4 Operator Algebra

Chapter 3 introduced algebras, i.e., vector spaces in which one can multiply two vectors to obtain a third vector. In this chapter, we want to investigate the algebra of linear transformations.

4.1 Algebra of $\mathrm{End}(\mathcal{V})$

The product in the algebra of the endomorphisms $\mathrm{End}(\mathcal{V})$ of a vector space $\mathcal{V}$ is defined as the composition of maps. In addition to the zero element, which is present in all algebras, $\mathrm{End}(\mathcal{V})$ has an identity element, $\mathbf{1}$, which satisfies the relation $\mathbf{1}|a\rangle = |a\rangle$ for all $|a\rangle \in \mathcal{V}$. Thus, $\mathrm{End}(\mathcal{V})$ is a unital algebra. With $\mathbf{1}$ in our possession, we can ask whether it is possible to find an operator $\mathbf{T}^{-1}$ with the property that $\mathbf{T}^{-1}\mathbf{T} = \mathbf{T}\mathbf{T}^{-1} = \mathbf{1}$. Generally speaking, only bijective mappings have inverses. Therefore, *only automorphisms of a vector space are invertible*.

Example 4.1.1 Let the linear operator $\mathbf{T} : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by

$$\mathbf{T}(x_1, x_2, x_3) = (x_1 + x_2, x_2 + x_3, x_1 + x_3).$$

We want to see whether $\mathbf{T}$ is invertible and, if so, find its inverse. $\mathbf{T}$ has an inverse if and only if it is bijective. By the comments after Theorem 2.3.13 this is the case if and only if $\mathbf{T}$ is either surjective or injective. The latter is equivalent to $\ker \mathbf{T} = |0\rangle$. But $\ker \mathbf{T}$ is the set of all vectors satisfying $\mathbf{T}(x_1, x_2, x_3) = (0, 0, 0)$, or

$$x_1 + x_2 = 0, \qquad x_2 + x_3 = 0, \qquad x_1 + x_3 = 0.$$

The reader may check that the unique solution to these equations is $x_1 = x_2 = x_3 = 0$. Thus, the only vector belonging to $\ker \mathbf{T}$ is the zero vector. Therefore, $\mathbf{T}$ has an inverse.

To find $\mathbf{T}^{-1}$ apply $\mathbf{T}^{-1}\mathbf{T} = \mathbf{1}$ to (x_1, x_2, x_3):

$$(x_1, x_2, x_3) = \mathbf{T}^{-1}\mathbf{T}(x_1, x_2, x_3) = \mathbf{T}^{-1}(x_1 + x_2, x_2 + x_3, x_1 + x_3).$$

S. Hassani, *Mathematical Physics*, DOI 10.1007/978-3-319-01195-0_4,

© Springer International Publishing Switzerland 2013

This equation demonstrates how $\mathbf{T}^{-1}$ acts on vectors. To make this more apparent, we let $x_1 + x_2 = x$, $x_2 + x_3 = y$, $x_1 + x_3 = z$, solve for x_1, x_2, and x_3 in terms of x, y, and z, and substitute in the preceding equation to obtain

$$\mathbf{T}^{-1}(x, y, z) = \frac{1}{2}(x - y + z, x + y - z, -x + y + z).$$

Rewriting this equation in terms of x_1, x_2, and x_3 gives

$$\mathbf{T}^{-1}(x_1, x_2, x_3) = \frac{1}{2}(x_1 - x_2 + x_3, x_1 + x_2 - x_3, -x_1 + x_2 + x_3).$$

We can easily verify that $\mathbf{T}^{-1}\mathbf{T} = \mathbf{1}$ and that $\mathbf{T}\mathbf{T}^{-1} = \mathbf{1}$.

Since $\mathrm{End}(\mathcal{V})$ is associative, Theorem 3.1.2 applies to it. Nevertheless, we restate it in the context of operators as a corollary in which we also include a generalization of Theorem 2.3.19:

Corollary 4.1.2 *The inverse of a linear operator is unique. If* $\mathbf{T}$ *and* $\mathbf{S}$ *are two invertible linear operators on* $\mathcal{V}$, *then* $\mathbf{TS}$ *is also invertible and*

$$(\mathbf{TS})^{-1} = \mathbf{S}^{-1}\mathbf{T}^{-1}.$$

The following proposition, whose straightforward proof is left as an exercise for the reader, turns out to be useful later on:

Proposition 4.1.3 *An endomorphism* $\mathbf{T} \in \mathrm{End}(\mathcal{V})$ *is invertible iff it sends a basis of* $\mathcal{V}$ *onto another basis of* $\mathcal{V}$.

Let $\mathcal{V}_1$ and $\mathcal{V}_2$ be vector spaces and $\mathcal{L}_1(\mathcal{V}_1)$ and $\mathcal{L}_2(\mathcal{V}_2)$ the set of their endomorphisms. A natural definition of $\mathcal{L}(\mathcal{V}_1 \otimes \mathcal{V}_2)$ is given by

$$\mathcal{L}(\mathcal{V}_1 \otimes \mathcal{V}_2) \cong (\mathcal{L}_1 \otimes \mathcal{L}_2)(\mathcal{V}_1 \otimes \mathcal{V}_2) \cong \mathcal{L}_1(\mathcal{V}_1) \otimes \mathcal{L}_2(\mathcal{V}_2). \tag{4.1}$$

In particular, if $\mathcal{V}_1 = \mathbb{C}$, $\mathcal{V}_2 = \mathcal{V}$ is a *real* vector space, then

$$\mathcal{L}(\mathbb{C} \otimes \mathcal{V}) \cong \mathcal{L}(\mathcal{V}^{\mathbb{C}}), \tag{4.2}$$

where $\mathcal{V}^{\mathbb{C}}$ is the complexification of $\mathcal{V}$ as given in Definition 2.4.8. It is important to note that

$$\mathcal{L}(\mathbb{C} \otimes \mathcal{V}) \ncong \mathbb{C} \otimes \mathcal{L}(\mathcal{V})$$

because $\mathcal{L}(\mathbb{C}) \ncong \mathbb{C}$.

4.1.1 Polynomials of Operators

From Sect. 3.6, we know that we can construct polynomials of $\mathbf{T} \in \mathrm{End}(\mathcal{V})$ such as

$$p(\mathbf{T}) = \alpha_0\mathbf{1} + \alpha_1\mathbf{T} + \alpha_2\mathbf{T}^2 + \cdots + \alpha_n\mathbf{T}^n.$$

We shall use these polynomials as starting points for constructing functions of operators.

Example 4.1.4 Let $\mathbf{T}_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear operator that rotates vectors in the xy-plane through the angle θ, that is,

$$\mathbf{T}_\theta(x, y) = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta).$$

We are interested in powers of $\mathbf{T}_\theta$:

$$\begin{aligned}
\mathbf{T}_\theta^2(x, y) &= \mathbf{T}_\theta\big(\overbrace{x\cos\theta - y\sin\theta}^{x'}, \overbrace{x\sin\theta + y\cos\theta}^{y'}\big) \\
&= \big(x'\cos\theta - y'\sin\theta, x'\sin\theta + y'\cos\theta\big) \\
&= \big((x\cos\theta - y\sin\theta)\cos\theta - (x\sin\theta + y\cos\theta)\sin\theta, \\
&\qquad (x\cos\theta - y\sin\theta)\sin\theta + (x\sin\theta + y\cos\theta)\cos\theta\big) \\
&= (x\cos 2\theta - y\sin 2\theta, x\sin 2\theta + y\cos 2\theta).
\end{aligned}$$

Thus, $\mathbf{T}^2$ rotates (x, y) by 2θ. Similarly, one can show that

$$\mathbf{T}_\theta^3(x, y) = (x\cos 3\theta - y\sin 3\theta, x\sin 3\theta + y\cos 3\theta),$$

and in general, $\mathbf{T}_\theta^n(x, y) = (x\cos n\theta - y\sin n\theta, x\sin n\theta + y\cos n\theta)$, which shows that $\mathbf{T}_\theta^n$ is a rotation of (x, y) through the angle $n\theta$, that is, $\mathbf{T}_\theta^n = \mathbf{T}_{n\theta}$. This result could have been guessed because $\mathbf{T}_\theta^n$ is equivalent to rotating (x, y) n times, each time by an angle θ.

Negative powers of an invertible linear operator $\mathbf{T}$ are defined by $\mathbf{T}^{-m} = (\mathbf{T}^{-1})^m$. The exponents of $\mathbf{T}$ satisfy the usual rules.[1] In particular, for any two integers m and n (positive or negative), $\mathbf{T}^m\mathbf{T}^n = \mathbf{T}^{m+n}$ and $(\mathbf{T}^m)^n = \mathbf{T}^{mn}$. The first relation implies that the inverse of $\mathbf{T}^m$ is $\mathbf{T}^{-m}$. One can further generalize the exponent to include fractions and ultimately all real numbers; but we need to wait until Chap. 6, in which we discuss the spectral decomposition theorem.

Example 4.1.5 Let us evaluate $\mathbf{T}_\theta^{-n}$ for the operator of the previous example. First, let us find $\mathbf{T}_\theta^{-1}$ (see Fig. 4.1). We are looking for an operator such that $\mathbf{T}_\theta^{-1}\mathbf{T}_\theta(x, y) = (x, y)$, or

$$\mathbf{T}_\theta^{-1}(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta) = (x, y). \tag{4.3}$$

We define $x' = x\cos\theta - y\sin\theta$ and $y' = x\sin\theta + y\cos\theta$ and solve x and y in terms of x' and y' to obtain $x = x'\cos\theta + y'\sin\theta$ and $y = -x'\sin\theta + y'\cos\theta$. Substituting for x and y in Eq. (4.3) yields

$$\mathbf{T}_\theta^{-1}\big(x', y'\big) = \big(x'\cos\theta + y'\sin\theta, -x'\sin\theta + y'\cos\theta\big).$$

[1] These rules apply to any associative algebra, not just to End($\mathcal{V}$).

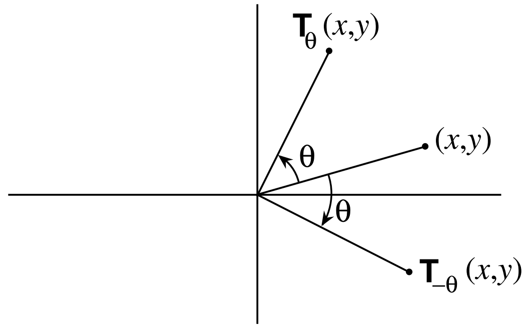

Fig. 4.1 The operator $\mathbf{T}_\theta$ and its inverse as they act on a point in the plane

Comparing this with the action of $\mathbf{T}_\theta$ in the previous example, we discover that the only difference between the two operators is the sign of the $\sin\theta$ term. We conclude that $\mathbf{T}_\theta^{-1}$ has the same effect as $\mathbf{T}_{-\theta}$. So we have

$$\mathbf{T}_\theta^{-1} = \mathbf{T}_{-\theta} \quad \text{and} \quad \mathbf{T}_\theta^{-n} = \big(\mathbf{T}_\theta^{-1}\big)^n = (\mathbf{T}_{-\theta})^n = \mathbf{T}_{-n\theta}.$$

It is instructive to verify that $\mathbf{T}_\theta^{-n}\mathbf{T}_\theta^n = \mathbf{1}$:

$$\begin{aligned}\mathbf{T}_\theta^{-n}\mathbf{T}_\theta^n(x,y) &= \mathbf{T}_\theta^{-n}\big(\overbrace{x\cos n\theta - y\sin n\theta}^{x'}, \overbrace{x\sin n\theta + y\cos n\theta}^{y'}\big) \\ &= \big(x'\cos n\theta + y'\sin n\theta, -x'\sin n\theta + y'\cos n\theta\big) \\ &= \big((x\cos n\theta - y\sin n\theta)\cos n\theta + (x\sin n\theta + y\cos n\theta)\sin n\theta, \\ &\quad -(x\cos n\theta - y\sin n\theta)\sin n\theta + (x\sin n\theta + y\cos n\theta)\cos n\theta\big) \\ &= \big(x\big(\cos^2 n\theta + \sin^2 n\theta\big), y\big(\sin^2 n\theta + \cos^2 n\theta\big)\big) = (x,y).\end{aligned}$$

Similarly, we can show that $\mathbf{T}_\theta^n\mathbf{T}_\theta^{-n}(x,y) = (x,y)$.

One has to keep in mind that $p(\mathbf{T})$ is not, in general, invertible, even if $\mathbf{T}$ is. In fact, the sum of two invertible operators is not necessarily invertible. For example, although $\mathbf{T}$ and $-\mathbf{T}$ are invertible, their sum, the zero operator, is not.

