusion, on the vector space V is a map $V \times V \to \mathbb{R}$ that assigns a real number $u \cdot v \in \mathbb{R}$ to every pair of vectors $u,\ v \in V$ satisfying the following three conditions:

(RIP1) The map is symmetric in both arguments, $u \cdot v = v \cdot u$.
(RIP2) The distributive law holds, $u \cdot (av + bw) = au \cdot v + bu \cdot w$.
(RIP3) If $u \cdot v = 0$ for all $v \in V$ then $u = 0$.

A real vector space V together with an inner product defined on it is called a **real inner product space**. The inner product is also distributive on the first argument for, by conditions (RIP1) and (RIP2),

$$(au + bv) \cdot w = w \cdot (au + bv) = aw \cdot u + bw \cdot v = au \cdot w + bv \cdot w.$$

We often refer to this linearity in both arguments by saying that the inner product is **bilinear**.

As a consequence of property (RIP3) the inner product is said to be **non-singular** and is often referred to as **pseudo-Euclidean**. Sometimes (RIP3) is replaced by the stronger condition

(RIP3′) $u \cdot u > 0$ for all vectors $u \neq 0$.

In this case the inner product is said to be **positive definite** or **Euclidean**, and a vector space V with such an inner product defined on it is called a **Euclidean vector space**. Condition (RIP3′) implies condition (RIP3), for if there exists a non-zero vector u such that $u \cdot v = 0$ for all $v \in V$ then $u \cdot u = 0$ (on setting $v = u$), which violates (RIP3′). Positive definiteness is therefore a stronger requirement than non-singularity.

Example 5.1 The space of ordinary 3-vectors **a**, **b**, etc. is a Euclidean vector space, often denoted $\mathbb{E}^3$, with respect to the usual scalar product

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3 = |\mathbf{a}|\,|\mathbf{b}| \cos\theta$$

where $|\mathbf{a}|$ is the length or magnitude of the vector **a** and θ is the angle between **a** and **b**. Conditions (RIP1) and (RIP2) are simple to verify, while (RIP3′) follows from

$$\mathbf{a} \cdot \mathbf{a} = \mathbf{a}^2 = (a_1)^2 + (a_2)^2 + (a_3)^2 > 0 \quad \text{if} \quad \mathbf{a} \neq \mathbf{0}.$$

This generalizes to a positive definite inner product on $\mathbb{R}^n$,

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = \sum_{i=1}^{n} a_i b_i,$$

the resulting Euclidean vector space denoted by $\mathbb{E}^n$.

The **magnitude** of a vector w is defined as $w \cdot w$. Note that in a pseudo-Euclidean space the magnitude of a non-vanishing vector may be negative or zero, but in a Euclidean space it is always a positive quantity. The *length* of a vector in a Euclidean space is defined to be the square root of the magnitude.

Two vectors u and v are said to be **orthogonal** if $u \cdot v = 0$. By requirement (RIP3) there is no non-zero vector u that is orthogonal to every vector in V. A pseudo-Euclidean inner product may allow for the existence of self-orthogonal or **null vectors** $u \neq 0$ having zero magnitude $u \cdot u = 0$, but this possibility is clearly ruled out in a Euclidean vector space.

In Chapter 9 we shall see that Einstein's special theory of relativity postulates a pseudo-Euclidean structure for space-time known as *Minkowski space*, in which null vectors play a significant role.

Components of a real inner product

Given a basis $\{e_1, \dots, e_n\}$ of an inner product space V, set

$$g_{ij} = e_i \cdot e_j = g_{ji}, \tag{5.1}$$

called the **components of the inner product with respect to the basis** $\{e_i\}$. The inner product is completely specified by the components of the symmetric matrix, for if $u = u^i e_i,\ v = v^j e_j$ are any pair of vectors then, on using (RIP1) and (RIP2), we have

$$u \cdot v = g_{ij}\, u^i\, v^j. \tag{5.2}$$

If we write the components of the inner product as a symmetric matrix

$$\mathsf{G} = [g_{ij}] = [e_i \cdot e_j] = [e_j \cdot e_i] = [g_{ji}] = \mathsf{G}^T,$$

and display the components of the vectors u and v in column form as $\mathbf{u} = [u^i]$ and $\mathbf{v} = [v^j]$, then the inner product can be written in matrix notation,

$$u \cdot v = \mathbf{u}^T \mathsf{G} \mathbf{v}.$$

Theorem 5.1 *The matrix* G *is non-singular if and only if condition (RIP3) holds.*

Proof: To prove the *if* part, assume that G is singular, $\det[g_{ij}] = 0$. Then there exists a non-trivial solution u^j to the linear system of equations

$$g_{ij} u^j \equiv \sum_{j=1}^{n} g_{ij} u^j = 0.$$

The vector $u = u^j e_j$ is non-zero and orthogonal to all $v = v^i e_i$,

$$u \cdot v = g(u, v) = g_{ij} u^i v^j = 0,$$

in contradiction to (RIP3).

Conversely, assume the matrix G is non-singular and that there exists a vector u violating (RIP3); $u \neq 0$ and $u \cdot v = 0$ for all $v \in V$. Then, by Eq. (5.2), we have

$$u_j v^j = 0 \quad \text{where} \quad u_j = g_{ij} u^i = g_{ji} u^i$$

for arbitrary values of v^j. Hence $u_j = 0$ for $j = 1, \dots, n$. However, this implies a non-trivial solution to the set of linear equations $g_{ji} u^i = 0$, which is contrary to the non-singularity assumption, $\det[g_{ij}] \neq 0$. ■

Orthonormal bases

Under a change of basis

$$e_i = A^j{}_i e'_j, \qquad e'_j = A'^k{}_j e_k, \tag{5.3}$$

the components g_{ij} transform by

$$\begin{aligned} g_{ij} = e_i \cdot e_j &= (A^k{}_i e'_k) \cdot (A^l{}_j e'_l) \\ &= A^k{}_i g'_{kl} A^l{}_j, \end{aligned} \tag{5.4}$$

where $g'_{kl} = e'_k \cdot e'_l$. In matrix notation this equation reads

$$\mathsf{G} = \mathsf{A}^T \, \mathsf{G}' \, \mathsf{A}. \tag{5.5}$$

Using $\mathsf{A}' = [A'^j{}_k] = \mathsf{A}^{-1}$, the transformed matrix G' can be written

$$\mathsf{G}' = \mathsf{A}'^T \, \mathsf{G} \, \mathsf{A}'. \tag{5.6}$$

An **orthonormal basis** $\{e_1, e_2, \dots, e_n\}$, for brevity written 'o.n. basis', consists of vectors all of magnitude ± 1 and orthogonal to each other in pairs,

$$g_{ij} = e_i \cdot e_j = \eta_i \delta_{ij} \quad \text{where} \quad \eta_i = \pm 1, \tag{5.7}$$

where the summation convention is temporarily suspended. We occasionally do this when a relation is referred to a specific class of bases.

Theorem 5.2 *In any finite dimensional real inner product space* $(V, \cdot)$*, with* $\dim V = n$*, there exists an orthonormal basis* $\{e_1, e_2, \dots, e_n\}$ *satisfying Eq. (5.7).*

Proof: The method is by a procedure called **Gram–Schmidt orthonormalization**, an algorithmic process for constructing an o.n. basis starting from any arbitrary basis $\{u_1, u_2, \dots, u_n\}$. For Euclidean inner products the procedure is relatively straightforward, but the possibility of vectors having zero magnitudes in general pseudo-Euclidean spaces makes for added complications.

Begin by choosing a vector u such that $u \cdot u \neq 0$. This is always possible because if $u \cdot u = 0$ for all $u \in V$, then for any pair of vectors u, v

$$0 = (u + v) \cdot (u + v) = u \cdot u + 2u \cdot v + v \cdot v = 2u \cdot v,$$

which contradicts the non-singularity condition (RIP3). For the first step of the Gram–Schmidt procedure we normalize this vector,

$$e_1 = \frac{u}{\sqrt{|u \cdot u)|}} \quad \text{and} \quad \eta_1 = e_1 \cdot e_1 = \pm 1.$$

In the Euclidean case any non-zero vector u will do for this first step, and $e_1 \cdot e_1 = 1$.

