14.2 Orthonormal Bases

In an inner-product space, we can demand that basis vectors have special properties. A list of vectors is said to be **orthonormal** if all vectors have norm one and are pairwise orthogonal. If $(e_1, \ldots, e_n)$ is a list of orthonormal vectors in V, then

$$\langle e_i, e_j \rangle \;=\; \delta_{ij}, \;\forall\, i, j = 1, \ldots, n. \tag{14.2.1}$$

We also have a simple expression for the norm of $a_1e_1 + \cdots + a_ne_n$, with a_i constants in the relevant field:

$$\begin{aligned} \|a_1e_1 + \cdots + a_ne_n\|^2 &= \langle a_1e_1 + \cdots + a_ne_n,\, a_1e_1 + \cdots + a_ne_n \rangle \\ &= \langle a_1e_1, a_1e_1 \rangle + \cdots + \langle a_ne_n, a_ne_n \rangle \\ &= |a_1|^2 + \cdots + |a_n|^2. \end{aligned} \tag{14.2.2}$$

This result implies the somewhat nontrivial fact that *the vectors in any orthonormal list are linearly independent*. Indeed, if $a_1e_1 + \cdots + a_ne_n = 0$, then its norm squared is zero and so is $|a_1|^2 + \cdots + |a_n|^2$. This implies all $a_i = 0$, thus proving the claim.

An **orthonormal basis** of V is a list of orthonormal vectors that is also a basis for V. Let $(e_1, \ldots, e_n)$ denote an orthonormal basis. Then any vector v can be written as

$$v \;=\; a_1e_1 + \cdots + a_ne_n, \tag{14.2.3}$$

for some constants a_i that can be calculated as follows:

$$a_i \;=\; \langle e_i, v \rangle. \tag{14.2.4}$$

Indeed,

$$\langle e_i, v \rangle = \sum_j \langle e_i, a_je_j \rangle = \sum_j a_j \langle e_i, e_j \rangle = \sum_j a_j \delta_{ij} = a_i. \tag{14.2.5}$$

Therefore, any vector v can be written as

$$v = \langle e_1, v\rangle e_1 + \cdots + \langle e_n, v\rangle e_n = \sum_i \langle e_i, v\rangle e_i. \tag{14.2.6}$$

To find an orthonormal basis on an inner-product space V, we can start with any basis and follow a procedure that yields the desired orthonormal basis. A little more generally, the Gram-Schmidt procedure achieves the following:

> *Gram-Schmidt: Given a list* $(v_1, \ldots, v_n)$ *of linearly independent vectors in V, one can construct a list* $(e_1, \ldots, e_n)$ *of orthonormal vectors such that both lists span the same subspace of V.*

The Gram-Schmidt procedure goes as follows. You take e_1 to be v_1, scaled to have unit norm:

$$e_1 = \frac{v_1}{\|v_1\|}. \tag{14.2.7}$$

Clearly, $\langle e_1, e_1\rangle = 1$. Then take

$$f_2 \equiv v_2 + \alpha\, e_1 \tag{14.2.8}$$

and fix the constant α to make f_2 orthogonal to e_1: $\langle e_1, f_2\rangle = 0$. The answer, as you can check, is

$$f_2 = v_2 - \langle e_1, v_2\rangle e_1. \tag{14.2.9}$$

This vector, divided by its norm, is set equal to e_2, the second vector in our orthonormal list:

$$e_2 = \frac{v_2 - \langle e_1, v_2\rangle e_1}{\|v_2 - \langle e_1, v_2\rangle e_1\|}. \tag{14.2.10}$$

In fact, we can write the general vector in a recursive fashion. If we have orthonormal $e_1, e_2, \ldots, e_{j-1}$, we can write the next orthonormal vector e_j as follows:

$$e_j = \frac{v_j - \langle e_1, v_j\rangle e_1 - \cdots - \langle e_{j-1}, v_j\rangle e_{j-1}}{\left\|v_j - \langle e_1, v_j\rangle e_1 - \cdots - \langle e_{j-1}, v_j\rangle e_{j-1}\right\|}. \tag{14.2.11}$$

It should be clear to you by inspection that this vector, as required, satisfies $\langle e_i, e_j\rangle = 0$ for all $i < j$ and that it has unit norm. The Gram-Schmidt procedure is quite practical.