4.1.2 Functions of Operators

We can go one step beyond polynomials of operators and, via Taylor expansion, define functions of them. Consider an ordinary function $f(x)$, which has the Taylor expansion

$$f(x) = \sum_{k=0}^{\infty} \frac{(x-x_0)^k}{k!} \frac{d^k f}{dx^k}\bigg|_{x=x_0}$$

in which x_0 is a point where $f(x)$ and all its derivatives are defined. To this function, there corresponds a function of the operator $\mathbf{T}$, defined as

$$f(\mathbf{T}) = \sum_{k=0}^{\infty} \left.\frac{d^k f}{dx^k}\right|_{x=x_0} \frac{(\mathbf{T} - x_0\mathbf{1})^k}{k!}. \tag{4.4}$$

Because this series is an infinite sum of operators, difficulties may arise concerning its convergence. However, as will be shown in Chap. 6, $f(\mathbf{T})$ is always defined for finite-dimensional vector spaces. In fact, it is always a polynomial in $\mathbf{T}$ (see also Problem 4.1). For the time being, we shall think of $f(\mathbf{T})$ as a formal infinite series. A simplification results when the function can be expanded about $x = 0$. In this case we obtain

$$f(\mathbf{T}) = \sum_{k=0}^{\infty} \left.\frac{d^k f}{dx^k}\right|_{x=0} \frac{\mathbf{T}^k}{k!}. \tag{4.5}$$

A widely used function is the exponential, whose expansion is easily found to be

$$e^{\mathbf{T}} \equiv \exp(\mathbf{T}) = \sum_{k=0}^{\infty} \frac{\mathbf{T}^k}{k!}. \tag{4.6}$$

Example 4.1.6 Let us evaluate $\exp(\alpha\mathbf{T})$ when $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^2$ is given by

$$\mathbf{T}(x, y) = (-y, x).$$

We can find a general formula for the action of $\mathbf{T}^n$ on (x, y). Start with $n = 2$:

$$\mathbf{T}^2(x, y) = \mathbf{T}(-y, x) = (-x, -y) = -(x, y) = -\mathbf{1}(x, y).$$

Thus, $\mathbf{T}^2 = -\mathbf{1}$. From $\mathbf{T}$ and $\mathbf{T}^2$ we can easily obtain higher powers of $\mathbf{T}$. For example: $\mathbf{T}^3 = \mathbf{T}(\mathbf{T}^2) = -\mathbf{T}$, $\mathbf{T}^4 = \mathbf{T}^2\mathbf{T}^2 = \mathbf{1}$, and in general,

$$\begin{aligned} \mathbf{T}^{2n} &= (-1)^n \mathbf{1} \quad \text{for } n = 0, 1, 2, \ldots \\ \mathbf{T}^{2n+1} &= (-1)^n \mathbf{T} \quad \text{for } n = 0, 1, 2, \ldots \end{aligned}$$

Thus,

$$\begin{aligned} \exp(\alpha\mathbf{T}) &= \sum_{n \text{ odd}} \frac{(\alpha\mathbf{T})^n}{n!} + \sum_{n \text{ even}} \frac{(\alpha\mathbf{T})^n}{n!} = \sum_{k=0}^{\infty} \frac{(\alpha\mathbf{T})^{2k+1}}{(2k+1)!} + \sum_{k=0}^{\infty} \frac{(\alpha\mathbf{T})^{2k}}{(2k)!} \\ &= \sum_{k=0}^{\infty} \frac{\alpha^{2k+1}\mathbf{T}^{2k+1}}{(2k+1)!} + \sum_{k=0}^{\infty} \frac{\alpha^{2k}\mathbf{T}^{2k}}{(2k)!} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k\alpha^{2k+1}}{(2k+1)!}\mathbf{T} + \sum_{k=0}^{\infty} \frac{(-1)^k\alpha^{2k}}{(2k)!}\mathbf{1} \end{aligned}$$

$$= \mathbf{T}\sum_{k=0}^{\infty}\frac{(-1)^k\alpha^{2k+1}}{(2k+1)!} + \mathbf{1}\sum_{k=0}^{\infty}\frac{(-1)^k\alpha^{2k}}{(2k)!}.$$

The two series are recognized as $\sin\alpha$ and $\cos\alpha$, respectively. Therefore, we get

$$e^{\alpha\mathbf{T}} = \mathbf{T}\sin\alpha + \mathbf{1}\cos\alpha,$$

which shows that $e^{\alpha\mathbf{T}}$ is a polynomial (of first degree) in $\mathbf{T}$.

The action of $e^{\alpha\mathbf{T}}$ on (x, y) is given by

$$\begin{aligned} e^{\alpha\mathbf{T}}(x, y) &= (\sin\alpha\mathbf{T} + \cos\alpha\mathbf{1})(x, y) = \sin\alpha\mathbf{T}(x, y) + \cos\alpha\mathbf{1}(x, y) \\ &= (\sin\alpha)(-y, x) + (\cos\alpha)(x, y) \\ &= (-y\sin\alpha, x\sin\alpha) + (x\cos\alpha, y\cos\alpha) \\ &= (x\cos\alpha - y\sin\alpha, x\sin\alpha + y\cos\alpha). \end{aligned}$$

generator of the rotation

The reader will recognize the final expression as a rotation in the xy-plane through an angle α. Thus, we can think of $e^{\alpha\mathbf{T}}$ as a rotation operator of angle α about the z-axis. In this context $\mathbf{T}$ is called the *generator* of the rotation.

4.1.3 Commutators

The result of multiplication of two operators depends on the order in which the operators appear. This means that if $\mathbf{T}, \mathbf{U} \in \mathcal{L}(\mathcal{V})$, then $\mathbf{TU} \in \mathcal{L}(\mathcal{V})$ and $\mathbf{UT} \in \mathcal{L}(\mathcal{V})$; however, in general $\mathbf{UT} \neq \mathbf{TU}$. When this is the case, we say that $\mathbf{U}$ and $\mathbf{T}$ *do not commute*. The extent to which two operators fail to commute is given in the following definition.

commutator defined

Definition 4.1.7 The **commutator** $[\mathbf{U}, \mathbf{T}]$ of the two operators $\mathbf{U}$ and $\mathbf{T}$ in $\mathcal{L}(\mathcal{V})$ is another operator in $\mathcal{L}(\mathcal{V})$, defined as

$$[\mathbf{U}, \mathbf{T}] \equiv \mathbf{UT} - \mathbf{TU}.$$

An immediate consequence of this definition is the following:

Proposition 4.1.8 *For* $\mathbf{S}, \mathbf{T}, \mathbf{U} \in \mathcal{L}(\mathcal{V})$ *and* $\alpha, \beta \in \mathbb{C}$ (*or* $\mathbb{R}$), *we have*

$$\begin{aligned} &[\mathbf{U}, \mathbf{T}] = -[\mathbf{T}, \mathbf{U}], && \textit{antisymmetry} \\ &[\alpha\mathbf{U}, \beta\mathbf{T}] = \alpha\beta[\mathbf{U}, \mathbf{T}], && \textit{linearity} \\ &[\mathbf{S}, \mathbf{T} + \mathbf{U}] = [\mathbf{S}, \mathbf{T}] + [\mathbf{S}, \mathbf{U}], && \textit{linearity in the right entry} \\ &[\mathbf{S} + \mathbf{T}, \mathbf{U}] = [\mathbf{S}, \mathbf{U}] + [\mathbf{T}, \mathbf{U}], && \textit{linearity in the left entry} \\ &[\mathbf{ST}, \mathbf{U}] = \mathbf{S}[\mathbf{T}, \mathbf{U}] + [\mathbf{S}, \mathbf{U}]\mathbf{T}, && \textit{right derivation property} \\ &[\mathbf{S}, \mathbf{TU}] = [\mathbf{S}, \mathbf{T}]\mathbf{U} + \mathbf{T}[\mathbf{S}, \mathbf{U}], && \textit{left derivation property} \\ &\big[[\mathbf{S}, \mathbf{T}], \mathbf{U}\big] + \big[[\mathbf{U}, \mathbf{S}], \mathbf{T}\big] + \big[[\mathbf{T}, \mathbf{U}], \mathbf{S}\big] = 0, && \textit{Jacobi identity} \end{aligned}$$

Proof In almost all cases the proof follows immediately from the definition. The only minor exceptions are the derivation properties. We prove the left derivation property:

$$[\mathbf{S},\mathbf{TU}] = \mathbf{S}(\mathbf{TU}) - (\mathbf{TU})\mathbf{S} = \mathbf{STU} - \mathbf{TUS} + \underbrace{\mathbf{TSU} - \mathbf{TSU}}_{\equiv 0}$$

$$= (\mathbf{ST} - \mathbf{TS})\mathbf{U} + \mathbf{T}(\mathbf{SU} - \mathbf{US}) = [\mathbf{S},\mathbf{T}]\mathbf{U} + \mathbf{T}[\mathbf{S},\mathbf{U}].$$

The right derivation property is proved in exactly the same way. □

A useful consequence of the definition and Proposition 4.1.8 is

$$\left[\mathbf{A},\mathbf{A}^m\right] = \mathbf{0} \quad \text{for } m = 0, \pm 1, \pm 2, \ldots.$$

In particular, $[\mathbf{A},\mathbf{1}] = \mathbf{0}$ and $[\mathbf{A},\mathbf{A}^{-1}] = \mathbf{0}$.

An example of the commutators of operators is that of $\mathbf{D}$ and $\mathbf{T}$ defined in Example 2.3.5. The reader is urged to verify that

$$[\mathbf{D},\mathbf{T}] = \mathbf{1} \tag{4.7}$$

4.2 Derivatives of Operators

Up to this point we have been discussing the algebraic properties of operators, static objects that obey certain algebraic rules and fulfill the static needs of some applications. However, physical quantities are dynamic, and if we want operators to represent physical quantities, we must allow them to change with time. This dynamism is best illustrated in quantum mechanics, where physical observables are represented by operators.

Let us consider a mapping $\mathbf{H}: \mathbb{R} \to \operatorname{End}(\mathcal{V})$, which[2] takes in a real number and gives out a linear operator on the vector space $\mathcal{V}$. We denote the image of $t \in \mathbb{R}$ by $\mathbf{H}(t)$, which acts on the underlying vector space $\mathcal{V}$. The physical meaning of this is that as t (usually time) varies, its image $\mathbf{H}(t)$ also varies. Therefore, for different values of t, we have *different* operators. In particular, $[\mathbf{H}(t),\mathbf{H}(t')] \neq \mathbf{0}$ for $t \neq t'$. A concrete example is an operator that is a linear combination of the operators $\mathbf{D}$ and $\mathbf{T}$ introduced in Example 2.3.5, with time-dependent scalars. To be specific, let $\mathbf{H}(t) = \mathbf{D}\cos\omega t + \mathbf{T}\sin\omega t$, where ω is a constant. As time passes, $\mathbf{H}(t)$ changes its identity from $\mathbf{D}$ to $\mathbf{T}$ and back to $\mathbf{D}$. Most of the time it has a hybrid identity! Since $\mathbf{D}$ and $\mathbf{T}$ do not commute [see Eq. (4.7)], values of $\mathbf{H}(t)$ for different times do not necessarily commute.

a time-dependent operator does not commute with itself at different times

Of particular interest are operators that can be written as $\exp(\mathbf{H}(t))$, where $\mathbf{H}(t)$ is a "simple" operator; i.e., the dependence of $\mathbf{H}(t)$ on t is sim-

[2]Strictly speaking, the domain of $\mathbf{H}$ must be an interval $[a,b]$ of the real line, because $\mathbf{H}$ may not be defined for all $\mathbb{R}$. However, for our purposes, such a fine distinction is not necessary.

pler than the corresponding dependence of $\exp(\mathbf{H}(t))$. We have already encountered such a situation in Example 4.1.6, where it was shown that the operation of rotation around the z-axis could be written as $\exp(\alpha\mathbf{T})$, and the action of $\mathbf{T}$ on (x, y) was a great deal simpler than the corresponding action of $\exp(\alpha\mathbf{T})$.

Euler's theorem and Euler angles

Such a state of affairs is very common in physics. In fact, it can be shown that many operators of physical interest can be written as a product of simpler operators, each being of the form $\exp(\alpha\mathbf{T})$. For example, we know from Euler's theorem in mechanics that an arbitrary rotation in three dimensions can be written as a product of three simpler rotations, each being a rotation through a so-called *Euler angle* about an axis.

derivative of an operator

Definition 4.2.1 For the mapping $\mathbf{H} : \mathbb{R} \to \mathrm{End}(\mathcal{V})$, we define the **derivative** as

$$\frac{d\mathbf{H}}{dt} = \lim_{\Delta t \to 0} \frac{\mathbf{H}(t + \Delta t) - \mathbf{H}(t)}{\Delta t}.$$

This derivative also belongs to $\mathrm{End}(\mathcal{V})$.

As long as we keep track of the order, practically all the rules of differentiation apply to operators. For example,

$$\frac{d}{dt}(\mathbf{UT}) = \frac{d\mathbf{U}}{dt}\mathbf{T} + \mathbf{U}\frac{d\mathbf{T}}{dt}.$$

We are not allowed to change the order of multiplication on the RHS, not even when both operators being multiplied are the same on the LHS. For instance, if we let $\mathbf{U} = \mathbf{T} = \mathbf{H}$ in the preceding equation, we obtain

$$\frac{d}{dt}\left(\mathbf{H}^2\right) = \frac{d\mathbf{H}}{dt}\mathbf{H} + \mathbf{H}\frac{d\mathbf{H}}{dt}.$$

This is *not*, in general, equal to $2\mathbf{H}\frac{d\mathbf{H}}{dt}$.

Example 4.2.2 Let us find the derivative of $\exp(t\mathbf{H})$, where $\mathbf{H}$ *is independent of* t. Using Definition 4.2.1, we have

$$\frac{d}{dt}\exp(t\mathbf{H}) = \lim_{\Delta t \to 0} \frac{\exp[(t + \Delta t)\mathbf{H}] - \exp(t\mathbf{H})}{\Delta t}.$$

However, for infinitesimal Δt we have

$$\begin{aligned}\exp\bigl[(t + \Delta t)\mathbf{H}\bigr] - \exp(t\mathbf{H}) &= e^{t\mathbf{H}}e^{\Delta t\mathbf{H}} - e^{t\mathbf{H}} \\ &= e^{t\mathbf{H}}(\mathbf{1} + \mathbf{H}\Delta t) - e^{t\mathbf{H}} = e^{t\mathbf{H}}\mathbf{H}\Delta t.\end{aligned}$$

Therefore,

$$\frac{d}{dt}\exp(t\mathbf{H}) = \lim_{\Delta t \to 0} \frac{e^{t\mathbf{H}}\mathbf{H}\Delta t}{\Delta t} = e^{t\mathbf{H}}\mathbf{H}.$$

Since $\mathbf{H}$ and $e^{t\mathbf{H}}$ commute,[3] we also have

$$\frac{d}{dt}\exp(t\mathbf{H}) = \mathbf{H}e^{t\mathbf{H}}.$$

Note that in deriving the equation for the derivative of $e^{t\mathbf{H}}$, we have used the relation $e^{t\mathbf{H}}e^{\Delta t\mathbf{H}} = e^{(t+\Delta t)\mathbf{H}}$. This may seem trivial, but it will be shown later that in general, $e^{\mathbf{S}+\mathbf{T}} \neq e^{\mathbf{S}}e^{\mathbf{T}}$.