Let V_1 be the subspace of V consisting of vectors orthogonal to e_1,

$$V_1 = \{w \in V \mid w \cdot e_1 = 0\}.$$

This is a vector subspace, for if w and w' are orthogonal to e_1 then so is any linear combination of the form $w + aw'$,

$$(w + aw') \cdot e_1 = w \cdot e_1 + aw' \cdot e_1 = 0.$$

For any $v \in V$, the vector $v' = v - ae_1 \in V_1$ where $a = \eta_1(v \cdot e_1)$, since $v' \cdot e_1 = v \cdot e_1 - (\eta_1)^2 v \cdot e_1 = 0$. Furthermore, the decomposition $v = a\, e_1 + v'$ into a component parallel to e_1 and a vector orthogonal to e_1 is unique, for if $v = a' e_1 + v''$ where $v'' \in V_1$ then

$$(a' - a)e_1 = v'' - v'.$$

Taking the inner product of both sides with e_1 gives firstly $a' = a$, and consequently $v'' = v'$.

The inner product restricted to V_1, as a map $V_1 \times V_1 \to \mathbb{R}$, is an inner product on the vector subspace V_1. Conditions (RIP1) and (RIP2) are trivially satisfied if the vectors u, v and w are restricted to vectors belonging to V_1. To show (RIP3), that this inner product is non-singular, let $v' \in V_1$ be a vector such that $v' \cdot w' = 0$ for all $w' \in V_1$. Then v' is orthogonal to every vector in $w \in V$ for, by the decomposition

$$w = \eta_1 (w \cdot e_1) e_1 + w',$$

we have $v' \cdot w = 0$. By condition (RIP3) for the inner product on V this implies $v' = 0$, as required.

Repeating the above argument, there exists a vector $u' \in V_1$ such that $u' \cdot u' \neq 0$. Set

$$e_2 = \frac{u'}{\sqrt{|u' \cdot u'|}}$$

and $\eta_2 = e_2 \cdot e_2 = \pm 1$. Clearly $e_2 \cdot e_1 = 0$ since $e_2 \in V_1$. Defining the subspace V_2 of vectors orthogonal to e_1 and e_2, the above argument can be used again to show that the restriction of the inner product to V_2 satisfies (RIP1)–(RIP3). Continue this procedure until n orthonormal vectors $\{e_1, e_2, \dots, e_n\}$ have been produced. These vectors must be linearly independent, for if there were a vanishing linear combination $a^i e_i = 0$, then performing the inner product of this equation with any e_j gives $a^j = 0$. By Theorem 3.3 these vectors form a basis of V. At this stage of the orthonormalization process $V_n = \{0\}$, as there can be no vector that is orthogonal to every $e_1, \dots, e_n$, and the procedure comes to an end. ■

The following theorem shows that for a fixed inner product space, apart from the order in which they appear, the coefficients η_i are the same in all orthonormal frames.

Theorem 5.3 (Sylvester) *The number of + and − signs among the η_i is independent of the choice of orthonormal basis.*

Proof: Let $\{e_i\}$ and $\{f_j\}$ be two orthonormal bases such that

$$e_1 \cdot e_1 = \dots = e_r \cdot e_r = +1, \qquad e_{r+1} \cdot e_{r+1} = \dots = e_n \cdot e_n = -1,$$
$$f_1 \cdot f_1 = \dots = f_s \cdot f_s = +1, \qquad f_{s+1} \cdot f_{s+1} = \dots = f_n \cdot f_n = -1.$$

If $s > r$ then the vectors $f_1, \dots, f_s$ and $e_{r+1}, \dots, e_n$ are a set of $s + n - r > n = \dim V$ vectors and there must be a non-trivial linear relation between them,

$$a^1 f_1 + \dots + a^s f_s + b^1 e_{r+1} + \dots + b^{n-r} e_n = 0.$$

The a^i cannot all vanish since the e_i form an l.i. set. Similarly, not all the b^j will vanish. Setting

$$u = a^1 f_1 + \dots + a^s f_s = -b^1 e_{r+1} - \dots - b^{n-r} e_n \neq 0$$

we have the contradiction

$$u \cdot u = \sum_{i=1}^{s} (a^i)^2 > 0 \quad \text{and} \quad u \cdot u = -\sum_{j=1}^{n-r} (b^j)^2 < 0.$$

Hence $r = s$ and the two bases must have exactly the same number of $+$ and $-$ signs. ■

If r is the number of $+$ signs and s the number of $-$ signs then their difference $r - s$ is called the **index** of the inner product. Sylvester's theorem shows that it is an invariant of the inner product space, independent of the choice of o.n. basis. For a Euclidean inner product, $r - s = n$, although the word 'Euclidean' is also applied to the negative definite case, $r - s = -n$. If $r - s = \pm(n - 2)$, the inner product is called **Minkowskian**.

Example 5.2 In a Euclidean space the Gram–Schmidt procedure is carried out as follows:

$$\begin{aligned}
&f_1 = u_1 && e_1 = \frac{f_1}{\sqrt{f_1 \cdot f_1}} && \eta_1 = e_1 \cdot e_1 = 1, \\
&f_2 = u_2 - (e_1 \cdot u_2)e_1 && e_2 = \frac{f_2}{\sqrt{f_2 \cdot f_2}} && \eta_2 = e_2 \cdot e_2 = 1, \\
&f_3 = u_3 - (e_1 \cdot u_3)e_1 - (e_2 \cdot u_3)e_2 && e_3 = \frac{f_3}{\sqrt{f_3 \cdot f_3}} && \eta_3 = e_3 \cdot e_3 = 1, \text{etc.}
\end{aligned}$$

Since each vector has positive magnitude, all denominators $\sqrt{f_i \cdot f_i} > 0$, and each step is well-defined. Each vector e_i is a unit vector and is orthogonal to each previous e_j $(j < i)$.

Example 5.3 Consider an inner product on a three-dimensional space having components in a basis u_1, u_2, u_3

$$\mathsf{G} = [g_{ij}] = [u_i \cdot u_j] = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

The procedure given in Example 5.2 obviously fails as each basis vector is a null vector, $u_1 \cdot u_1 = u_2 \cdot u_2 = u_3 \cdot u_3 = 0$, and cannot be normalized to a unit vector.

Firstly, we find a vector u such that $u \cdot u \neq 0$. Any vector of the form $u = u_1 + au_2$ with $a \neq 0$ will do, since

$$u \cdot u = u_1 \cdot u_1 + 2au_1 \cdot u_2 + u_2 \cdot u_2 = 2a.$$

Setting $a = 1$ gives $u = u_1 + u_2$ and $u \cdot u = 2$. The first step in the orthonormalization process is then

$$e_1 = \frac{1}{\sqrt{2}}(u_1 + u_2), \qquad \eta_1 = e_1 \cdot e_1 = 1.$$

There is of course a significant element of arbitrariness in this as the choice of u is by no means unique; for example, choosing $a = \frac{1}{2}$ leads to $e_1 = u_1 + \frac{1}{2}u_2$.

The subspace V_1 of vectors orthogonal to e_1 consists of vectors of the form $v = au_1 + bu_2 + cu_3$ such that

$$v \cdot e_1 \propto v \cdot u = (au_1 + bu_2 + cu_3) \cdot (u_1 + u_2) = a + b + 2c = 0.$$

Setting, for example, $c = 0$ and $a = -b = 1$ results in $v = u_1 - u_2$. The magnitude of v is $v \cdot v = -2$ and normalizing gives

$$e_2 = \frac{1}{\sqrt{2}}(u_1 - u_2), \qquad \eta_2 = e_2 \cdot e_2 = -1, \qquad e_2 \cdot e_1 = 0.$$

Finally, we need a vector $w = au_1 + bu_2 + cu_3$ that is orthogonal to both e_1 and e_2. These two requirements imply that $a = b = -c$, and setting $c = 1$ results in $w = u_1 + u_2 - u_3$. Normalizing w results in

$$\begin{aligned} w \cdot w &= (u_1 + u_2 - u_3) \cdot (u_1 + u_2 - u_3) = -2 \\ \Longrightarrow e_3 &= \frac{1}{\sqrt{2}}(u_1 + u_2 - u_3), \quad \eta_3 = e_3 \cdot e_3 = -1. \end{aligned}$$

The components of the inner product in this o.n. basis are therefore

$$\mathsf{G}' = [g'_{ij}] = [e_i \cdot e_j] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

The index of the inner product is $1 - 2 = -1$.