If we have an orthonormal basis $(e_1, \ldots, e_n)$ for a vector space V, there is a simple formula for the matrix elements of any operator $T \in \mathcal{L}(V)$. Consider the inner product

$$\langle e_i, Te_j \rangle = \langle e_i, \sum_k T_{kj} e_k \rangle = \sum_k T_{kj} \langle e_i, e_k \rangle = \sum_k T_{kj} \delta_{ik} = T_{ij}. \tag{14.2.12}$$

We thus have

$$\boxed{T_{ij} = \langle e_i, Te_j \rangle \text{ in an orthonormal basis.}} \tag{14.2.13}$$

This formula is so familiar that one could be led to believe that an inner product is required to define the matrix elements of an operator. We know better: a basis suffices. If the basis is orthonormal, the simple formula above is available.

Example 14.4 *An orthonormal basis of Hermitian matrices in two dimensions.*

Two-by-two Hermitian matrices are important in quantum mechanics, mostly because they define the most general Hamiltonian for a quantum system with two basis states. In fact, we showed in example 13.7 that 2×2 Hermitian matrices form a real vector space of dimension four, with basis vectors $(\mathbb{1}, \sigma_1, \sigma_2, \sigma_3)$. Now we can demonstrate that with the inner product in (14.1.27) this is in fact an orthonormal basis of operators. As noted in exercise 14.2, for Hermitian matrices the inner product is real and thus suitable for a real vector space.

To manipulate the operators, it is convenient to use an index μ that runs over four values, zero to three, so that we can use the value zero for the identity matrix:

$$\sigma_\mu = \{\sigma_0, \sigma_1, \sigma_2, \sigma_3\}, \quad \sigma_0 = \mathbb{1}, \quad \mu = 0, 1, 2, 3. \tag{14.2.14}$$

With the inner product introduced before, the basis vectors are orthonormal:

$$\langle \sigma_\mu, \sigma_\nu \rangle = \tfrac{1}{2}\text{tr}(\sigma_\mu^\dagger \sigma_\nu) = \tfrac{1}{2}\text{tr}(\sigma_\mu \sigma_\nu) = \delta_{\mu\nu}. \tag{14.2.15}$$

Here we used the Hermiticity of σ_μ, and the last step, giving us the Kronecker delta, is verified by explicit computation:

$$\begin{aligned}
\langle\sigma_0,\sigma_0\rangle &= \tfrac{1}{2}\mathrm{tr}(\mathbb{1}\mathbb{1}) = \tfrac{1}{2}\mathrm{tr}\mathbb{1} = \tfrac{1}{2}\cdot 2 = 1,\\
\langle\sigma_0,\sigma_i\rangle &= \tfrac{1}{2}\mathrm{tr}(\mathbb{1}\sigma_i) = \tfrac{1}{2}\mathrm{tr}\sigma_i = 0,\\
\langle\sigma_i,\sigma_j\rangle &= \tfrac{1}{2}\mathrm{tr}(\sigma_i\sigma_j) = \tfrac{1}{2}\mathrm{tr}(\delta_{ij}\mathbb{1} + i\epsilon_{ijk}\sigma_k) = \tfrac{1}{2}\delta_{ij}\mathrm{tr}\mathbb{1} = \delta_{ij}.
\end{aligned} \tag{14.2.16}$$

On account of (14.2.3) and (14.2.4), we know that any Hermitian matrix $M = M^\dagger$ can be written as

$$M = m_0\mathbb{1} + m_1\sigma_1 + m_2\sigma_2 + m_3\sigma_3, \quad \text{with} \quad m_\mu = \langle\sigma_\mu, M\rangle = \tfrac{1}{2}\mathrm{tr}(\sigma_\mu M). \tag{14.2.17}$$

As a result, we have

$$M = \tfrac{1}{2}(\mathrm{tr}M)\,\mathbb{1} + \tfrac{1}{2}\sum_{i=1}^{3}\mathrm{tr}(\sigma_i M)\,\sigma_i, \tag{14.2.18}$$

showing how to write any Hermitian matrix as a superposition of $\mathbb{1}$ and Pauli matrices.

□