Now let us evaluate the derivative of a more general time-dependent operator, $\exp[\mathbf{H}(t)]$:

$$\frac{d}{dt}\exp\bigl[\mathbf{H}(t)\bigr] = \lim_{\Delta t\to 0}\frac{\exp[\mathbf{H}(t+\Delta t)] - \exp[\mathbf{H}(t)]}{\Delta t}.$$

If $\mathbf{H}(t)$ possesses a derivative, we have, to the first order in Δt,

$$\mathbf{H}(t+\Delta t) = \mathbf{H}(t) + \Delta t\frac{d\mathbf{H}}{dt},$$

and we can write $\exp[\mathbf{H}(t+\Delta t)] = \exp[\mathbf{H}(t) + \Delta t\, d\mathbf{H}/dt]$. It is very tempting to factor out the $\exp[\mathbf{H}(t)]$ and expand the remaining part. However, as we will see presently, this is not possible in general. As preparation, consider the following example, which concerns the integration of an operator.

Example 4.2.3 The Schrödinger equation

$$i\frac{\partial}{\partial t}\bigl|\psi(t)\bigr\rangle = \mathbf{H}\bigl|\psi(t)\bigr\rangle$$

can be turned into an *operator differential equation* as follows. Define the so-called **evolution operator** $\mathbf{U}(t)$ by $|\psi(t)\rangle = \mathbf{U}(t)|\psi(0)\rangle$, and substitute in the Schrödinger equation to obtain

evolution operator

$$i\frac{\partial}{\partial t}\mathbf{U}(t)\bigl|\psi(0)\bigr\rangle = \mathbf{H}\mathbf{U}(t)\bigl|\psi(0)\bigr\rangle.$$

Ignoring the arbitrary vector $|\psi(0)\rangle$ results in $d\mathbf{U}/dt = -i\mathbf{H}\mathbf{U}(t)$, a differential equation in $\mathbf{U}(t)$, where $\mathbf{H}$ is not dependent on t. We can find a solution to such an equation by repeated differentiation followed by Taylor series expansion. Thus,

$$\begin{aligned}
\frac{d^2\mathbf{U}}{dt^2} &= -i\mathbf{H}\frac{d\mathbf{U}}{dt} = -i\mathbf{H}\bigl[-i\mathbf{H}\mathbf{U}(t)\bigr] = (-i\mathbf{H})^2\mathbf{U}(t),\\
\frac{d^3\mathbf{U}}{dt^3} &= \frac{d}{dt}\bigl[(-i\mathbf{H})^2\mathbf{U}(t)\bigr] = (-i\mathbf{H})^2\frac{d\mathbf{U}}{dt} = (-i\mathbf{H})^3\mathbf{U}(t).
\end{aligned}$$

[3] This is a consequence of a more general result that if two operators commute, any pair of functions of those operators also commute (see Problem 4.14).

In general $d^n\mathbf{U}/dt^n = (-i\mathbf{H})^n\mathbf{U}(t)$. Assuming that $\mathbf{U}(t)$ is well-defined at $t = 0$, the above relations say that all derivatives of $\mathbf{U}(t)$ are also well-defined at $t = 0$. Therefore, we can expand $\mathbf{U}(t)$ around $t = 0$ to obtain

$$\begin{aligned}\mathbf{U}(t) &= \sum_{n=0}^{\infty}\frac{t^n}{n!}\left(\frac{d^n\mathbf{U}}{dt^n}\right)_{t=0} = \sum_{n=0}^{\infty}\frac{t^n}{n!}(-i\mathbf{H})^n\mathbf{U}(0)\\ &= \left(\sum_{n=0}^{\infty}\frac{(-it\mathbf{H})^n}{n!}\right)\mathbf{U}(0) = e^{-it\mathbf{H}}\mathbf{U}(0).\end{aligned}$$

But $\mathbf{U}(0) = \mathbf{1}$ by the definition of $\mathbf{U}(t)$. Hence, $\mathbf{U}(t) = e^{-it\mathbf{H}}$, and

$$\left|\psi(t)\right\rangle = e^{-it\mathbf{H}}\left|\psi(0)\right\rangle.$$

Let us see under what conditions we have $\exp(\mathbf{S}+\mathbf{T}) = \exp(\mathbf{S})\exp(\mathbf{T})$. We consider only the case where the commutator of the two operators commutes with both of them:

$$\big[\mathbf{T}, [\mathbf{S}, \mathbf{T}]\big] = \mathbf{0} = \big[\mathbf{S}, [\mathbf{S}, \mathbf{T}]\big].$$

Now consider the operator $\mathbf{U}(t) = e^{t\mathbf{S}}e^{t\mathbf{T}}e^{-t(\mathbf{S}+\mathbf{T})}$ and differentiate it using the result of Example 4.2.2 and the product rule for differentiation:

$$\begin{aligned}\frac{d\mathbf{U}}{dt} &= \mathbf{S}e^{t\mathbf{S}}e^{t\mathbf{T}}e^{-t(\mathbf{S}+\mathbf{T})} + e^{t\mathbf{S}}\mathbf{T}e^{t\mathbf{T}}e^{-t(\mathbf{S}+\mathbf{T})} - e^{t\mathbf{S}}e^{t\mathbf{T}}(\mathbf{S}+\mathbf{T})e^{-t(\mathbf{S}+\mathbf{T})}\\ &= \mathbf{S}e^{t\mathbf{S}}e^{t\mathbf{T}}e^{-t(\mathbf{S}+\mathbf{T})} - e^{t\mathbf{S}}e^{t\mathbf{T}}\mathbf{S}e^{-t(\mathbf{S}+\mathbf{T})}.\end{aligned} \tag{4.8}$$

The three factors of $\mathbf{U}(t)$ are present in all terms; however, they are not always next to one another. We can switch the operators if we introduce a commutator. For instance, $e^{t\mathbf{T}}\mathbf{S} = \mathbf{S}e^{t\mathbf{T}} + [e^{t\mathbf{T}}, \mathbf{S}]$.

It is left as a problem for the reader to show that if $[\mathbf{S}, \mathbf{T}]$ commutes with $\mathbf{S}$ and $\mathbf{T}$, then $[e^{t\mathbf{T}}, \mathbf{S}] = -t[\mathbf{S}, \mathbf{T}]e^{t\mathbf{T}}$, and therefore, $e^{t\mathbf{T}}\mathbf{S} = \mathbf{S}e^{t\mathbf{T}} - t[\mathbf{S}, \mathbf{T}]e^{t\mathbf{T}}$. Substituting this in Eq. (4.8) and noting that $e^{t\mathbf{S}}\mathbf{S} = \mathbf{S}e^{t\mathbf{S}}$ yields $d\mathbf{U}/dt = t[\mathbf{S}, \mathbf{T}]\mathbf{U}(t)$. The solution to this equation is

$$\mathbf{U}(t) = \exp\left(\frac{t^2}{2}[\mathbf{S}, \mathbf{T}]\right) \quad \Rightarrow \quad e^{t\mathbf{S}}e^{t\mathbf{T}}e^{-t(\mathbf{S}+\mathbf{T})} = \exp\left(\frac{t^2}{2}[\mathbf{S}, \mathbf{T}]\right)$$

because $\mathbf{U}(0) = \mathbf{1}$. We thus have the following:

Baker-Campbell-Hausdorff formula

Proposition 4.2.4 *Let* $\mathbf{S}, \mathbf{T} \in \mathcal{L}(\mathcal{V})$. *If* $[\mathbf{S}, [\mathbf{S}, \mathbf{T}]] = \mathbf{0} = [\mathbf{T}, [\mathbf{S}, \mathbf{T}]]$, *then the* ***Baker–Campbell–Hausdorff formula*** *holds*:

$$e^{t\mathbf{S}}e^{t\mathbf{T}}e^{-(t^2/2)[\mathbf{S},\mathbf{T}]} = e^{t(\mathbf{S}+\mathbf{T})}. \tag{4.9}$$

In particular, $e^{t\mathbf{S}}e^{t\mathbf{T}} = e^{t(\mathbf{S}+\mathbf{T})}$ *if and only if* $[\mathbf{S}, \mathbf{T}] = \mathbf{0}$.

If $t = 1$, Eq. (4.9) reduces to

$$e^{\mathbf{S}}e^{\mathbf{T}}e^{-(1/2)[\mathbf{S},\mathbf{T}]} = e^{\mathbf{S}+\mathbf{T}}. \tag{4.10}$$

Now assume that both $\mathbf{H}(t)$ and its derivative commute with $[\mathbf{H}, d\mathbf{H}/dt]$. Letting $\mathbf{S} = \mathbf{H}(t)$ and $\mathbf{T} = \Delta t d\mathbf{H}/dt$ in (4.10), we obtain

$$\begin{aligned} e^{\mathbf{H}(t+\Delta t)} &= e^{\mathbf{H}(t)+\Delta t d\mathbf{H}/dt} \\ &= e^{\mathbf{H}(t)} e^{\Delta t (d\mathbf{H}/dt)} e^{-[\mathbf{H}(t), \Delta t d\mathbf{H}/dt]/2}. \end{aligned}$$

For infinitesimal Δt, this yields

$$\begin{aligned} e^{\mathbf{H}(t+\Delta t)} &= e^{\mathbf{H}(t)} \left(\mathbf{1} + \Delta t \frac{d\mathbf{H}}{dt}\right)\left(\mathbf{1} - \frac{1}{2}\Delta t \left[\mathbf{H}(t), \frac{d\mathbf{H}}{dt}\right]\right) \\ &= e^{\mathbf{H}(t)} \left\{\mathbf{1} + \Delta t \frac{d\mathbf{H}}{dt} - \frac{1}{2}\Delta t \left[\mathbf{H}(t), \frac{d\mathbf{H}}{dt}\right]\right\}, \end{aligned}$$

and we have

$$\frac{d}{dt} e^{\mathbf{H}(t)} = e^{\mathbf{H}} \frac{d\mathbf{H}}{dt} - \frac{1}{2} e^{\mathbf{H}} \left[\mathbf{H}, \frac{d\mathbf{H}}{dt}\right].$$

We can also write

$$\begin{aligned} e^{\mathbf{H}(t+\Delta t)} &= e^{[\mathbf{H}(t)+\Delta t d\mathbf{H}/dt]} = e^{[\Delta t d\mathbf{H}/dt + \mathbf{H}(t)]} \\ &= e^{[\Delta t d\mathbf{H}/dt]} e^{\mathbf{H}(t)} e^{-[\Delta t d\mathbf{H}/dt, \mathbf{H}(t)]/2}, \end{aligned}$$

which yields

$$\frac{d}{dt} e^{\mathbf{H}(t)} = \frac{d\mathbf{H}}{dt} e^{\mathbf{H}} + \frac{1}{2} e^{\mathbf{H}} \left[\mathbf{H}, \frac{d\mathbf{H}}{dt}\right].$$

Adding the above two expressions and dividing by 2 yields the following symmetric expression for the derivative:

$$\frac{d}{dt} e^{\mathbf{H}(t)} = \frac{1}{2}\left(\frac{d\mathbf{H}}{dt} e^{\mathbf{H}} + e^{\mathbf{H}} \frac{d\mathbf{H}}{dt}\right) = \frac{1}{2}\left\{\frac{d\mathbf{H}}{dt}, e^{\mathbf{H}}\right\},$$

where $\{\mathbf{S}, \mathbf{T}\} \equiv \mathbf{ST} + \mathbf{TS}$ is called the **anticommutator** of the operators $\mathbf{S}$ and $\mathbf{T}$. We, therefore, have the following proposition.

anticommutator

Proposition 4.2.5 *Let* $\mathbf{H} : \mathbb{R} \to \mathcal{L}(\mathcal{V})$ *and assume that* $\mathbf{H}$ *and its derivative commute with* $[\mathbf{H}, d\mathbf{H}/dt]$. *Then*

$$\frac{d}{dt} e^{\mathbf{H}(t)} = \frac{1}{2}\left\{\frac{d\mathbf{H}}{dt}, e^{\mathbf{H}}\right\}.$$

In particular, if $[\mathbf{H}, d\mathbf{H}/dt] = 0$, *then*

$$\frac{d}{dt} e^{\mathbf{H}(t)} = \frac{d\mathbf{H}}{dt} e^{\mathbf{H}} = e^{\mathbf{H}} \frac{d\mathbf{H}}{dt}.$$

A frequently encountered operator is $\mathbf{F}(t) = e^{t\mathbf{A}} \mathbf{B} e^{-t\mathbf{A}}$, where $\mathbf{A}$ and $\mathbf{B}$ are t-independent. It is straightforward to show that

$$\frac{d\mathbf{F}}{dt} = [\mathbf{A}, \mathbf{F}(t)] \quad \text{and} \quad \frac{d}{dt}[\mathbf{A}, \mathbf{F}(t)] = \left[\mathbf{A}, \frac{d\mathbf{F}}{dt}\right].$$