Any pair of orthonormal bases $\{e_i\}$ and $\{e'_i\}$ are connected by a basis transformation

$$e_i = L^j{}_i e'_j,$$

such that

$$g_{ij} = e_i \cdot e_j = e'_i \cdot e'_j = g'_{ij} = \mathrm{diag}(\eta_1, \dots, \eta_n).$$

From Eq. (5.4) we have

$$g_{ij} = g_{kl} L^k{}_i L^l{}_j, \tag{5.8}$$

or its matrix equivalent

$$\mathsf{G} = \mathsf{L}^T \mathsf{G} \mathsf{L}. \tag{5.9}$$

For a Euclidean metric $\mathsf{G} = \mathsf{I}$, and L is an orthogonal transformation, while for a Minkoswkian metric with $n = 4$ the transformations are *Lorentz* transformations discussed in Section 2.7. As was shown in Chapter 2, these transformations form the groups $O(n)$ and $O(3, 1)$ respectively. The general pseudo-orthogonal inner product results in a group $O(p, q)$ of **pseudo-orthogonal transformations of type** (p, q).

Problems

Problem 5.1 Let $(V, \cdot)$ be a real Euclidean inner product space and denote the length of a vector $x \in V$ by $|x| = \sqrt{x \cdot x}$. Show that two vectors u and v are orthogonal iff $|u + v|^2 = |u|^2 + |v|^2$.

Problem 5.2 Let

$$\mathsf{G} = [g_{ij}] = [u_i \cdot u_j] = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

be the components of a real inner product with respect to a basis u_1, u_2, u_3. Use Gram–Schmidt orthogonalization to find an orthonormal basis e_1, e_2, e_3, expressed in terms of the vectors u_i, and find the index of this inner product.

Problem 5.3 Let G be the symmetric matrix of components of a real inner product with respect to a basis u_1, u_2, u_3,

$$\mathsf{G} = [g_{ij}] = [u_i \cdot u_j] = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -2 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Using Gram–Schmidt orthogonalization, find an orthonormal basis e_1, e_2, e_3 expressed in terms of the vectors u_i.

Problem 5.4 Define the concept of a 'symmetric operator' $S : V \to V$ as one that satisfies

$$(Su) \cdot v = u \cdot (Sv) \qquad \text{for all } u, v \in V.$$

Show that this results in the component equation

$$S^k_{\ i} g_{kj} = g_{ik} S^k_{\ j},$$

equivalent to the matrix equation

$$\mathsf{S}^T \mathsf{G} = \mathsf{G}\mathsf{S}.$$

Show that for an orthonormal basis in a Euclidean space this results in the usual notion of symmetry, but fails for pseudo-Euclidean spaces.

Problem 5.5 Let V be a Minkowskian vector space of dimension n with index $n-2$ and let $k \neq 0$ be a null vector ($k \cdot k = 0$) in V.

(a) Show that there is an orthonormal basis $e_1, \dots, e_n$ such that

$$k = e_1 - e_n.$$

(b) Show that if u is a 'timelike' vector, defined as a vector with negative magnitude $u \cdot u < 0$, then u is not orthogonal to k.
(c) Show that if v is a null vector such that $v \cdot k = 0$, then $v \propto k$.
(d) If $n \geq 4$ which of these statements generalize to a space of index $n-4$?

5.2 Complex inner product spaces

We now consider a complex vector space V, which in the first instance may be infinite dimensional. Vectors will continue to be denoted by lower case Roman letters such as u and v, but complex scalars will be denoted by Greek letters such as α, β, ... from the early part of the alphabet. The word **inner product**, or **scalar product**, on a complex vector space

V will be reserved for a map $V \times V \to \mathbb{C}$ that assigns to every pair of vectors $u, v \in V$ a complex scalar $\langle u \,|\, v\rangle$ satisfying

(IP1) $\langle u \,|\, v\rangle = \overline{\langle v \,|\, u\rangle}$.
(IP2) $\langle u \,|\, \alpha v + \beta w\rangle = \alpha\langle u \,|\, v\rangle + \beta\langle u \,|\, w\rangle$ for all complex numbers α, β.
(IP3) $\langle u \,|\, u\rangle \geq 0$ and $\langle u \,|\, u\rangle = 0$ iff $u = 0$.

The condition (IP1) implies $\langle u \,|\, u\rangle$ is always real, a necessary condition for (IP3) to make any sense. From (IP1) and (IP2)

$$\begin{aligned}\langle \alpha v + \beta w \,|\, u\rangle &= \overline{\langle u \,|\, \alpha v + \beta w\rangle} \\ &= \overline{\alpha\langle u \,|\, v\rangle + \beta\langle u \,|\, w\rangle} \\ &= \overline{\alpha}\overline{\langle u \,|\, v\rangle} + \overline{\beta}\,\overline{\langle w \,|\, v\rangle},\end{aligned}$$

so that

$$\langle \alpha v + \beta w \,|\, u\rangle = \overline{\alpha}\langle v \,|\, u\rangle + \overline{\beta}\langle w \,|\, u\rangle. \tag{5.10}$$

This property is often described by saying that the inner product is **antilinear** with respect to the first argument.

A complex vector space with an inner product will simply be called an **inner product space**. If V is finite dimensional it is often called a **finite dimensional Hilbert space**, but for infinite dimensional spaces the term *Hilbert space* only applies if the space is *complete* (see Chapter 13).

Mathematicians more commonly adopt a notation (u, v) in place of our angular bracket notation, and demand linearity in the *first* argument, with antilinearity in the second. Our conventions follow that which is most popular with physicists and takes its origins in Dirac's 'bra' and 'ket' terminology for quantum mechanics (see Chapter 14).

Example 5.4 On $\mathbb{C}^n$ set

$$\langle (\alpha_1, \dots, \alpha_n) \,|\, (\beta_1, \dots, \beta_n)\rangle = \sum_{i=1}^{n} \overline{\alpha_i}\beta_i.$$

Conditions (IP1)–(IP3) are easily verified. We shall see directly that this is the archetypal finite dimensional inner product space. Every finite dimensional inner product space has a basis such that the inner product takes this form.

Example 5.5 A complex-valued function $\varphi : [0, 1] \to \mathbb{C}$ is said to be *continuous* if both the real and imaginary parts of the function $\varphi(x) = f(x) + ig(x)$ are continuous. Let $\mathcal{C}[0, 1]$ be the set of continuous complex-valued functions on the real line interval $[0, 1]$, and define an inner product

$$\langle \varphi \,|\, \psi\rangle = \int_0^1 \overline{\varphi(x)}\psi(x)\,\mathrm{d}x.$$

Conditions (IP1) and (IP2) are simple to prove, but in order to show (IP3) it is necessary to show that

$$\int_0^1 |f(x)|^2 + |g(x)|^2\,\mathrm{d}x = 0 \implies f(x) = g(x) = 0, \quad \forall x \in [0, 1].$$

If $f(a) \neq 0$ for some $0 \leq a \leq 1$ then, by continuity, there exists an interval $[a - \epsilon, a]$ or an interval $[a, a + \epsilon]$ on which $|f(x)| > \frac{1}{2}|f(a)|$. Then

$$\int_0^1 |\varphi(x)|^2 \, dx > \tfrac{1}{2}\epsilon |f(a)|^2 + \int_0^1 |g(x)|^2 \, dx > 0.$$

Hence $f(x) = 0$ for all $x \in [0, 1]$. The proof that $g(x) = 0$ is essentially identical.

Example 5.6 A complex-valued function on the real line, $\varphi : \mathbb{R} \to \mathbb{C}$, is said to be square integrable if $|\varphi|^2$ is an integrable function on any closed interval of $\mathbb{R}$ and $\int_{-\infty}^{\infty} |\varphi(x)|^2 dx < \infty$. The set $L^2(\mathbb{R})$ of square integrable complex-valued functions on the real line is a complex vector space, for if α is a complex constant and φ and ψ are any pair of square integrable functions, then

$$\int_{-\infty}^{\infty} |\varphi(x) + \alpha\psi(x)|^2 \, dx \leq \int_{-\infty}^{\infty} |\varphi(x)|^2 \, dx + |\alpha|^2 \int_{-\infty}^{\infty} |\psi(x)|^2 \, dx < \infty.$$

On $L^2(\mathbb{R})$ define the inner product

$$\langle \varphi \,|\, \psi \rangle = \int_{-\infty}^{\infty} \overline{\varphi(x)} \psi(x) \, dx.$$

This is well-defined for any pair of square integrable functions φ and ψ for, after some algebraic manipulation, we find that

$$\overline{\varphi}\psi = \tfrac{1}{2}\big(|\varphi + \psi|^2 - i|\varphi + i\psi|^2 - (1 - i)(|\varphi|^2 + |\psi|^2)\big).$$

Hence the integral of the left-hand side is equal to a sum of integrals on the right-hand side, each of which has been shown to exist.

The properties (IP1) and (IP2) are trivial to show but the proof of (IP3) along the lines given in Example 5.5 will not suffice here since we do not stipulate continuity for the functions in $L^2(\mathbb{R})$. For example, the function $f(x)$ defined by $f(x) = 0$ for all $x \neq 0$ and $f(0) = 1$ is a positive non-zero function whose integral vanishes. The remedy is to 'identify' any two real functions f and g having the property that $\int_{-\infty}^{\infty} |f(x) - g(x)|^2 dx = 0$. Such a pair of functions will be said to be equal **almost everywhere**, and $L^2(\mathbb{R})$ must be interpreted as consisting of equivalence classes of complex-valued functions whose real and imaginary parts are equal almost everywhere. A more complete discussion will be given in Chapter 13, Example 13.4.

Exercise: Show that the relation $f \equiv g$ iff $\int_{-\infty}^{\infty} |f(x) - g(x)|^2 dx = 0$ is an equivalence relation on $L^2(\mathbb{R})$.

Norm of a vector

The **norm** of a vector u in an inner product space, denoted $\|u\|$, is defined to be the non-negative real number

$$\|u\| = \sqrt{\langle u \,|\, u \rangle} \geq 0. \tag{5.11}$$

From (IP2) and Eq. (5.10) it follows immediately that

$$\|\alpha u\| = |\alpha|\,\|u\|. \tag{5.12}$$

Theorem 5.4 (Cauchy–Schwarz inequality) *For any pair of vectors u, v in an inner product space*

$$|\langle u \,|\, v\rangle| \;\le\; \|u\|\,\|v\|. \tag{5.13}$$

Proof: By (IP3), (IP2) and Eq. (5.10) we have for all $\lambda \in \mathbb{C}$

$$\begin{aligned} 0 &\le \langle u + \lambda v \,|\, u + \lambda v\rangle \\ &= \langle u \,|\, u\rangle + \lambda\langle u \,|\, v\rangle + \overline{\lambda}\langle v \,|\, u\rangle + \lambda\overline{\lambda}\langle v \,|\, v\rangle. \end{aligned}$$

Substituting the particular value

$$\lambda = -\frac{\langle v \,|\, u\rangle}{\langle v \,|\, v\rangle}$$

gives the inequality

$$\begin{aligned} 0 &\le \langle u \,|\, u\rangle - \frac{\langle v \,|\, u\rangle\langle u \,|\, v\rangle}{\langle v \,|\, v\rangle} - \frac{\overline{\langle v \,|\, u\rangle}\langle v \,|\, u\rangle}{\langle v \,|\, v\rangle} + \frac{|\langle v \,|\, u\rangle|^2}{\langle v \,|\, v\rangle} \\ &= \langle u \,|\, u\rangle - \frac{|\langle v \,|\, u\rangle|^2}{\langle v \,|\, v\rangle}. \end{aligned}$$

Hence, from (IP1),

$$|\langle u \,|\, v\rangle|^2 = |\langle v \,|\, u\rangle|^2 \ge \langle u \,|\, u\rangle\langle v \,|\, v\rangle$$

and the desired result follows from (5.11) on taking the square roots of both sides of this inequality. ■

Corollary 5.5 *Equality in (5.13) can only result if u and v are proportional to each other,*

$$|\langle u \,|\, v\rangle| = \|u\|\,\|v\| \iff u = \alpha v, \qquad \text{for some } \alpha \in \mathbb{C}.$$

Proof: If $u = \alpha v$ then from (5.11) and (5.12) we have

$$|\langle u \,|\, v\rangle| = |\langle \alpha v \,|\, v\rangle| = |\alpha|\langle v \,|\, v\rangle = \|\alpha v\|\|v\| = \|u\|\|v\|.$$

Conversely, if $|\langle u \,|\, v\rangle| = \|u\|\,\|v\|$ then,

$$|\langle u \,|\, v\rangle|^2 = \|u\|^2\,\|v\|^2$$

and reversing the steps in the proof of Lemma 5.4 with inequalities replaced by equalities gives

$$\langle u + \lambda v \,|\, u + \lambda v\rangle = 0 \quad \text{where} \quad \lambda = -\frac{\langle v \,|\, u\rangle}{\langle v \,|\, v\rangle}.$$

By (IP3) we conclude that $u = -\lambda v$ and the proposition follows with $\alpha = -\lambda$. ■

Theorem 5.6 (Triangle inequality) *For any pair of vectors u and v in an inner product space,*

$$\|u+v\| \le \|u\| + \|v\|. \tag{5.14}$$

Proof:

$$\begin{aligned}(\|u+v\|)^2 &= \langle u+v \,|\, u+v\rangle \\ &= \langle u \,|\, u\rangle + \langle v \,|\, v\rangle + \langle v \,|\, u\rangle + \langle u \,|\, v\rangle \\ &= \|u\|^2 + \|v\|^2 + 2\mathrm{Re}(\langle u \,|\, v\rangle) \\ &\le \|u\|^2 + \|v\|^2 + 2|\langle u \,|\, v\rangle| \\ &\le \|u\|^2 + \|v\|^2 + 2\|u\|\|v\| \quad \text{by Eq. (5.13)} \\ &= (\|u\| + \|v\|)^2 .\end{aligned}$$

The triangle inequality (5.14) follows on taking square roots. ■

Orthonormal bases

Let V be a finite dimensional inner product space with basis $e_1, e_2, \dots, e_n$. Define the **components** of the inner product with respect to this basis to be

$$h_{ij} = \langle e_i \,|\, e_j\rangle = \overline{\langle e_j \,|\, e_i\rangle} = \overline{h_{ji}}. \tag{5.15}$$

The matrix of components $\mathsf{H} = [h_{ij}]$ is clearly hermitian,

$$\mathsf{H} = \mathsf{H}^\dagger \quad \text{where} \quad \mathsf{H}^\dagger = \overline{\mathsf{H}}^T .$$

Under a change of basis (5.3) we have

$$\langle e_i \,|\, e_j\rangle = \overline{A^k_{\ i}}\langle e'_k \,|\, e'_m\rangle A^m_{\ j}$$

and the components of the inner product transform as

$$h_{ij} = \overline{A^k_{\ i}} h'_{km} A^m_{\ j}. \tag{5.16}$$

An identical argument can be used to express the primed components in terms of unprimed components,

$$h'_{ij} = \langle e'_i \,|\, e'_j\rangle = \overline{A'^k_{\ i}} h_{km} A'^m_{\ j}. \tag{5.17}$$

These equations have matrix equivalents,

$$\mathsf{H} = \mathsf{A}^\dagger \mathsf{H}' \mathsf{A}, \qquad \mathsf{H}' = \mathsf{A}'^\dagger \mathsf{H} \mathsf{A}', \tag{5.18}$$

where $\mathsf{A}' = \mathsf{A}^{-1}$.

Exercise: Show that the hermitian nature of the matrix H is unchanged by a transformation (5.18).

Two vectors u and v are said to be **orthogonal** if $\langle u \,|\, v\rangle = 0$. A basis $e_1, e_2, \dots, e_n$ is called an **orthonormal basis** if the vectors all have unit norm and are orthogonal to each other,

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

Equivalently, a basis is orthonormal if the matrix of components of the inner product with respect to the basis is the unit matrix, $\mathsf{H} = \mathsf{I}$.