Using these results, we can write

$$\frac{d^2\mathbf{F}}{dt^2} = \frac{d}{dt}\big[\mathbf{A}, \mathbf{F}(t)\big] = \big[\mathbf{A}, \big[\mathbf{A}, \mathbf{F}(t)\big]\big] \equiv \mathbf{A}^2\big[\mathbf{F}(t)\big],$$

and in general, $d^n\mathbf{F}/dt^n = \mathbf{A}^n[\mathbf{F}(t)]$, where $\mathbf{A}^n[\mathbf{F}(t)]$ is defined inductively as $\mathbf{A}^n[\mathbf{F}(t)] = [\mathbf{A}, \mathbf{A}^{n-1}[\mathbf{F}(t)]]$, with $\mathbf{A}^0[\mathbf{F}(t)] \equiv \mathbf{F}(t)$. For example,

$$\mathbf{A}^3\big[\mathbf{F}(t)\big] = \big[\mathbf{A}, \mathbf{A}^2\big[\mathbf{F}(t)\big]\big] = \big[\mathbf{A}, \big[\mathbf{A}, \mathbf{A}\big[\mathbf{F}(t)\big]\big]\big] = \big[\mathbf{A}, \big[\mathbf{A}, \big[\mathbf{A}, \mathbf{F}(t)\big]\big]\big].$$

Evaluating $\mathbf{F}(t)$ and all its derivatives at $t = 0$ and substituting in the Taylor expansion about $t = 0$, we get

$$\mathbf{F}(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \frac{d^n\mathbf{F}}{dt^n}\bigg|_{t=0} = \sum_{n=0}^{\infty} \frac{t^n}{n!}\mathbf{A}^n\big[\mathbf{F}(0)\big] = \sum_{n=0}^{\infty} \frac{t^n}{n!}\mathbf{A}^n[\mathbf{B}].$$

That is,

$$e^{t\mathbf{A}}\mathbf{B}e^{-t\mathbf{A}} = \sum_{n=0}^{\infty} \frac{t^n}{n!}\mathbf{A}^n[\mathbf{B}] \equiv \mathbf{B} + t[\mathbf{A}, \mathbf{B}] + \frac{t^2}{2!}\big[\mathbf{A}, [\mathbf{A}, \mathbf{B}]\big] + \cdots.$$

Sometimes this is written symbolically as

$$e^{t\mathbf{A}}\mathbf{B}e^{-t\mathbf{A}} = \left(\sum_{n=0}^{\infty} \frac{t^n}{n!}\mathbf{A}^n\right)[\mathbf{B}] \equiv e^{t\mathbf{A}}[\mathbf{B}],$$

where the RHS is merely an abbreviation of the infinite sum in the middle.

For $t = 1$ we obtain a widely used formula:

$$e^{\mathbf{A}}\mathbf{B}e^{-\mathbf{A}} = e^{\mathbf{A}}[\mathbf{B}] = \left(\sum_{n=0}^{\infty} \frac{1}{n!}\mathbf{A}^n\right)[\mathbf{B}] \equiv \mathbf{B} + [\mathbf{A}, \mathbf{B}] + \frac{1}{2!}\big[\mathbf{A}, [\mathbf{A}, \mathbf{B}]\big] + \cdots.$$

If $\mathbf{A}$ commutes with $[\mathbf{A}, \mathbf{B}]$, then the infinite series truncates at the second term, and we have

$$e^{t\mathbf{A}}\mathbf{B}e^{-t\mathbf{A}} = \mathbf{B} + t[\mathbf{A}, \mathbf{B}].$$

For instance, if $\mathbf{A}$ and $\mathbf{B}$ are replaced by $\mathbf{D}$ and $\mathbf{T}$ of Example 2.3.5 and Eq. (4.7), we get

$$e^{t\mathbf{D}}\mathbf{T}e^{-t\mathbf{D}} = \mathbf{T} + t[\mathbf{D}, \mathbf{T}] = \mathbf{T} + t\mathbf{1}.$$

generator of translation

The RHS shows that the operator $\mathbf{T}$ has been *translated* by an amount t (more precisely, by t times the unit operator). We therefore call $\exp(t\mathbf{D})$ the translation operator of $\mathbf{T}$ by t, and we call $\mathbf{D}$ the *generator* of translation. With a little modification $\mathbf{T}$ and $\mathbf{D}$ become, respectively, the position and momentum operators in quantum mechanics. Thus,

momentum as generator of translation

Box 4.2.6 *Momentum is the generator of translation in quantum mechanics.*

But more of this later!

4.3 Conjugation of Operators

We have discussed the notion of the dual of a vector in conjunction with inner products. We now incorporate linear operators into this notion. Let $|b\rangle, |c\rangle \in \mathcal{V}$ and assume that $|c\rangle = \mathbf{T}|b\rangle$. We know that there are linear functionals in the dual space $\mathcal{V}^*$ that are associated with $(|b\rangle)^\dagger = \langle b|$ and $(|c\rangle)^\dagger = \langle c|$. Is there a linear operator belonging to $\mathcal{L}(\mathcal{V}^*)$ that somehow corresponds to $\mathbf{T}$? In other words, can we find a linear operator that relates $\langle b|$ and $\langle c|$ just as $\mathbf{T}$ relates $|b\rangle$ and $|c\rangle$? The answer comes in the following definition.

adjoint of an operator

Definition 4.3.1 Let $\mathbf{T} \in \mathcal{L}(\mathcal{V})$ and $|a\rangle, |b\rangle \in \mathcal{V}$. The **adjoint**, or **hermitian conjugate**, of $\mathbf{T}$ is denoted by $\mathbf{T}^\dagger$ and defined by[4]

$$\langle a|\mathbf{T}|b\rangle^* = \langle b|\mathbf{T}^\dagger|a\rangle. \tag{4.11}$$

The LHS of Eq. (4.11) can be written as $\langle a \mid c\rangle^*$ or $\langle c \mid a\rangle$, in which case we can identify

$$\langle c| = \langle b|\mathbf{T}^\dagger \quad \Rightarrow \quad \bigl(\mathbf{T}|b\rangle\bigr)^\dagger = \langle b|\mathbf{T}^\dagger. \tag{4.12}$$

This equation is sometimes used as the definition of the hermitian conjugate. From Eq. (4.11), the reader may easily verify that $\mathbf{1}^\dagger = \mathbf{1}$. Thus, using the unit operator for $\mathbf{T}$, (4.12) justifies Eq. (2.26).

Some of the properties of conjugation are listed in the following theorem, whose proof is left as an exercise.

Theorem 4.3.2 *Let* $\mathbf{U}, \mathbf{T} \in \mathrm{End}(\mathcal{V})$ *and* $\alpha \in \mathbb{C}$. *Then*

1. $(\mathbf{U}+\mathbf{T})^\dagger = \mathbf{U}^\dagger + \mathbf{T}^\dagger$.
2. $(\mathbf{U}\mathbf{T})^\dagger = \mathbf{T}^\dagger\mathbf{U}^\dagger$.
3. $(\alpha\mathbf{T})^\dagger = \alpha^*\mathbf{T}^\dagger$.
4. $\bigl((\mathbf{T})^\dagger\bigr)^\dagger = \mathbf{T}$.

The last identity holds for finite-dimensional vector spaces; it does not apply to infinite-dimensional vector spaces in general.

In previous examples dealing with linear operators $\mathbf{T} : \mathbb{R}^n \to \mathbb{R}^n$, an element of $\mathbb{R}^n$ was denoted by a row vector, such as (x, y) for $\mathbb{R}^2$ and (x, y, z) for $\mathbb{R}^3$. There was no confusion, because we were operating only in $\mathcal{V}$. However, since elements of both $\mathcal{V}$ and $\mathcal{V}^*$ are required when discussing $\mathbf{T}$, $\mathbf{T}^*$, and $\mathbf{T}^\dagger$, it is helpful to make a distinction between them. We therefore resort to the convention introduced in Example 2.2.3 by which

Box 4.3.3 *Kets are represented as column vectors and bras as row vectors.*

[4]With the notion of adjoint already introduced in Definition 2.4.3, we should probably not use the same name for $\mathbf{T}^\dagger$. However, the adjoint as defined in Definition 2.4.3 is rarely used in physics. Furthermore, both the notation and the context will make it clear which adjoint one is talking about. Therefore, there is no risk of confusion.

Example 4.3.4 Let us find the hermitian conjugate of the operator $\mathsf{T} : \mathbb{C}^3 \to \mathbb{C}^3$ given by

$$\mathsf{T}\begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{pmatrix} = \begin{pmatrix}\alpha_1 - i\alpha_2 + \alpha_3\\ i\alpha_1 - \alpha_3\\ \alpha_1 - \alpha_2 + i\alpha_3\end{pmatrix}.$$

Introduce

$$|a\rangle = \begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{pmatrix} \quad \text{and} \quad |b\rangle = \begin{pmatrix}\beta_1\\ \beta_2\\ \beta_3\end{pmatrix}$$

with dual vectors $\langle a| = (\alpha_1^* \; \alpha_2^* \; \alpha_3^*)$ and $\langle b| = (\beta_1^* \; \beta_2^* \; \beta_3^*)$, respectively. We use Eq. (4.11) to find $\mathsf{T}^\dagger$:

$$\begin{aligned}
&\langle b|\mathsf{T}^\dagger|a\rangle \\
&\quad = \langle a|\mathsf{T}|b\rangle^* = \left[\begin{pmatrix}\alpha_1^* & \alpha_2^* & \alpha_3^*\end{pmatrix}\mathsf{T}\begin{pmatrix}\beta_1\\ \beta_2\\ \beta_3\end{pmatrix}\right]^* \\
&\quad = \left[\begin{pmatrix}\alpha_1^* & \alpha_2^* & \alpha_3^*\end{pmatrix}\begin{pmatrix}\beta_1 - i\beta_2 + \beta_3\\ i\beta_1 - \beta_3\\ \beta_1 - \beta_2 + i\beta_3\end{pmatrix}\right]^* \\
&\quad = \left[\alpha_1^*\beta_1 - i\alpha_1^*\beta_2 + \alpha_1^*\beta_3 + i\alpha_2^*\beta_1 - \alpha_2^*\beta_3 + \alpha_3^*\beta_1 - \alpha_3^*\beta_2 + i\alpha_3^*\beta_3\right]^*.
\end{aligned}$$

Taking the complex conjugate of all the numbers inside the square brackets, we find

$$\langle b|\mathsf{T}^\dagger|a\rangle = \underbrace{\begin{pmatrix}\beta_1^* & \beta_2^* & \beta_3^*\end{pmatrix}}_{=\langle b|}\begin{pmatrix}\alpha_1 - i\alpha_2 + \alpha_3\\ i\alpha_1 - \alpha_3\\ \alpha_1 - \alpha_2 - i\alpha_3\end{pmatrix}.$$

Therefore, we obtain

$$\mathsf{T}^\dagger\begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{pmatrix} = \begin{pmatrix}\alpha_1 - i\alpha_2 + \alpha_3\\ i\alpha_1 - \alpha_3\\ \alpha_1 - \alpha_2 - i\alpha_3\end{pmatrix}.$$

4.3.1 Hermitian Operators

The process of conjugation of linear operators looks much like conjugation of complex numbers. Equation (4.11) alludes to this fact, and Theorem 4.3.2 provides further evidence. It is therefore natural to look for operators that are counterparts of real numbers. One can define complex conjugation for operators and thereby construct real operators. However, these real operators will not be interesting because—as it turns out—they completely ignore the complex character of the vector space. The following alternative definition makes use of *hermitian* conjugation, and the result will have much wider application than is allowed by a mere *complex* conjugation.

Definition 4.3.5 A linear operator $\mathbf{H} \in \mathcal{L}(\mathcal{V})$ is called **hermitian**, or **self-adjoint**, if $\mathbf{H}^\dagger = \mathbf{H}$. Similarly, $\mathbf{A} \in \mathcal{L}(\mathcal{V})$ is called **anti-hermitian** if $\mathbf{A}^\dagger = -\mathbf{A}$.

hermitian and anti-hermitian operators

Historical Notes

Charles Hermite (1822–1901), one of the most eminent French mathematicians of the nineteenth century, was particularly distinguished for the clean elegance and high artistic quality of his work. As a student, he courted disaster by neglecting his routine assigned work to study the classic masters of mathematics; and though he nearly failed his examinations, he became a first-rate creative mathematician while still in his early twenties. In 1870 he was appointed to a professorship at the Sorbonne, where he trained a whole generation of well-known French mathematicians, including Picard, Borel, and Poincaré. The character of his mind is suggested by a remark of Poincaré: "Talk with M. Hermite. He never evokes a concrete image, yet you soon perceive that the most abstract entities are to him like living creatures." He disliked geometry, but was strongly attracted to number theory and analysis, and his favorite subject was elliptic functions, where these two fields touch in many remarkable ways. Earlier in the century the Norwegian genius Abel had proved that the general equation of the fifth degree cannot be solved by functions involving only rational operations and root extractions. One of Hermite's most surprising achievements (in 1858) was to show that this equation can be solved by elliptic functions. His 1873 proof of the transcendence of e was another high point of his career.[5] If he had been willing to dig even deeper into this vein, he could probably have disposed of π as well, but apparently he had enough of a good thing. As he wrote to a friend, "I shall risk nothing on an attempt to prove the transcendence of the number π. If others undertake this enterprise, no one will be happier than I at their success, but believe me, my dear friend, it will not fail to cost them some efforts." As it turned out, Lindemann's proof nine years later rested on extending Hermite's method.

Charles Hermite
1822–1901

Several of his purely mathematical discoveries had unexpected applications many years later to mathematical physics. For example, the Hermitian forms and matrices that he invented in connection with certain problems of number theory turned out to be crucial for Heisenberg's 1925 formulation of quantum mechanics, and Hermite polynomials (see Chap. 8) are useful in solving Schrödinger's wave equation.

The following observations strengthen the above conjecture that conjugation of complex numbers and hermitian conjugation of operators are somehow related.

Definition 4.3.6 The **expectation value** $\langle \mathbf{T} \rangle_a$ of an operator $\mathbf{T}$ in the "state" $|a\rangle$ is a complex number defined by $\langle \mathbf{T} \rangle_a = \langle a|\mathbf{T}|a\rangle$.

expectation value

The complex conjugate of the expectation value is[6]

$$\langle \mathbf{T} \rangle^* = \langle a|\mathbf{T}|a\rangle^* = \langle a|\mathbf{T}^\dagger|a\rangle.$$

[5] Transcendental numbers are those that are not roots of polynomials with integer coefficients.