Starting with an arbitrary basis $\{u_1, u_2, \dots, u_n\}$, it is always possible to construct an orthonormal basis by a process known as **Schmidt orthonormalization**, which closely mirrors the Gram–Schmidt process for Euclidean inner products, outlined in Example 5.2. Sequentially, the steps are:

1. Set $f_1 = u_1$, then $e_1 = \dfrac{f_1}{\|f_1\|}$.
2. Set $f_2 = u_2 - \langle e_1 \,|\, u_2 \rangle e_1$, which is orthogonal to e_1 since

$$\langle e_1 \,|\, f_2 \rangle = \langle e_1 \,|\, u_2 \rangle - \langle e_1 \,|\, u_2 \rangle \|e_1\| = 0.$$

   Normalize f_2 by setting $e_2 = f_2/\|f_2\|$.
3. Set $f_3 = u_3 - \langle e_1 \,|\, u_3 \rangle e_1 - \langle e_2 \,|\, u_3 \rangle e_2$, which is orthogonal to both e_1 and e_2. Normalize to give $e_3 = \dfrac{f_3}{\|f_3\|}$.
4. Continue in this way until

$$f_n = u_n - \sum_{i=1}^{n-1} \langle e_i \,|\, u_n \rangle e_i \quad \text{and} \quad e_n = \frac{f_n}{\|f_n\|}.$$

Since each vector e_i is a unit vector and is orthogonal to all the e_j for $j < i$ defined by previous steps, they form an o.n. set. It is easily seen that any vector $v = v^i u_i$ of V is a linear combination of the e_i since each u_j is a linear combination of $e_1, \dots, e_j$. Hence the vectors $\{e_i\}$ form a basis by Theorem 3.3 since they span V and are n in number.

With respect to an orthonormal basis the inner product of any pair of vectors $u = u^i e_i$ and $v = v^j e_j$ is given by

$$\langle u \,|\, v \rangle = \langle u^i e_i \,|\, v^j e_j \rangle = \overline{u^i} v^j \langle e_i \,|\, e_j \rangle = \overline{u^i} v^j \delta_{ij}.$$

Hence

$$\langle u \,|\, v \rangle \sum_{i=1}^{n} \overline{u^i} v^i = \overline{u^1} v^1 + \overline{u^2} v^2 + \dots + \overline{u^n} v^n,$$

which is equivalent to the standard inner product defined on $\mathbb{C}^n$ in Example 5.4.

Example 5.7 Let an inner product have the following components in a basis u_1, u_2, u_3:

$$\begin{array}{lll} h_{11} = \langle u_1 \,|\, u_1 \rangle = 1 & h_{12} = \langle u_1 \,|\, u_2 \rangle = 0 & h_{13} = \langle u_1 \,|\, u_3 \rangle = \frac{1}{2}(1+i) \\ h_{21} = \langle u_2 \,|\, u_1 \rangle = 0 & h_{22} = \langle u_2 \,|\, u_2 \rangle = 2 & h_{23} = \langle u_2 \,|\, u_3 \rangle = 0 \\ h_{31} = \langle u_3 \,|\, u_1 \rangle = \frac{1}{2}(1-i) & h_{32} = \langle u_3 \,|\, u_2 \rangle = 0 & h_{33} = \langle u_3 \,|\, u_3 \rangle = 1. \end{array}$$

Before proceeding it is important to realize that this inner product does in fact satisfy the positive definite condition (IP3). This would not be true, for example, if we had given $h_{13} = \overline{h_{31}} = 1 + i$, for then the vector $v = u_1 - \dfrac{1-i}{\sqrt{2}} u_3$ would have negative norm $\langle v \,|\, v\rangle = 2(1 - \sqrt{2}) < 0$.

In the above inner product, begin by setting $e_1 = f_1 = u_1$. The next vector is

$$f_2 = u_2 - \langle e_1 \,|\, u_2\rangle e_1 = u_2, \qquad e_2 = \frac{u_2}{\|u_2\|} = \frac{1}{\sqrt{2}} u_2.$$

The last step is to set

$$f_3 = u_3 - \langle e_1 \,|\, u_3\rangle e_1 - \langle e_2 \,|\, u_3\rangle e_2 = u_3 - \tfrac{1}{2}(1+i)u_1$$

which has norm squared

$$\begin{aligned}(\|f_3\|)^2 &= \langle u_3 \,|\, u_3\rangle - \tfrac{1}{2}(1-i)\langle u_1 \,|\, u_3\rangle - \tfrac{1}{2}(1+i)\langle u_3 \,|\, u_1\rangle + \tfrac{1}{4}(1-i)(1+i)\langle u_1 \,|\, u_1\rangle \\ &= 1 - \tfrac{1}{4}(1-i)(1+i) - \tfrac{1}{4}(1+i)(1-i) + \tfrac{1}{2} = \tfrac{1}{2}.\end{aligned}$$

Hence $e_3 = f_3/\|f_3\| = \sqrt{2}(u_3 - (1+i)u_1)$ completes the orthonormal basis.

The Schmidt orthonormalization procedure actually provides a good method for proving positive definiteness, since the process breaks down at some stage, producing a vector with non-positive norm if the inner product does not satisfy (IP3).

Exercise: Try to perform the Schmidt orthonormalization on the above inner product suggested with the change $h_{13} = \overline{h_{31}} = 1 + i$, and watch it break down!

Unitary transformations

A linear operator $U : V \to V$ on an inner product space is said to be **unitary** if it preserves inner products,

$$\langle Uu \,|\, Uv\rangle = \langle u \,|\, v\rangle, \quad \forall\, u, v \in V. \tag{5.19}$$

Unitary operators clearly preserve the norm of any vector v,

$$\|Uv\| = \sqrt{\langle Uv \,|\, Uv\rangle} = \sqrt{\langle v \,|\, v\rangle} = \|v\|.$$

In fact it can be shown that a linear operator U is unitary if and only if it is norm preserving (see Problem 5.7).

A unitary operator U transforms any orthonormal basis $\{e_i\}$ into another o.n. basis $e'_i = Ue_i$, since

$$\langle e'_i \,|\, e'_j\rangle = \langle Ue_i \,|\, Ue_j\rangle = \langle e_i \,|\, e_j\rangle = \delta_{ij}. \tag{5.20}$$

The set $\{e'_1, \dots, e'_n\}$ is linearly independent, and is thus a basis, for if $\alpha^i e'_i = 0$ then $\langle e'_j \,|\, \alpha^i e'_i\rangle = \alpha^j = 0$. The map U is onto since every vector $u = u'^i e'_i = U(u'^i e_i)$, and one-to-one since $Uv = 0 \Rightarrow v^i e'_i = 0 \Rightarrow v = v^i e_i = 0$. Hence every unitary operator U is invertible.

With respect to an orthonormal basis $\{e_i\}$ the components of the linear transformation U, defined by $Ue_i = U^k_i e_k$, form a unitary matrix $\mathsf{U} = [U^k_i]$:

$$\begin{aligned}\delta_{ij} = \langle Ue_i \,|\, Ue_j\rangle &= \overline{U^k_i} U^m_j \langle e_k \,|\, e_m\rangle \\ &= \overline{U^k_i} U^m_j \delta_{km} = \sum_{k=1}^{n} \overline{U^k_i} U^k_j,\end{aligned}$$

or, in terms of matrices,

$$\mathsf{I} = \mathsf{U}^\dagger \mathsf{U}.$$

If $\{e_i\}$ and $\{e'_j\}$ are any pair of orthonormal bases, then the linear operator U defined by $e'_i = Ue_i$ is unitary since for any pair of vectors $u = u^i e_i$ and $v = v^j e_j$

$$\begin{aligned}\langle Uu \,|\, Uv\rangle &= \overline{u^i} v^j \langle e'_i \,|\, e'_j\rangle \\ &= \overline{u^i} v^j \delta_{ij} \\ &= \overline{u^i} v^j \langle e_i \,|\, e_j\rangle \;=\; \langle u \,|\, v\rangle.\end{aligned}$$

Thus all orthonormal bases are uniquely related by unitary transformations.

In the language of Section 3.6 this is the *active view*, wherein vectors are ‘physically’ moved about in the inner product space by the unitary transformation. In the related *passive view*, the change of basis is given by (5.3) – it is the *components* of vectors that are transformed, not the vectors themselves. If both bases are orthonormal the components of an inner product, given by Eq. (5.16), are $h_{ij} = h'_{ij} = \delta_{ij}$, and setting $A^k_i = U^k_i$ in Eq. (5.18) implies the matrix $\mathsf{U} = [U^k_i]$ is unitary,

$$\mathsf{I} = \mathsf{H} = \mathsf{U}^\dagger \mathsf{H}' \mathsf{U} = \mathsf{U}^\dagger \mathsf{I} \mathsf{U} = \mathsf{U}^\dagger \mathsf{U}.$$

Thus, from both the active and passive viewpoint, orthonormal bases are related by unitary matrices.

Problems

Problem 5.6 Show that the norm defined by an inner product satisfies the **parallelogram law**

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$

Problem 5.7 On an inner product space show that

$$4\langle u \,|\, v\rangle = \|u + v\|^2 - \|u - v\|^2 - i\|u + iv\|^2 + i\|u - iv\|^2.$$

Hence show that a linear transformation $U : V \to V$ is unitary iff it is norm preserving,

$$\langle Uu \,|\, Uv\rangle = \langle u \,|\, v\rangle, \quad \forall\, u, v \in V \iff \|Uv\| = \|v\|, \quad \forall\, v \in V.$$

Problem 5.8 Show that a pair of vectors u and v in a complex inner product space are orthogonal iff

$$\|\alpha u + \beta v\|^2 = \|\alpha u\|^2 + \|\beta v\|^2, \quad \forall\, \alpha, \beta \in \mathbb{C}.$$

Find a non-orthogonal pair of vectors u and v in a complex inner product space such that $\|u + v\|^2 = \|u\|^2 + \|v\|^2$.

Problem 5.9 Show that the formula

$$\langle \mathsf{A} \,|\, \mathsf{B} \rangle = \mathrm{tr}(\mathsf{B}\,\mathsf{A}^\dagger)$$

defines an inner product on the vector space of $m \times n$ complex matrices $M(m, n)$.

(a) Calculate $\|\mathsf{I}_n\|$ where I_n is the $n \times n$ identity matrix.
(b) What characterizes matrices orthogonal to I_n?
(c) Show that all unitary $n \times n$ matrices U have the same norm with respect to this inner product.

Problem 5.10 Let S and T be complex inner product spaces and let $U : S \to T$ be a linear map such that $\|Ux\| = \|x\|$. Prove that

$$\langle Ux \,|\, Uy \rangle = \langle x \,|\, y \rangle \quad \text{for all } x, y \in S.$$

Problem 5.11 Let V be a complex vector space with an 'indefinite inner product', defined as an inner product that satisfies (IP1), (IP2) but with (IP3) replaced by the non-singularity condition

(IP3′) $\langle u \,|\, v \rangle = 0$ for all $v \in V$ implies that $u = 0$.

(a) Show that similar results to Theorems 5.2 and 5.3 can be proved for such an indefinite inner product.
(b) If there are p $+1$'s along the diagonal and q -1's, find the defining relations for the group of transformations $U(p, q)$ between orthonormal basis.

Problem 5.12 If V is an inner product space, an operator $K : V \to V$ is called **self-adjoint** if

$$\langle u \,|\, Kv \rangle = \langle Ku \,|\, v \rangle$$

for any pair of vectors $u, v \in V$. Let $\{e_i\}$ be an arbitrary basis, having $\langle e_i \,|\, e_j \rangle = h_{ij}$, and set $Ke_k = K^j_{\ k} e_j$. Show that if $\mathsf{H} = [h_{ij}]$ and $\mathsf{K} = [K^k_{\ j}]$ then

$$\mathsf{H}\mathsf{K} = \mathsf{K}^\dagger \mathsf{H} = (\mathsf{H}\mathsf{K})^\dagger.$$

If $\{e_i\}$ is an orthonormal basis, show that K is a hermitian matrix.

5.3 Representations of finite groups

If G is a finite group, it turns out that every finite dimensional representation is equivalent to a representation by *unitary* transformations on an inner product space – known as a **unitary representation**. For, let T be a representation on any finite dimensional vector space V, and let $\{e_i\}$ be any basis of V. Define an inner product $(u|v)$ on V by setting $\{e_i\}$ to be an orthonormal set,

$$(u|v) = \sum_{i=1}^{n} \overline{u^i} v^i \quad \text{where} \quad u = u^i e_i, \ v = v^j e_j \quad (u^i, v^j \in \mathbb{C}). \tag{5.21}$$

Of course there is no reason why the linear transformations $T(g)$ should be unitary with respect to this inner product, but they *will* be unitary with respect to the inner product $\langle u \,|\, v \rangle$

formed by 'averaging over the group',

$$\langle u \,|\, v\rangle = \frac{1}{|G|}\sum_{a\in G}(T(a)u|T(a)v), \tag{5.22}$$

where $|G|$ is the *order* of the group G (the number of elements in G). This follows from

$$\begin{aligned}\langle T(g)u \,|\, T(g)v\rangle &= \frac{1}{|G|}\sum_{a\in G}(T(a)T(g)u|T(a)T(g)v)\\ &= \frac{1}{|G|}\sum_{a\in G}(T(ag)u|T(ag)v)\\ &= \frac{1}{|G|}\sum_{b\in G}(T(b)u|T(b)v)\\ &= \langle u \,|\, v\rangle\end{aligned}$$

since, as a ranges over the group G, so does $b = ag$ for any fixed $g \in G$.

Theorem 5.7 *Any finite dimensional representation of a finite group G is completely reducible into a direct sum of irreducible representations.*

Proof: Using the above device we may assume that the representation is unitary on a finite dimensional Hilbert space V with inner product $\langle\cdot\,|\,\cdot\rangle$. If W is a vector subspace of V, define its **orthogonal complement** $W^\perp$ to be the set of vectors orthogonal to W,

$$W^\perp = \{u \mid \langle u \,|\, w\rangle = 0,\ \forall w \in W\}.$$

$W^\perp$ is clearly a vector subspace, for if α is an arbitrary complex number then

$$u, v \in W^\perp \implies \langle u + \alpha v \,|\, w\rangle = \langle u \,|\, w\rangle + \alpha\langle v \,|\, w\rangle = 0 \implies u + \alpha v \in W^\perp.$$

By selecting an orthonormal basis such that the first dim W vectors belong to W, it follows that the remaining vectors of the basis span $W^\perp$. Hence W and $W^\perp$ are orthogonal and complementary subspaces, $H = W \oplus W^\perp$. If W is a G-invariant subspace, then $W^\perp$ is also G-invariant, For, if $u \in W^\perp$ then for any $w \in W$,

$$\begin{aligned}0 = \langle T(g)u \,|\, w\rangle &= \langle T(g)u \,|\, T(g)T(g)^{-1}w\rangle\\ &= \langle u \,|\, T(g^{-1})w\rangle \quad \text{since } T(g) \text{ is unitary}\\ &= 0\end{aligned}$$

since $T(g^{-1})w \in W$ by the G-invariance of W. Hence $T(g)u \in W^\perp$.

Now pick W to be the G-invariant subspace of V of smallest dimension, not counting the trivial subspace $\{0\}$. The representation induced on W must be irreducible since it can have no proper G-invariant subspaces, as they would need to have smaller dimension. If $W = V$ then the representation T is irreducible. If $W \neq V$ its orthogonal complement $W^\perp$ is either irreducible, in which case the proof is finished, or it has a non-trivial invariant subspace W'. Again pick the invariant subspace of smallest dimension and continue in this fashion until V is a direct sum of irreducible subspaces,

$$V = W \oplus W' \oplus W'' \oplus \cdots$$

The representation T decomposes into subrepresentations $T\big|_{W^{(i)}}$. ■

Orthogonality relations

The components of the matrices of irreducible group representatives satisfy a number of important orthogonality relationships, which are the cornerstone of the classification procedure of group representations. We will give just a few of these relations; others can be found in [4, 5].