[6] When no risk of confusion exists, it is common to drop the subscript "a" and write $\langle \mathbf{T} \rangle$ for the expectation value of $\mathbf{T}$.

In words, $\mathbf{T}^\dagger$, the *hermitian conjugate* of $\mathbf{T}$, has an expectation value that is the *complex conjugate* of the latter's expectation value. In particular, if $\mathbf{T}$ is hermitian—is equal to its hermitian conjugate—its expectation value will be real.

What is the analogue of the known fact that a complex number is the sum of a real number and a pure imaginary one? The decomposition

$$\mathbf{T} = \frac{1}{2}\big(\mathbf{T} + \mathbf{T}^\dagger\big) + \frac{1}{2}\big(\mathbf{T} - \mathbf{T}^\dagger\big) \equiv \mathbf{X} + \mathbf{A}$$

shows that *any operator can be written as a sum of a hermitian operator* $\mathbf{X} = \frac{1}{2}(\mathbf{T} + \mathbf{T}^\dagger)$ *and an anti-hermitian operator* $\mathbf{A} = \frac{1}{2}(\mathbf{T} - \mathbf{T}^\dagger)$.

We can go even further, because any anti-hermitian operator $\mathbf{A}$ can be written as $\mathbf{A} = i(-i\mathbf{A})$ in which $-i\mathbf{A}$ is hermitian: $(-i\mathbf{A})^\dagger = (-i)^*\mathbf{A}^\dagger = i(-\mathbf{A}) = -i\mathbf{A}$. Denoting $-i\mathbf{A}$ by $\mathbf{Y}$, we write $\mathbf{T} = \mathbf{X} + i\mathbf{Y}$, where both $\mathbf{X}$ and $\mathbf{Y}$ are *hermitian*. This is the analogue of the decomposition $z = x + iy$ in which both x and y are *real*.

Clearly, we should expect some departures from a perfect correspondence. This is due to a lack of commutativity among operators. For instance, although the product of two real numbers is real, the product of two hermitian operators is not, in general, hermitian:

$$(\mathbf{TU})^\dagger = \mathbf{U}^\dagger\mathbf{T}^\dagger = \mathbf{UT} \neq \mathbf{TU}.$$

We have seen the relation between expectation values and conjugation properties of operators. The following theorem completely characterizes hermitian operators in terms of their expectation values:

Theorem 4.3.7 *A linear map* $\mathbf{H}$ *on a complex inner product space is hermitian if and only if* $\langle a|\mathbf{H}|a\rangle$ *is real for all* $|a\rangle$.

Proof We have already pointed out that a hermitian operator has real expectation values. Conversely, assume that $\langle a|\mathbf{H}|a\rangle$ is real for all $|a\rangle$. Then

$$\langle a|\mathbf{H}|a\rangle = \langle a|\mathbf{H}|a\rangle^* = \langle a|\mathbf{H}^\dagger|a\rangle \quad \Leftrightarrow \quad \langle a|\mathbf{H} - \mathbf{H}^\dagger|a\rangle = 0 \quad \forall|a\rangle.$$

By Theorem 2.3.8 we must have $\mathbf{H} - \mathbf{H}^\dagger = \mathbf{0}$. □

Example 4.3.8 In this example, we illustrate the result of the above theorem with 2×2 matrices. The matrix $\mathsf{H} = \left(\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right)$ is hermitian[7] and acts on $\mathbb{C}^2$. Let us take an arbitrary vector $|a\rangle = \left(\begin{smallmatrix} \alpha_1 \\ \alpha_2 \end{smallmatrix}\right)$ and evaluate $\langle a|\mathsf{H}|a\rangle$. We have

$$\mathsf{H}|a\rangle = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} -i\alpha_2 \\ i\alpha_1 \end{pmatrix}.$$

[7]We assume that the reader has a casual familiarity with hermitian *matrices*. Think of an $n \times n$ matrix as a linear operator that acts on column vectors whose elements are components of vectors defined in the standard basis of $\mathbb{C}^n$ or $\mathbb{R}^n$. A hermitian matrix then becomes a hermitian operator.

Therefore,

$$\langle a|\mathbf{H}|a\rangle = \begin{pmatrix}\alpha_1^* & \alpha_2^*\end{pmatrix}\begin{pmatrix}-i\alpha_2 \\ i\alpha_1\end{pmatrix} = -i\alpha_1^*\alpha_2 + i\alpha_2^*\alpha_1$$
$$= i\alpha_2^*\alpha_1 + \left(i\alpha_2^*\alpha_1\right)^* = 2\,\mathrm{Re}\left(i\alpha_2^*\alpha_1\right),$$

and $\langle a|\mathbf{H}|a\rangle$ is real.

For the most general 2×2 hermitian matrix $\mathsf{H} = \begin{pmatrix}\alpha & \beta \\ \beta^* & \gamma\end{pmatrix}$, where α and γ are real, we have

$$\mathbf{H}|a\rangle = \begin{pmatrix}\alpha & \beta \\ \beta^* & \gamma\end{pmatrix}\begin{pmatrix}\alpha_1 \\ \alpha_2\end{pmatrix} = \begin{pmatrix}\alpha\alpha_1 + \beta\alpha_2 \\ \beta^*\alpha_1 + \gamma\alpha_2\end{pmatrix}$$

and

$$\langle a|\mathbf{H}|a\rangle = \begin{pmatrix}\alpha_1^* & \alpha_2^*\end{pmatrix}\begin{pmatrix}\alpha\alpha_1 + \beta\alpha_2 \\ \beta^*\alpha_1 + \gamma\alpha_2\end{pmatrix} = \alpha_1^*(\alpha\alpha_1 + \beta\alpha_2) + \alpha_2^*\left(\beta^*\alpha_1 + \gamma\alpha_2\right)$$
$$= \alpha|\alpha_1|^2 + \alpha_1^*\beta\alpha_2 + \alpha_2^*\beta^*\alpha_1 + \gamma|\alpha_2|^2$$
$$= \alpha|\alpha_1|^2 + \gamma|\alpha_2|^2 + 2\,\mathrm{Re}\left(\alpha_1^*\beta\alpha_2\right).$$

Again $\langle a|\mathbf{H}|a\rangle$ is real.

Definition 4.3.9 An operator $\mathbf{A}$ on an inner product space is called **positive** (written $\mathbf{A} \geq 0$) if $\langle a|\mathbf{A}|a\rangle \geq 0$ for all $|a\rangle \neq |0\rangle$. By Theorem 4.3.7 $\mathbf{A}$ is necessarily hermitian. We say $\mathbf{A}$ is **strictly positive** or **positive definite** (written $\mathbf{A} > 0$) if $\langle a|\mathbf{A}|a\rangle > 0$ for all $|a\rangle \neq |0\rangle$.

positive, strictly positive, and positive definite operators

Theorem 4.3.10 *A strictly positive operator* $\mathbf{T}$ *is invertible.*

Proof By Proposition 2.3.14, it is sufficient to prove that $\ker\mathbf{T} = \{|0\rangle\}$. If $|0\rangle \neq |a\rangle \in \ker\mathbf{T}$, then $\langle a|\mathbf{T}|a\rangle = 0$, contradicting the fact that $\mathbf{T}$ is strictly positive. □

Example 4.3.11 An example of a positive operator is the square of a hermitian operator.[8] We note that for any hermitian operator $\mathbf{H}$ and any vector $|a\rangle$, we have $\langle a|\mathbf{H}^2|a\rangle = \langle a|\mathbf{H}^\dagger\mathbf{H}|a\rangle = \langle\mathbf{H}a \mid \mathbf{H}a\rangle \geq 0$ because of the positive definiteness of the inner product.

From the discussion of the example above, we conclude that the square of an *invertible* hermitian operator is positive definite.

[8]This is further evidence that hermitian operators are analogues of real numbers: The square of any real number is positive.

4.3.2 Unitary Operators

The reader may be familiar with two- and three-dimensional rigid rotations and the fact that they preserve distances and the scalar product.[9] Can this be generalized to complex inner product spaces? Let $|a\rangle, |b\rangle \in \mathcal{V}$, and let $\mathbf{U}$ be an operator on $\mathcal{V}$ that preserves the scalar product; that is, given $|b'\rangle = \mathbf{U}|b\rangle$ and $|a'\rangle = \mathbf{U}|a\rangle$, then $\langle a' \mid b'\rangle = \langle a \mid b\rangle$. This yields

$$\langle a' \mid b'\rangle = \big(\langle a|\mathbf{U}^\dagger\big)\big(\mathbf{U}|b\rangle\big) = \langle a|\mathbf{U}^\dagger\mathbf{U}|b\rangle = \langle a \mid b\rangle = \langle a|\mathbf{1}|b\rangle.$$

Since this is true for arbitrary $|a\rangle$ and $|b\rangle$, we obtain $\mathbf{U}^\dagger\mathbf{U} = \mathbf{1}$. In the next chapter, when we introduce the concept of the determinant of operators, we shall see that this relation implies that $\mathbf{U}$ and $\mathbf{U}^\dagger$ are both invertible,[10] with each one being the inverse of the other.

unitary operators

Definition 4.3.12 Let $\mathcal{V}$ be a finite-dimensional inner product space. An operator $\mathbf{U}$ is called a **unitary operator** if $\mathbf{U}^\dagger = \mathbf{U}^{-1}$. Unitary operators preserve the inner product of $\mathcal{V}$.

Example 4.3.13 The linear transformation $\mathbf{T} : \mathbb{C}^3 \to \mathbb{C}^3$ given by

$$\mathbf{T}\begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{pmatrix} = \begin{pmatrix}(\alpha_1 - i\alpha_2)/\sqrt{2}\\ (\alpha_1 + i\alpha_2 - 2\alpha_3)/\sqrt{6}\\ \{\alpha_1 - \alpha_2 + \alpha_3 + i(\alpha_1 + \alpha_2 + \alpha_3)\}/\sqrt{6}\end{pmatrix}$$

is unitary. In fact, let

$$|a\rangle = \begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{pmatrix} \quad \text{and} \quad |b\rangle = \begin{pmatrix}\beta_1\\ \beta_2\\ \beta_3\end{pmatrix}$$

with dual vectors $\langle a| = (\alpha_1^* \; \alpha_2^* \; \alpha_3^*)$ and $\langle b| = (\beta_1^* \; \beta_2^* \; \beta_3^*)$, respectively. We use Eq. (4.11) and the procedure of Example 4.3.4 to find $\mathbf{T}^\dagger$. The result is

$$\mathbf{T}^\dagger\begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{pmatrix} = \begin{pmatrix}\frac{\alpha_1}{\sqrt{2}} + \frac{\alpha_2}{\sqrt{6}} + \frac{\alpha_3(1-i)}{\sqrt{6}}\\ \frac{i\alpha_1}{\sqrt{2}} - \frac{i\alpha_2}{\sqrt{6}} - \frac{\alpha_3(1+i)}{\sqrt{6}}\\ -\frac{2\alpha_2}{\sqrt{6}} + \frac{\alpha_3(1-i)}{\sqrt{6}}\end{pmatrix},$$

and we can verify that

$$\mathbf{T}\mathbf{T}^\dagger\begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{pmatrix} = \begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{pmatrix}.$$

[9] We have also encountered isometries, which are more general than unitary operators. The word "unitary" is usually reserved for isometries on sesquilinear (hermitian) inner product spaces.

[10] This implication holds only for finite-dimensional vector spaces.

Thus $\mathbf{T}\mathbf{T}^\dagger = \mathbf{1}$. Similarly, we can show that $\mathbf{T}^\dagger\mathbf{T} = \mathbf{1}$ and therefore that $\mathbf{T}$ is unitary.

4.4 Idempotents

We have already considered the decomposition of a vector space into subspaces in Sect. 2.1.3. We have also pointed out the significance of subspaces resulting from the fact that physics frequently takes place not inside the whole vector space, but in one of its subspaces. For instance, the study of projectile motion teaches us that it is very convenient to "project" the motion onto the horizontal and vertical axes and to study these projections separately. It is, therefore, appropriate to ask how we can go from a full space to one of its subspaces in the context of linear operators.

Let us first consider a simple example. A point in the plane is designated by the coordinates (x, y). A subspace of the plane is the x-axis. Is there a *linear* operator,[11] say $\mathbf{P}_x$, that acts on such a point and somehow sends it into that subspace? Of course, there are many operators from $\mathbb{R}^2$ to $\mathbb{R}$. However, we are looking for a specific one. We want $\mathbf{P}_x$ to *project* the point onto the x-axis. Such an operator has to act on (x, y) and produce $(x, 0)$: $\mathbf{P}_x(x, y) = (x, 0)$. Therefore, if the point already lies on the x-axis, $\mathbf{P}_x$ does not change it. In particular, if we apply $\mathbf{P}_x$ twice, we get the same result as if we apply it only once. And this is true for any point in the plane. Therefore, our operator must have the property $\mathbf{P}_x^2 = \mathbf{P}_x$, i.e., it must be an idempotent of the algebra $\mathrm{End}(\mathcal{V})$.

Suppose that $\mathcal{V}$ is the direct sum of r of its subspaces:

$$\mathcal{V} = \mathcal{U}_1 \oplus \cdots \oplus \mathcal{U}_r = \bigoplus_{i=1}^{r} \mathcal{U}_i.$$

For any $|v\rangle \in \mathcal{V}$, define $\mathbf{P}_j$ by $\mathbf{P}_j|v\rangle = |v_j\rangle$, where $|v_j\rangle$ is the component of $|v\rangle$ in $\mathcal{U}_j$. It is easy (and instructive) to prove that $\mathbf{P}_j$ is a linear operator; i.e., that $\mathbf{P}_j \in \mathrm{End}(\mathcal{V})$. Moreover, since $\mathbf{P}_j|v_j\rangle = |v_j\rangle$ for all $|v_j\rangle \in \mathcal{U}_j$, we have

$$\mathbf{P}_j^2|v\rangle = \mathbf{P}_j\mathbf{P}_j|v\rangle = \mathbf{P}_j|v_j\rangle = |v_j\rangle = \mathbf{P}_j|v\rangle$$

for all $|v\rangle \in \mathcal{V}$. It follows that $\mathbf{P}_j^2 = \mathbf{P}_j$, i.e., that $\mathbf{P}_j$ is an idempotent.

Next note that, for $j \neq k$,

$$\mathbf{P}_j\mathbf{P}_k|v\rangle = \mathbf{P}_j|v_k\rangle = |0\rangle,$$

because $|v_k\rangle$ has no component in $\mathcal{U}_j$. Since this is true for all $j \neq k$ and all $|v\rangle \in \mathcal{V}$, we have

$$\mathbf{P}_j\mathbf{P}_k = \mathbf{P}_k\mathbf{P}_j = \mathbf{0},$$

[11] We want this operator to preserve the vector-space structure of the plane and the axis.

i.e., that the idempotents are orthogonal. Furthermore, since

$$\mathbf{1}|v\rangle = |v\rangle = \sum_{j=1}^{r} |v_j\rangle = \sum_{j=1}^{r} \mathbf{P}_j|v\rangle = \left(\sum_{j=1}^{r} \mathbf{P}_j\right)|v\rangle$$

for all $|v\rangle \in \mathcal{V}$, we must have $\sum_{j=1}^{r} \mathbf{P}_j = \mathbf{1}$. We summarize the foregoing observation as

Proposition 4.4.1 *Let $\mathcal{V}$ be the direct sum of $\{\mathcal{U}_i\}_{i=1}^r$ and let $\mathbf{P}_j \in \mathcal{L}(\mathcal{V})$ be defined by $\mathbf{P}_j|v\rangle = |v_j\rangle$ with $|v_j\rangle$ the component of $|v\rangle$ in $\mathcal{U}_j$. Then $\{\mathbf{P}_i\}_{i=1}^r$ is a complete set of orthogonal idempotents, i.e.,*

$$\mathbf{P}_i\mathbf{P}_j = \delta_{ij}\mathbf{P}_i \text{ (no sum)} \quad \text{and} \quad \sum_{j=1}^{r} \mathbf{P}_j = \mathbf{1}.$$

If $\mathbf{T} \in \mathrm{End}(\mathcal{V})$ and $\{\mathbf{P}_i\}_{i=1}^r$ is a complete set of orthogonal idempotents corresponding to $\{\mathcal{U}_i\}_{i=1}^r$, then $\mathbf{T}$ can be written as a sum of operators which restrict to the subspaces. More precisely, multiply the sum in Proposition 4.4.1 by $\mathbf{T}$ on the right to get

$$\mathbf{T} = \sum_{j=1}^{r} \mathbf{P}_j\mathbf{T} \equiv \sum_{j=1}^{r} \mathbf{T}_j, \quad \mathbf{T}_j = \mathbf{P}_j\mathbf{T} \tag{4.13}$$

and note that $\mathbf{T}_j \in \mathrm{End}(\mathcal{U}_j)$.

4.4.1 Projection Operators

When the vector space carries an inner product, it is useful to demand hermiticity for the idempotents:

projection operators

Definition 4.4.2 A hermitian idempotent of $\mathrm{End}(\mathcal{V})$ is called a **projection operator**.

Consider two projection operators $\mathbf{P}_1$ and $\mathbf{P}_2$. We want to investigate conditions under which $\mathbf{P}_1 + \mathbf{P}_2$ becomes a projection operator. By definition,

$$\mathbf{P}_1 + \mathbf{P}_2 = (\mathbf{P}_1 + \mathbf{P}_2)^2 = \mathbf{P}_1^2 + \mathbf{P}_1\mathbf{P}_2 + \mathbf{P}_2\mathbf{P}_1 + \mathbf{P}_2^2 = \mathbf{P}_1 + \mathbf{P}_1\mathbf{P}_2 + \mathbf{P}_2\mathbf{P}_1 + \mathbf{P}_2.$$

So $\mathbf{P}_1 + \mathbf{P}_2$ is a projection operator if and only if

$$\mathbf{P}_1\mathbf{P}_2 + \mathbf{P}_2\mathbf{P}_1 = 0. \tag{4.14}$$

Multiply this on the left by $\mathbf{P}_1$ to get

$$\mathbf{P}_1^2\mathbf{P}_2 + \mathbf{P}_1\mathbf{P}_2\mathbf{P}_1 = 0 \quad \Rightarrow \quad \mathbf{P}_1\mathbf{P}_2 + \mathbf{P}_1\mathbf{P}_2\mathbf{P}_1 = 0.$$

Now multiply the same equation on the right by $\mathbf{P}_1$ to get

$$\mathbf{P}_1\mathbf{P}_2\mathbf{P}_1 + \mathbf{P}_2\mathbf{P}_1^2 = 0 \quad \Rightarrow \quad \mathbf{P}_1\mathbf{P}_2\mathbf{P}_1 + \mathbf{P}_2\mathbf{P}_1 = 0.$$

These last two equations yield

$$\mathbf{P}_1\mathbf{P}_2 - \mathbf{P}_2\mathbf{P}_1 = 0. \tag{4.15}$$

The solution to Eqs. (4.14) and (4.15) is $\mathbf{P}_1\mathbf{P}_2 = \mathbf{P}_2\mathbf{P}_1 = 0$. We therefore have the following result.

Proposition 4.4.3 *Let $\mathbf{P}_1, \mathbf{P}_2 \in \mathcal{L}(\mathcal{V})$ be projection operators. Then $\mathbf{P}_1 + \mathbf{P}_2$ is a projection operator if and only if $\mathbf{P}_1$ and $\mathbf{P}_2$ are orthogonal.*

More generally, if there is a set $\{\mathbf{P}_i\}_{i=1}^m$ of projection operators satisfying

$$\mathbf{P}_i\mathbf{P}_j = \begin{cases} \mathbf{P}_i & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$$

then $\mathbf{P} = \sum_{i=1}^m \mathbf{P}_i$ is also a projection operator.

Given a normal vector $|e\rangle$, one can show easily that $\mathbf{P} = |e\rangle\langle e|$ is a projection operator:

- $\mathbf{P}$ is hermitian: $\mathbf{P}^\dagger = (|e\rangle\langle e|)^\dagger = (\langle e|)^\dagger(|e\rangle)^\dagger = |e\rangle\langle e|$.
- $\mathbf{P}$ equals its square: $\mathbf{P}^2 = (|e\rangle\langle e|)(|e\rangle\langle e|) = |e\rangle \underbrace{\langle e|e\rangle}_{=1} \langle e| = |e\rangle\langle e|$.

Let $|y\rangle$ be any nonzero vector in an inner product space $\mathcal{V}$. The normal vector $|e_y\rangle$ along $|y\rangle$ is $|e_y\rangle = |y\rangle/\sqrt{\langle y \mid y\rangle}$. From this, we construct the projection operator $\mathbf{P}_y$ along $|y\rangle$:

$$\mathbf{P}_y = |e_y\rangle\langle e_y| = \frac{|y\rangle}{\sqrt{\langle y \mid y\rangle}} \frac{\langle y|}{\sqrt{\langle y \mid y\rangle}} = \frac{|y\rangle\langle y|}{\langle y \mid y\rangle}.$$

Given any vector $|x\rangle$, $\mathbf{P}_y|x\rangle$ is the component of $|x\rangle$ along $|y\rangle$. Hence,

$$|x\rangle - \mathbf{P}_y|x\rangle \quad \text{or} \quad (\mathbf{1} - \mathbf{P}_y)|x\rangle$$

is the component of $|x\rangle$ perpendicular to $|y\rangle$. The reflection of $|x\rangle$ in $|y\rangle$ is therefore (see Fig. 4.2)

$$\mathbf{P}_y|x\rangle - (\mathbf{1} - \mathbf{P}_y)|x\rangle = 2\mathbf{P}_y|x\rangle - |x\rangle = (2\mathbf{P}_y - \mathbf{1})|x\rangle = 2\frac{\langle y \mid x\rangle}{\langle y \mid y\rangle}|y\rangle - |x\rangle.$$

As shown in Fig. 4.2, from a two- or three-dimenstional geometric point of view, it is clear that the negative of this last vector is the reflection in the plane perpendicular to $|y\rangle$.[12] We generalize this to any vector space:

Definition 4.4.4 For any nonzero vector $|y\rangle$ the projection operator $\mathbf{P}_y$ along $|y\rangle$ and the **reflection operator** $\mathbf{R}_y$ in the plane perpendicular to $|y\rangle$

reflection operator

[12] One can note more directly—also shown in Fig. 4.2—that in three-dimensional geometry, if one adds to $|x\rangle$ twice the negative of its projection on $|y\rangle$, one gets the reflection of $|x\rangle$ in the plane perpendicular to $|y\rangle$.

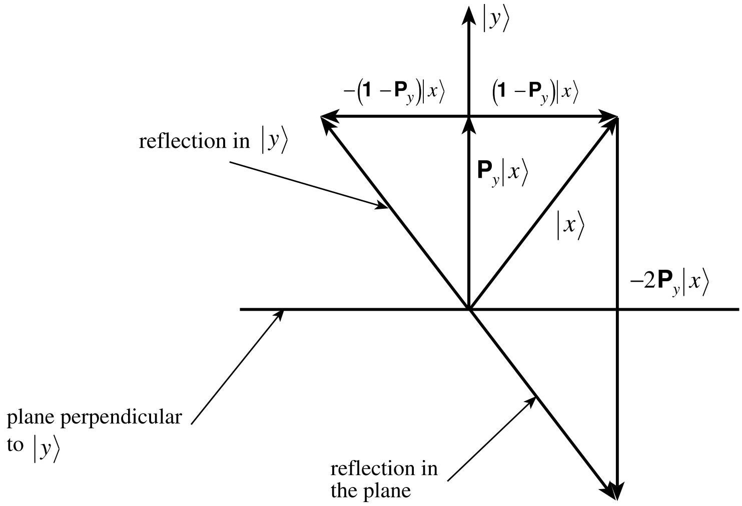

Fig. 4.2 The vectors $|x\rangle$ and $|y\rangle$ and the reflections of $|x\rangle$ in $|y\rangle$ and in the plane perpendicular to $|y\rangle$

are given by

$$\mathbf{P}_y = \frac{|y\rangle\langle y|}{\langle y|y\rangle} \quad \text{and} \quad \mathbf{R}_y = \mathbf{1} - 2\mathbf{P}_y = \mathbf{1} - 2\frac{|y\rangle\langle y|}{\langle y|y\rangle}.$$

For any other vector $|x\rangle$, the component $|x\rangle_y$ of $|x\rangle$ along $|y\rangle$ and its reflection $|x\rangle_{r,y}$ in the plane perpendicular to $|y\rangle$ are given by

$$|x\rangle_y = \mathbf{P}_y|x\rangle = \frac{\langle y|x\rangle}{\langle y|y\rangle}|y\rangle \quad \text{and} \quad |x\rangle_{r,y} = \mathbf{R}_y|x\rangle = |x\rangle - 2\frac{\langle y|x\rangle}{\langle y|y\rangle}|y\rangle.$$

The relations $|y\rangle_y = |y\rangle$ and $|y\rangle_{r,y} = -|y\rangle$ confirm our intuitive geometrical expectation.

Example 4.4.5 Let $\mathcal{V}$ be a one-dimensional vector space. Let $|a\rangle$ be any nonzero vector in $\mathcal{V}$. Any other vector $|x\rangle$ can be written as $|x\rangle = \alpha|a\rangle$ for some number α. Then

$$\mathbf{P}_a|x\rangle = \frac{|a\rangle\langle a|}{\langle a|a\rangle}|x\rangle = \frac{\langle a|x\rangle}{\langle a|a\rangle}|a\rangle = \frac{\alpha\langle a|a\rangle}{\langle a|a\rangle}|a\rangle = \alpha|a\rangle = |x\rangle.$$

Since this is true for all $|a\rangle$ and $|x\rangle$, we conclude that $\mathbf{P}_y = \mathbf{1}$ for any $|y\rangle$ in a one-dimensional vector space. The reflection operator can also be found. Therefore,

$$\mathbf{P}_y = \mathbf{1} \quad \text{and} \quad \mathbf{R}_y = -\mathbf{1}$$

in a one-dimensional vector space such as the complex vector space $\mathbb{C}$ and the real vector space $\mathbb{R}$.

We can take an orthonormal basis $B = \{|e_i\rangle\}_{i=1}^N$ and construct a set of projection operators $\{\mathbf{P}_i = |e_i\rangle\langle e_i|\}_{i=1}^N$. The operators $\mathbf{P}_i$ are mutually orthogonal. Thus, their sum $\sum_{i=1}^N \mathbf{P}_i$ is also a projection operator.

Proposition 4.4.6 *Let $B = \{|e_i\rangle\}_{i=1}^N$ be an orthonormal basis for $\mathcal{V}_N$. Then the set $\{\mathbf{P}_i = |e_i\rangle\langle e_i|\}_{i=1}^N$ consists of mutually orthogonal projection operators, and*

completeness relation

$$\sum_{i=1}^N \mathbf{P}_i = \sum_{i=1}^N |e_i\rangle\langle e_i| = \mathbf{1}.$$

This relation is called the ***completeness relation****.*

Proof The proof is left as Problem 4.26. □

If we choose only the first $m < N$ vectors, then the projection operator $\mathbf{P}^{(m)} \equiv \sum_{i=1}^m |e_i\rangle\langle e_i|$ projects arbitrary vectors into the subspace spanned by the first m basis vectors $\{|e_i\rangle\}_{i=1}^m$. In other words, when $\mathbf{P}^{(m)}$ acts on any vector $|a\rangle \in \mathcal{V}$, the result will be a linear combination of only the first m vectors. The simple proof of this fact is left as an exercise. These points are illustrated in the following example.