Let T_1 and T_2 be irreducible representations of a finite group G on complex vector spaces V_1 and V_2 respectively. If $\{e_i \mid i = 1, \dots, n_1 = \dim V_1\}$ and $\{f_a \mid a = 1, \dots, n_2 = \dim V_2\}$ are bases of these two vector spaces, we will write the representative matrices as $\mathsf{T}_1(g) = [T_{(1)}{}^j{}_i]$ and $\mathsf{T}_2(g) = [T_{(2)}{}^b{}_a]$ where

$$T_1(g)e_i = T_{(1)}{}^j{}_i e_j \quad \text{and} \quad T_2(g)f_a = T_{(2)}{}^b{}_a f_b.$$

If $A : V_1 \to V_2$ is any linear map, define its 'group average' $\tilde{A} : V_1 \to V_2$ to be the linear map

$$\tilde{A} = \frac{1}{|G|} \sum_{g \in G} T_2(g) A T_1(g^{-1}).$$

Then if h is any element of the group G,

$$\begin{aligned} T_2(h)\,\tilde{A} T_1(h^{-1}) &= \frac{1}{|G|} \sum_{g \in G} T_2(hg) A T_1((hg)^{-1}) \\ &= \frac{1}{|G|} \sum_{g' \in G} T_2(g') A T_1((g')^{-1}) \\ &= \tilde{A}. \end{aligned}$$

Hence $\tilde{A}$ is an intertwining operator,

$$T_2(h)\tilde{A} = \tilde{A} T_1(h) \qquad \text{for all } h \in G,$$

and by Schur's lemma, Theorem 4.6, if $T_1 \nsim T_2$ then $\tilde{A} = 0$. On the other hand, from the corollary to Schur's lemma, 4.7, if $V_1 = V_2 = V$ and $T_1 = T_2$ then $\tilde{A} = c\,\mathrm{id}_V$. The matrix version of this equation with respect to any basis of V is $\tilde{\mathsf{A}} = c\,\mathsf{I}$, and taking the trace gives

$$c = \frac{1}{n} \operatorname{tr} \tilde{\mathsf{A}}.$$

However

$$\begin{aligned} \operatorname{tr} \tilde{\mathsf{A}} &= \frac{1}{|G|} \operatorname{tr} \sum_{g \in G} \mathsf{T}(g)\mathsf{A}\mathsf{T}^{-1}(g) \\ &= \frac{1}{|G|} \sum_{g \in G} \operatorname{tr}(\mathsf{T}^{-1}(g)\mathsf{T}(g)\mathsf{A}) \\ &= \frac{1}{|G|} \sum_{g \in G} \operatorname{tr} \mathsf{A} = \operatorname{tr} \mathsf{A}, \end{aligned}$$

whence

$$c = \frac{1}{n} \operatorname{tr} \mathsf{A}.$$

If $T_1 \not\sim T_2$, expressing A and $\tilde{A}$ in terms of the bases e_1 and f_a,

$$Ae_i = A^a_{\ i} f_a \quad \text{and} \quad \tilde{A}e_i = \tilde{A}^a_{\ i} f_a$$

the above consequence of Schur's lemma can be written

$$\tilde{A}^a_{\ i} = \frac{1}{|G|} \sum_{g\in G} T_{(2)}{}^a_{\ b}(g) A^b_{\ j} T_{(1)}{}^j_{\ i}(g^{-1}) = 0.$$

As A is an arbitrary operator the matrix elements $A^b_{\ j}$ are arbitrary complex numbers, so that

$$\frac{1}{|G|} \sum_{g\in G} T_{(2)}{}^a_{\ b}(g) T_{(1)}{}^j_{\ i}(g^{-1}) = 0. \tag{5.23}$$

If $T_1 = T_2 = T$ and $n = \dim V$ is the degree of the representation we have

$$\tilde{A}^j_{\ i} = \frac{1}{|G|} \sum_{g\in G} T^j_{\ k}(g) A^k_{\ l} T^l_{\ i}(g^{-1}) = \frac{1}{n} A^k_{\ k} \delta^j_{\ i}.$$

As $A^k_{\ l}$ are arbitrary,

$$\frac{1}{|G|} \sum_{g\in G} T^j_{\ k}(g) T^l_{\ i}(g^{-1}) = \frac{1}{n} \delta^j_{\ i} \delta^l_{\ k}. \tag{5.24}$$

If $\langle \cdot \,|\, \cdot \rangle$ is the invariant inner product defined by a representation T on a vector space V by (5.22), and $\{e_i\}$ is any basis such that

$$\langle e_i \,|\, e_j \rangle = \delta_{ji},$$

then the unitary condition $\langle T(g)u \,|\, T(g)v \rangle = \langle u \,|\, v \rangle$ implies

$$\sum_k \overline{T_{ki}(g)} T_{kj}(g) = \delta_{ij},$$

where indices on T are all lowered. In matrices

$$\mathsf{T}^\dagger(g)\mathsf{T}(g) = \mathsf{I},$$

whence

$$\mathsf{T}(g^{-1}) = (\mathsf{T}(g))^{-1} = \mathsf{T}^\dagger(g),$$

or equivalently

$$T_{ji}(g^{-1}) = \overline{T_{ij}(g)}. \tag{5.25}$$

Substituting this relation for T_1 in place of T into (5.23), with all indices now lowered, gives

$$\frac{1}{|G|} \sum_{g\in G} \overline{T_{(1)ij}(g)} T_{(2)ab}(g) = 0. \tag{5.26}$$

Similarly if $T_1 = T_2 = T$, Eqs. (5.25) and (5.24) give

$$\frac{1}{|G|} \sum_{g\in G} \overline{T_{ij}(g)} T_{kl}(g) = \frac{1}{n} \delta_{ik} \delta_{jl}. \tag{5.27}$$

The left-hand sides of Eqs. (5.26) and (5.27) have the appearance of an inner product, and this is in fact so. Let $\mathcal{F}(G)$ be the space of all complex-valued functions on G,

$$\mathcal{F}(G) = \{\phi \mid \phi : G \to \mathbb{C}\}$$

with inner product

$$(\phi, \psi) = \frac{1}{|G|} \sum_{a \in G} \overline{\phi(a)} \psi(a). \tag{5.28}$$

It is easy to verify that the requirements (IP1)–(IP3) hold for this inner product, namely

$$\begin{aligned}
(\phi, \psi) &= \overline{(\psi, \phi)}, \\
(\phi, \alpha\psi) &= \alpha(\phi, \psi), \\
(\phi, \phi) &\geq 0 \quad \text{and} \quad (\phi, \phi) = 0 \quad \text{iff} \quad \phi = 0.
\end{aligned}$$

The matrix components T_{ji} of any representation with respect to an o.n. basis form a set of n^2 complex-valued functions on G, and Eqs. (5.26) and (5.27) read

$$(T_{(1)ij},\ T_{(2)ab}) = 0 \quad \text{if} \quad T_1 \nsim T_2, \tag{5.29}$$

and

$$(T_{ij},\ T_{kl}) = \frac{1}{n} \delta_{ik} \delta_{jl}. \tag{5.30}$$

Example 5.8 Consider the group S_3 with notation as in Example 4.8. The invariant inner product (5.22) on the space spanned by e_1, e_2 and e_3 is given by

$$\begin{aligned}
\langle u \mid v \rangle &= \tfrac{1}{6} \sum_{\pi} \sum_{i=1}^{3} \overline{(T(\pi)u)^i} (T(\pi)v)^i \\
&= \tfrac{1}{6} \big(\overline{u^1} v^1 + \overline{u^2} v^2 + \overline{u^3} v^3 + \overline{u^3} v^3 + \overline{u^1} v^1 + \overline{u^2} v^2 + \dots \big) \\
&= \overline{u^1} v^1 + \overline{u^2} v^2 + \overline{u^3} v^3.
\end{aligned}$$

Hence $\{e_i\}$ forms an orthonormal basis for this inner product,

$$h_{ij} = \langle e_i \mid e_j \rangle = \delta_{ij}.$$

It is only because S_3 runs through *all* permutations that the averaging process gives the same result as the inner product defined by (5.21). A similar conclusion would hold for the action of S_n on an n-dimensional space spanned by $\{e_1, \dots, e_n\}$, but these vectors would *not* in general be orthonormal with respect to the inner product (5.22) defined by an arbitrary subgroup of S_n.

As seen in Example 4.8, the vector $f_1 = e_1 + e_2 + e_3$ spans an invariant subspace with respect to this representation of S_3. As in Example 4.8, the vectors f_1, $f_2 = e_1 - e_2$ and $f_3 = e_1 + e_2 - 2e_3$ are mutually orthogonal,

$$\langle f_1 \mid f_2 \rangle = \langle f_1 \mid f_3 \rangle = \langle f_2 \mid f_3 \rangle = 0.$$

Hence the subspace spanned by f_2 and f_3 is orthogonal to f_1, and from the proof of Theorem 5.7, it is also invariant. Form an o.n. set by normalizing their lengths to unity,

$$f_1' = \frac{f_1}{\sqrt{3}}, \qquad f_2' = \frac{f_2}{\sqrt{2}}, \qquad f_3' = \frac{f_3}{\sqrt{6}}.$$

The representation T_1 on the one-dimensional subspace spanned by f_1' is clearly the trivial one, whereby every group element is mapped to the number 1,

$$T_1(\pi) = 1 \qquad \text{for all } \pi \in S_3.$$

The matrices of the representation T_2 on the invariant subspace spanned by $h_1 = f_2'$ and $h_2 = f_3'$ are easily found from the 2×2 parts of the matrices given in Example 4.8 by transforming to the renormalized basis,

$$\mathsf{T}_2(e) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \mathsf{T}_2(\pi_1) = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \qquad \mathsf{T}_2(\pi_2) = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix},$$

$$\mathsf{T}_2(\pi_3) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \mathsf{T}_2(\pi_4) = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \qquad \mathsf{T}_2(\pi_5) = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

It is straightforward to verify (5.29):

$$\big(T_{(1)11},\ T_{(2)11}\big) = \frac{1}{6}\Big(1 - \frac{1}{2} - \frac{1}{2} - 1 + \frac{1}{2} + \frac{1}{2}\Big) = 0,$$

$$\big(T_{(1)11},\ T_{(2)12}\big) = \frac{1}{6}\Big(0 - \frac{\sqrt{3}}{2} + \frac{\sqrt{3}}{2} + 0 + \frac{\sqrt{3}}{2} - \frac{\sqrt{3}}{2}\Big) = 0, \text{ etc.}$$

From the exercise following Example 4.8 the representation T_2 is an irreducible representation with $n = 2$ and the relations (5.30) are verified as follows:

$$\big(T_{(2)11},\ T_{(2)11}\big) = \frac{1}{6}\Big(1^2 + \Big(-\frac{1}{2}\Big)^2 + \Big(-\frac{1}{2}\Big)^2 + (-1)^2 + \frac{1}{2}^2 + \frac{1}{2}^2\Big) = \frac{3}{6} = \frac{1}{2}\delta_{11}\delta_{11} = \frac{1}{2}$$

$$\big(T_{(2)12},\ T_{(2)12}\big) = \frac{1}{6}\Big(\Big(-\frac{\sqrt{3}}{2}\Big)^2 + \Big(\frac{\sqrt{3}}{2}\Big)^2 + \Big(\frac{\sqrt{3}}{2}\Big)^2 + \Big(-\frac{\sqrt{3}}{2}\Big)^2\Big) = \frac{1}{2} = \frac{1}{2}\delta_{11}\delta_{22}$$

$$\big(T_{(2)11},\ T_{(2)12}\big) = \frac{1}{6}\Big(1 \cdot 0 - \frac{1}{2} \cdot \Big(-\frac{\sqrt{3}}{2}\Big) - \frac{1}{2} \cdot \frac{\sqrt{3}}{2} - 1 \cdot 0 + \frac{1}{2} \cdot \frac{\sqrt{3}}{2} + \frac{1}{2} \cdot \Big(-\frac{\sqrt{3}}{2}\Big)\Big)$$

$$= 0 = \frac{1}{2}\delta_{11}\delta_{12}, \text{ etc.}$$

Theorem 5.8 *There are a finite number N of inequivalent irreducible representations of a finite group, and*

$$\sum_{\mu=1}^{N} (n_\mu)^2 \le |G|, \tag{5.31}$$

where $n_\mu = \dim V_\mu$ are the degrees of the inequivalent representations.

Proof: Let $T_1 : G \to GL(V_1),\ T_2 : G \to GL(V_2), \dots$ be inequivalent irreducible representations of G. If the basis on each vector space V_μ is chosen to be orthonormal with respect to the inner product $\langle \cdot \,|\, \cdot \rangle$ for each $\mu = 1, 2, \dots$, then (5.29) and (5.30) may be summarized as the single equation

$$\left(T_{(\mu)ij},\ T_{(\nu)ab}\right) = \frac{1}{n_\mu}\delta_{\mu\nu}\delta^{ia}\delta_{jb}.$$

Hence for each μ the $T_{(\mu)ij}$ consist of $(n_\mu)^2$ mutually orthogonal functions in $\mathcal{F}(G)$ that are orthogonal to all $T_{(\nu)ab}$ for $\nu \neq \mu$. There cannot therefore be more than $\dim \mathcal{F}(G) = |G|$ of these, giving the desired inequality (5.31). Clearly there are at most a finite number N of such representations. ■

It may in fact be shown that the inequality (5.31) can be replaced by equality, a very useful identity in the enumeration of irreducible representations of a finite group. Details of the proof as well as further orthogonality relations and applications of group representation theory to physics may be found in [4, 5].

Problems

Problem 5.13 For a function $\phi : G \to \mathbb{C}$, if we set $g\phi$ to be the function $(g\phi)(a) = \phi(g^{-1}a)$ show that $(gg')\phi = g(g'\phi)$. Show that the inner product (5.28) is G-invariant, $(g\phi, g\psi) = (\phi, \psi)$ for all $g \in G$.

Problem 5.14 Let the **character** of a representation T of a group G on a vector space V be the function $\chi : G \to \mathbb{C}$ defined by

$$\chi(g) = \operatorname{tr} T(g) = T^i_i(g).$$

(a) Show that the character is independent of the choice of basis and is a member of $\mathcal{F}(G)$, and that characters of equivalent representations are identical. Show that $\chi(e) = \dim V$.

(b) Any complex-valued function on G that is constant on conjugacy classes (see Section 2.4) is called a **central function**. Show that characters are central functions.

(c) Show that with respect to the inner product (5.28), characters of any pair of inequivalent irreducible representations $T_1 \not\sim T_2$ are orthogonal to each other, $(\chi_1,\ \chi_2) = 0$, while the character of any irreducible representation T has unit norm $(\chi,\ \chi) = 1$.

(d) From Theorem 5.8 and Theorem 5.7 every unitary representation T can be decomposed into a direct sum of inequivalent irreducible unitary representations $T_\mu : G \to GL(V_\mu)$,

$$T \sim m_1 T_1 \oplus m_2 T_2 \oplus \cdots \oplus m_N T_N \quad (m_\mu \geq 0).$$

Show that the **multiplicities** m_μ of the representations T_μ are given by

$$m_\mu = (\chi,\ \chi_\mu) = \frac{1}{|G|}\sum_{g\in G} \overline{\chi(g)}\chi_\mu(g)$$

and T is irreducible if and only if its character has unit magnitude, $(\chi,\ \chi) = 1$. Show that T and T' have no irreducible representations in common in their decompositions if and only if their characters are orthogonal.

References

[1] P. R. Halmos. *Finite-dimensional Vector Spaces*. New York, D. Van Nostrand Company, 1958.
[2] P. R. Halmos. *Introduction to Hilbert Space*. New York, Chelsea Publishing Company, 1951.
[3] L. H. Loomis and S. Sternberg. *Advanced Calculus*. Reading, Mass., Addison-Wesley, 1968.
[4] M. Hammermesh. *Group Theory and its Applications to Physical Problems*. Reading, Mass., Addison-Wesley, 1962.
[5] S. Sternberg. *Group Theory and Physics*. Cambridge, Cambridge University Press, 1994.