Example 4.4.7 Consider three orthonormal vectors $\{|e_i\rangle\}_{i=1}^3 \in \mathbb{R}^3$ given by

$$|e_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix}, \qquad |e_2\rangle = \frac{1}{\sqrt{6}}\begin{pmatrix}1\\-1\\2\end{pmatrix}, \qquad |e_3\rangle = \frac{1}{\sqrt{3}}\begin{pmatrix}-1\\1\\1\end{pmatrix}.$$

The projection operators associated with each of these can be obtained by noting that $\langle e_i|$ is a row vector. Therefore,

$$\mathbf{P}_1 = |e_1\rangle\langle e_1| = \frac{1}{2}\begin{pmatrix}1\\1\\0\end{pmatrix}\begin{pmatrix}1 & 1 & 0\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & 1 & 0\\1 & 1 & 0\\0 & 0 & 0\end{pmatrix}.$$

Similarly,

$$\mathbf{P}_2 = \frac{1}{6}\begin{pmatrix}1\\-1\\2\end{pmatrix}\begin{pmatrix}1 & -1 & 2\end{pmatrix} = \frac{1}{6}\begin{pmatrix}1 & -1 & 2\\-1 & 1 & -2\\2 & -2 & 4\end{pmatrix}$$

and

$$\mathbf{P}_3 = \frac{1}{3}\begin{pmatrix}-1\\1\\1\end{pmatrix}\begin{pmatrix}-1 & 1 & 1\end{pmatrix} = \frac{1}{3}\begin{pmatrix}1 & -1 & -1\\-1 & 1 & 1\\-1 & 1 & 1\end{pmatrix}.$$

Note that $\mathbf{P}_i$ projects onto the line along $|e_i\rangle$. This can be tested by letting $\mathbf{P}_i$ act on an arbitrary vector and showing that the resulting vector is perpendicular to the other two vectors. For example, let $\mathbf{P}_2$ act on an arbitrary column vector:

$$|a\rangle \equiv \mathbf{P}_2\begin{pmatrix}x\\y\\z\end{pmatrix} = \frac{1}{6}\begin{pmatrix}1 & -1 & 2\\-1 & 1 & -2\\2 & -2 & 4\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix} = \frac{1}{6}\begin{pmatrix}x-y+2z\\-x+y-2z\\2x-2y+4z\end{pmatrix}.$$

We verify that $|a\rangle$ is perpendicular to both $|e_1\rangle$ and $|e_3\rangle$:

$$\langle e_1|a\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1 & 0\end{pmatrix}\frac{1}{6}\begin{pmatrix}x-y+2z\\ -x+y-2z\\ 2x-2y+4z\end{pmatrix} = 0.$$

Similarly, $\langle e_3|a\rangle = 0$. So indeed, $|a\rangle$ is along $|e_2\rangle$.

We can find the operator that projects onto the plane formed by $|e_1\rangle$ and $|e_2\rangle$. This is

$$\mathbf{P}_1 + \mathbf{P}_2 = \frac{1}{3}\begin{pmatrix}2 & 1 & 1\\ 1 & 2 & -1\\ 1 & -1 & 2\end{pmatrix}.$$

When this operator acts on an arbitrary column vector, it produces a vector lying in the plane of $|e_1\rangle$ and $|e_2\rangle$, or perpendicular to $|e_3\rangle$:

$$|b\rangle \equiv (\mathbf{P}_1 + \mathbf{P}_2)\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \frac{1}{3}\begin{pmatrix}2 & 1 & 1\\ 1 & 2 & -1\\ 1 & -1 & 2\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \frac{1}{3}\begin{pmatrix}2x+y+z\\ x+2y-z\\ x-y+2z\end{pmatrix}.$$

It is easy to show that $\langle e_3|b\rangle = 0$. The operators that project onto the other two planes are obtained similarly. Finally, we verify easily that

$$\mathbf{P}_1 + \mathbf{P}_2 + \mathbf{P}_3 = \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} = \mathbf{1},$$

i.e., that completeness relation holds.

Example 4.4.8 We want to find the most general projection and reflection operators in a real two-dimensional vector space $\mathcal{V}$. Without loss of generality, we assume that $\mathcal{V} = \mathbb{R}^2$, and consider a vector

$$|y\rangle = \begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix}.$$

Then

$$\mathbf{P}_y = \frac{|y\rangle\langle y|}{\langle y|y\rangle} = \frac{1}{\eta_1^2+\eta_2^2}\begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix}\begin{pmatrix}\eta_1 & \eta_2\end{pmatrix} = \frac{1}{\eta_1^2+\eta_2^2}\begin{pmatrix}\eta_1^2 & \eta_1\eta_2\\ \eta_1\eta_2 & \eta_2^2\end{pmatrix}.$$

Let

$$\frac{\eta_1^2}{\eta_1^2+\eta_2^2} \equiv \cos^2\alpha \quad\Rightarrow\quad \frac{\eta_1}{\sqrt{\eta_1^2+\eta_2^2}} = \pm\cos\alpha,$$

and

$$\frac{\eta_2^2}{\eta_1^2+\eta_2^2} \equiv \sin^2\alpha \quad\Rightarrow\quad \frac{\eta_2}{\sqrt{\eta_1^2+\eta_2^2}} = \pm\sin\alpha.$$

If the product $\eta_1\eta_2$ is negative, we can define a new angle which is the negative of the old angle. This will change the sign of $\eta_1\eta_2$ and make it positive without changing the signs of η_1^2 and η_2^2. Thus the most general projection operator in $\mathbb{R}^2$ is

$$\mathbf{P}_y = \begin{pmatrix} \cos^2\alpha & \sin\alpha\cos\alpha \\ \sin\alpha\cos\alpha & \sin^2\alpha \end{pmatrix}. \tag{4.16}$$

Now that we have the projection operator along $|y\rangle$, we can construct the reflection operator in a plane perpendicular to $|y\rangle$:

$$\begin{aligned}\mathbf{R}_y = \mathbf{1} - 2\mathbf{P}_y &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - 2\begin{pmatrix} \cos\alpha^2 & \sin\alpha\cos\alpha \\ \sin\alpha\cos\alpha & \sin\alpha^2 \end{pmatrix} \\ &= \begin{pmatrix} 1-2\cos\alpha^2 & -2\sin\alpha\cos\alpha \\ -2\sin\alpha\cos\alpha & 1-2\sin\alpha^2 \end{pmatrix} = \begin{pmatrix} -\cos 2\alpha & -\sin 2\alpha \\ -\sin 2\alpha & \cos 2\alpha \end{pmatrix}.\end{aligned}$$

Defining $\phi = -2\alpha$, we see that a general reflection is of the form

$$\mathbf{R}_\phi = \begin{pmatrix} -\cos\phi & \sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}.$$

It is interesting to note that

The product of two reflections in $\mathbb{R}^2$ is a rotation.

$$\mathbf{R}_{\phi_1}\mathbf{R}_{\phi_2} = \begin{pmatrix} \cos(\phi_2-\phi_1) & -\sin(\phi_2-\phi_1) \\ \sin(\phi_2-\phi_1) & \cos(\phi_2-\phi_1) \end{pmatrix}.$$

This matrix describes a rotation of angle $\phi_2 - \phi_1$ in $\mathbb{R}^2$ (see Problem 5.9 in Chap. 5), which is clearly an isometry. We have just shown that the product of two reflections is an isometry. It turns out that this is a general property of isometries: they can always be expressed as a product of reflections (see Theorem 26.5.17).

4.5 Representation of Algebras

The operator algebra, i.e., the algebra $\mathcal{L}(\mathcal{V})$ of the endomorphisms of a vector space, plays a significant role in physical applications as demonstrated so far in this chapter. These (abstract) operators take on a concrete (numerical) look once they are identified as matrices, the topic of Chap. 5. This suggests making an identification of any given algebra with a (sub)algebra of $\mathcal{L}(\mathcal{V})$, which subsequently could be identified as a collection of numbers—what physicists are after—constituting the rows and columns of matrices. The vague notion of "identification" is made precise by the concept of homomorphism of algebras.

For the following definition, it is convenient to introduce some notation. Let both $\mathbb{F}$ and $\mathbb{K}$ denote either $\mathbb{R}$ or $\mathbb{C}$ with the condition that $\mathbb{F} \subseteq \mathbb{K}$. So, for instance, when $\mathbb{F} = \mathbb{R}$, then $\mathbb{K}$ can be either $\mathbb{R}$ and $\mathbb{C}$; but when $\mathbb{F} = \mathbb{C}$, then $\mathbb{K}$ can be only $\mathbb{C}$. If $\mathcal{V}$ is a vector space over $\mathbb{K}$, we denote the algebra of its endomorphisms by $\mathcal{L}_{\mathbb{K}}(\mathcal{V})$ or $\mathrm{End}_{\mathbb{K}}(\mathcal{V})$. When there is no danger of confusion, we remove the subscript $\mathbb{K}$.

representation of an algebra in a vector space

Definition 4.5.1 Let $\mathcal{A}$ be an associative algebra over $\mathbb{F}$ with identity $\mathbf{1}_A$. A $\mathbb{K}$-**representation** of $\mathcal{A}$ in a vector space $\mathcal{V}$ over $\mathbb{K}$ is a homomorphism $\rho : \mathcal{A} \to \mathrm{End}_{\mathbb{K}}(\mathcal{V})$ such that $\rho(\mathbf{1}_A) = \mathbf{1}$, where $\mathbf{1}$ is the unit operator in $\mathrm{End}_{\mathbb{K}}(\mathcal{V})$. The representation ρ is said to be **faithful** if it is injective.

Proposition 4.5.2 *A nontrivial representation of a simple algebra is faithful.*

Proof Since any representation is a homomorphism, the proof follows immediately from Proposition 3.2.14. □

a representation of the algebra of quaternions

Example 4.5.3 Let $\mathcal{A} = \mathbb{H}$, the (real) algebra of the quaternions, i.e., elements of the form

$$q = q_0\mathbf{e}_0 + q_1\mathbf{e}_1 + q_2\mathbf{e}_2 + q_3\mathbf{e}_3, \quad q_i \in \mathbb{R}$$

as given in Example 3.1.16. Let $\mathcal{V} = \mathbb{R}^4$. Consider $\rho : \mathbb{H} \to \mathrm{End}_{\mathbb{R}}(\mathbb{R}^4)$, the real representation of quaternions given by

$$\mathbf{T}_q|x\rangle \equiv \mathbf{T}_q \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} q_0x_1 - q_1x_2 - q_2x_3 - q_3x_4 \\ q_1x_1 + q_0x_2 - q_3x_3 + q_2x_4 \\ q_2x_1 + q_3x_2 + q_0x_3 - q_1x_4 \\ q_3x_1 - q_2x_2 + q_1x_3 + q_0x_4 \end{pmatrix}$$

$$= \begin{pmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & -q_3 & q_2 \\ q_2 & q_3 & q_0 & -q_1 \\ q_3 & -q_2 & q_1 & q_0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

where $\mathbf{T}_q \equiv \rho(q) \in \mathcal{L}(\mathbb{R}^4)$, and for convenience, we have written the element of $\mathbb{R}^4$ as a column vector and introduced the matrix of q's. With this matrix, we can simply write

$$\rho(q) = \begin{pmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & -q_3 & q_2 \\ q_2 & q_3 & q_0 & -q_1 \\ q_3 & -q_2 & q_1 & q_0 \end{pmatrix}. \tag{4.17}$$

Using this matrix, it is straightforward, but slightly tedious, to show directly that $\rho(qp) = \rho(q)\rho(p)$. Hence, ρ is indeed a representation of $\mathbb{H}$. However, instead, we calculate the matrices corresponding to the basis vectors of $\mathbb{H}$. Since $q_0 = 1$ and $q_1 = q_2 = q_3 = 0$ for $\mathbf{e}_0$, we get $\rho(\mathbf{e}_0) = \mathbf{1}$, as we should, and as is evident from (4.17). Similarly, we can calculate the matrices of the other basis vectors. The results are given below

$$\rho(\mathbf{e}_0) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \rho(\mathbf{e}_1) = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

$$\rho(\mathbf{e}_2) = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \qquad \rho(\mathbf{e}_3) = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

It is now easier to check that $\rho(\mathbf{e}_i\mathbf{e}_j) = \rho(\mathbf{e}_i)\rho(\mathbf{e}_j)$ for $i, j = 0, 1, 2, 3$, and hence, by Proposition 3.1.19, that ρ is indeed a representation.

Definition 4.5.4 A subspace $\mathcal{W}$ of $\mathcal{V}$ is called **stable** or **invariant** under a representation $\rho : \mathcal{A} \to \mathrm{End}_{\mathbb{C}}(\mathcal{V}) \equiv \mathrm{End}(\mathcal{V})$ if $\rho(\mathbf{a})|w\rangle$ is in $\mathcal{W}$ for all $|w\rangle \in \mathcal{W}$ and all $\mathbf{a} \in \mathcal{A}$. A representation ρ is called **irreducible** if the only stable subspaces are $\mathcal{W} = \mathcal{V}$ and $\mathcal{W} = \{|0\rangle_V\}$.

stable, invariant, and irreducible subspaces

Problem 4.34 shows that if ρ is surjective, then it is irreducible.

Proposition 4.5.5 *Let $\rho : \mathcal{A} \to \mathrm{End}(\mathcal{V})$ be a representation and $|v\rangle$ an arbitrary nonzero vector in $\mathcal{V}$. Then*

$$\mathcal{W} \equiv \rho(\mathcal{A})|v\rangle = \bigl\{|w\rangle \in \mathcal{V} \mid |w\rangle = \rho(\mathbf{a})|v\rangle \text{ for some } \mathbf{a} \in \mathcal{A}\bigr\}$$

is a stable subspace of $\mathcal{V}$. In particular, ρ is irreducible if and only if $\rho(\mathcal{A})|v\rangle = \mathcal{V}$.

Proof The straightforward but instructive proof is the content of Problem 4.37. □

Isomorphic vector spaces are indistinguishable. So can be their corresponding representations. More precisely,

Definition 4.5.6 Suppose $\mathbf{T} : \mathcal{V}_1 \cong \mathcal{V}_2$ is an isomorphism. Two representations $\rho_1 : \mathcal{A} \to \mathrm{End}(\mathcal{V}_1)$ and $\rho_2 : \mathcal{A} \to \mathrm{End}(\mathcal{V}_2)$ are called **equivalent** if

$$\mathbf{T} \circ \rho_1(\mathbf{a}) = \rho_2(\mathbf{a}) \circ \mathbf{T} \quad \text{for all } \mathbf{a} \in \mathcal{A}.$$

We write $\rho_1 \sim \rho_2$ to indicate the equivalence of the two representations.

equivalent representations

Sometimes the condition of equivalence is written as

$$\rho_1(\mathbf{a}) = \mathbf{T}^{-1} \circ \rho_2(\mathbf{a}) \circ \mathbf{T} \quad \text{or} \quad \rho_2(\mathbf{a}) = \mathbf{T} \circ \rho_1(\mathbf{a}) \circ \mathbf{T}^{-1}. \tag{4.18}$$

Just as we can combine two vector spaces to get a new vector space, we can combine two representations to obtain a new representation.

Definition 4.5.7 Let ρ and η be representations of $\mathcal{A}$ in $\mathcal{U}$ and $\mathcal{V}$, respectively. We define representations $\rho \oplus \eta$, called the **direct sum**, and $\rho \otimes \eta$, called the **tensor product**, of ρ and η, respectively in $\mathcal{U} \oplus \mathcal{V}$ and $\mathcal{U} \otimes \mathcal{V}$ by

$$(\rho \oplus \eta)(\mathbf{a}) = \rho(\mathbf{a}) \oplus \eta(\mathbf{a}) \quad \mathbf{a} \in \mathcal{A}$$
$$(\rho \otimes \eta)(\mathbf{a}) = \rho(\mathbf{a}) \otimes \eta(\mathbf{a}) \quad \mathbf{a} \in \mathcal{A}.$$

It should be obvious that if $\rho_1 \sim \rho_2$ and $\eta_1 \sim \eta_2$, then $\rho_1 \oplus \eta_1 \sim \rho_2 \oplus \eta_2$ and $\rho_1 \otimes \eta_1 \sim \rho_2 \otimes \eta_2$.

using $|a\rangle$ and **a** in certain representations of an algebra

Since an algebra $\mathcal{A}$ is also a vector space, it is possible to come up with representations of the form $\rho : \mathcal{A} \to \operatorname{End}(\mathcal{A})$. When there is no danger of confusion, we designate members of $\mathcal{A}$ as a ket when they are considered simply as vectors, but use bold face type when the same member participates in an algebra multiplication. Thus $|a\rangle \in \mathcal{A}$ when the member is considered as a vector and $\mathbf{a} \in \mathcal{A}$ when the same member is one of the factors in a product.

regular representation of an algebra

Definition 4.5.8 The **regular representation** of $\mathcal{A}$ in $\mathcal{A}$ is the representation $\rho_L : \mathcal{A} \to \operatorname{End}(\mathcal{A})$ given by $\rho_L(\mathbf{a})|b\rangle = \mathbf{ab}$.

It is trivial to show that this is indeed a representation, i.e., that $\rho_L(\mathbf{ab}) = \rho_L(\mathbf{a})\rho_L(\mathbf{b})$.

If $\mathcal{A}$ is unital, then $\rho_L(\mathbf{a}) = \rho_L(\mathbf{a}')$ implies that $\rho_L(\mathbf{a})|1\rangle = \rho_L(\mathbf{a}')|1\rangle$ or that $\mathbf{a1} = \mathbf{a}'\mathbf{1}$, namely that $\mathbf{a} = \mathbf{a}'$, indicating that ρ_L is injective, and the representation faithful, or $\rho_L(\mathcal{A}) \cong \mathcal{A}$.

ρ_L is simply the left-multiplication of $\mathcal{A}$. What about the right-multiplication? If we set $\rho_R(\mathbf{a})|b\rangle = \mathbf{ba}$, then

$$\begin{aligned}\rho_R(\mathbf{ab})|c\rangle &= \mathbf{cab} = (\mathbf{ca})\mathbf{b} = \big(\rho_R(\mathbf{a})|c\rangle\big)\mathbf{b} = \rho_R(\mathbf{b})\big(\rho_R(\mathbf{a})|c\rangle\big) \\ &= \big(\rho_R(\mathbf{b})\rho_R(\mathbf{a})\big)|c\rangle.\end{aligned}$$

Hence, $\rho_R(\mathbf{ab}) = \rho_R(\mathbf{b})\rho_R(\mathbf{a})$. Again if $\mathcal{A}$ is unital, then ρ_R is faithful and $\rho_R(\mathcal{A}) \cong \mathcal{A}^{\mathrm{op}}$, where $\mathcal{A}^{\mathrm{op}}$ is the algebra opposite to $\mathcal{A}$ given in Definition 3.1.8.

Theorem 4.5.9 *Let $\mathcal{L}$ be a minimal left ideal of an algebra $\mathcal{A}$. Then the representation $\rho^{(\mathcal{L})} : \mathcal{A} \to \operatorname{End}(\mathcal{L})$, the regular representation of $\mathcal{A}$ in $\mathcal{L}$, given by*

$$\rho^{(\mathcal{L})}(\mathbf{a})|x\rangle \equiv \rho_L(\mathbf{a})|x\rangle = \mathbf{ax} \quad \textit{for } \mathbf{a} \in \mathcal{A} \textit{ and } |x\rangle \equiv \mathbf{x} \in \mathcal{L},$$

is irreducible.

Proof We note that $\rho^{(\mathcal{L})}(\mathcal{A})|x\rangle = \mathcal{A}\mathbf{x}$. Since $\mathcal{L}$ is minimal, $\mathcal{A}\mathbf{x} = \mathcal{L}$ by Theorem 3.2.6, and $\rho^{(\mathcal{L})}(\mathcal{A})|x\rangle = \mathcal{L}$. By Proposition 4.5.5, $\rho^{(\mathcal{L})}$ is irreducible. □

Theorem 4.5.10 *All irreducible representations of a simple algebra $\mathcal{A}$ are faithful and equivalent to $\rho^{(\mathcal{L})}$, the regular representation of $\mathcal{A}$ in the minimal left ideal $\mathcal{L}$.*

Proof The faithfulness is a consequence of Proposition 4.5.2. Let $\rho : \mathcal{A} \to \mathrm{End}(\mathcal{V})$ be an irreducible representation. For $\mathbf{x} \in \mathcal{L}$ and a vector $|e\rangle \in \mathcal{V}$, let $\rho(\mathbf{x})|e\rangle = |v\rangle$. Then

$$\rho(\mathcal{L})|e\rangle = \rho(\mathcal{A}\mathbf{x})|e\rangle = \rho(\mathcal{A})\rho(\mathbf{x})|e\rangle = \rho(\mathcal{A})|v\rangle = \mathcal{V}. \tag{4.19}$$

The first equality follows from Theorem 3.2.6, and the last equality from Proposition 4.5.5.

Now consider a linear map $\mathbf{T} : \mathcal{L} \to \mathcal{V}$ given by $\mathbf{T}(\mathbf{y}) = \rho(\mathbf{y})|e\rangle$, and note that by Eq. (4.19),

$$\mathbf{T}(\mathcal{L}) = \rho(\mathcal{L})|e\rangle = \mathcal{V}.$$

Therefore, $\mathbf{T}$ is surjective. Now let $\mathbf{z}$ be a nonzero member of $\mathcal{L}$. If $\mathbf{z} \in \ker \mathbf{T}$, then by Theorem 3.2.6, $\mathcal{L} = \mathcal{A}\mathbf{z}$ and

$$\mathbf{T}(\mathcal{L}) = \rho(\mathcal{L})|e\rangle = \rho(\mathcal{A}\mathbf{z})|e\rangle = \rho(\mathcal{A})\rho(\mathbf{z})|e\rangle = \rho(\mathcal{A})\mathbf{T}(\mathbf{z}) = \{\mathbf{0}\}$$

which contradicts the previous equation. Therefore, $\ker \mathbf{T} = \{\mathbf{0}\}$ and $\mathbf{T}$ is injective, hence, bijective.

To complete the proof, we have to show that

$$\mathbf{T} \circ \rho^{(\mathcal{L})}(\mathbf{a}) = \rho(\mathbf{a}) \circ \mathbf{T} \quad \text{for all } \mathbf{a} \in \mathcal{A}.$$

If $\mathbf{y} \in \mathcal{L}$, then the right-hand side gives

$$\rho(\mathbf{a}) \circ \mathbf{T}(\mathbf{y}) = \rho(\mathbf{a})\rho(\mathbf{y})|e\rangle = \rho(\mathbf{ay})|e\rangle = \mathbf{T}(\mathbf{ay}),$$

while the left-hand side yields

$$\big(\mathbf{T} \circ \rho^{(\mathcal{L})}(\mathbf{a})\big)\mathbf{y} = \mathbf{T}\big(\rho^{(\mathcal{L})}(\mathbf{a})\mathbf{y}\big) = \mathbf{T}(\mathbf{ay}).$$

This completes the proof. □

A consequence of this theorem is

Corollary 4.5.11 *All minimal left ideals of a simple algebra are isomorphic.*

Proof If $\mathcal{L}'$ is another left ideal of $\mathcal{A}$, then let $\mathcal{V} = \mathcal{L}'$ in Theorem 4.5.10. Then $\mathbf{T}$ of the theorem establishes an isomorphism between $\mathcal{L}$ and $\mathcal{L}'$. □

Theorem 4.5.12 *Two irreducible representations of a semi-simple algebra are equivalent if and only if they have the same kernel.*

Proof Recall from Theorem 3.5.25 that a semi-simple algebra $\mathcal{A}$ is the direct sum of simple algebras, each component being an ideal of $\mathcal{A}$. Let

$$\mathcal{A} = \mathcal{I}_1 \oplus \mathcal{I}_2 \oplus \cdots \oplus \mathcal{I}_r = \bigoplus_{i=1}^{r} \mathcal{I}_i$$

and $\rho : \mathcal{A} \to \operatorname{End}(\mathcal{V})$ be an irreducible representation. Assume there is $\mathbf{0} \neq \mathbf{x}_p \in \mathcal{I}_p$ for some p and $|e\rangle \in \mathcal{V}$ such that $|v\rangle \equiv \rho(\mathbf{x}_p)|e\rangle \neq |0\rangle$. Then since ρ is irreducible, by Proposition 4.5.5

$$\mathcal{V} = \rho(\mathcal{A})|v\rangle = \rho(\mathcal{A})\rho(\mathbf{x}_p)|e\rangle = \rho(\mathcal{A}\mathbf{x}_p)|e\rangle \subseteq \rho(\mathcal{I}_p)|e\rangle.$$

But obviously, $\rho(\mathcal{I}_p)|e\rangle \subseteq \mathcal{V}$. Hence,

$$\rho(\mathcal{I}_p)|e\rangle = \mathcal{V}, \tag{4.20}$$

which also indicates that any $|x\rangle \in \mathcal{V}$ can be written as $|x\rangle = \rho(\mathbf{y}_p)|e\rangle$ for some $\mathbf{y}_p \in \mathcal{I}_p$. Now since, $\mathcal{I}_p\mathcal{I}_k = \mathcal{I}_k\mathcal{I}_p = \{\mathbf{0}\}$ for $k \neq p$, we have

$$\rho(\mathbf{z}_k)|x\rangle = \rho(\mathbf{z}_k)\rho(\mathbf{y}_p)|e\rangle = \rho(\mathbf{z}_k\mathbf{y}_p)|e\rangle = \rho(\mathbf{0})|e\rangle = |0\rangle$$

for all $|x\rangle \in \mathcal{V}$. It follows that $\rho(\mathbf{z}_k)$ is the zero operator, i.e., $\mathbf{z}_k \in \ker\rho$ for all $k \neq p$, or

$$\bigoplus_{\substack{i=1\\ i\neq p}}^{r} \mathcal{I}_i \subseteq \ker\rho.$$

Now let $\rho|_{\mathcal{I}_p} : \mathcal{I}_p \to \operatorname{End}(\mathcal{V})$ be the restriction of ρ to $\mathcal{I}_p$, i.e., a representation of $\mathcal{I}_p$ in $\mathcal{V}$. Then $\mathbf{T} : \mathcal{I}_p \to \mathcal{V}$ given by $\mathbf{T}(\mathbf{z}_p) = \rho(\mathbf{z}_p)|e\rangle$ is an isomorphism by the proof of Theorem 4.5.10. Hence, $\rho(\mathbf{z}_p) = \mathbf{0}$ implies that $\mathbf{z}_p = \mathbf{0}$, i.e., that $\mathcal{I}_p \cap \ker\rho = \{\mathbf{0}\}$. This yields

$$\bigoplus_{\substack{i=1\\ i\neq p}}^{r} \mathcal{I}_i = \ker\rho.$$

Let $\rho_1 : \mathcal{A} \to \operatorname{End}(\mathcal{V}_1)$ and $\rho_2 : \mathcal{A} \to \operatorname{End}(\mathcal{V}_2)$ be two irreducible representations of a semi-simple algebra $\mathcal{A}$. Assume further that ρ_1 and ρ_2 have the same kernel; i.e. that for some $1 \leq p \leq r$,

$$\ker\rho_1 = \ker\rho_2 = \bigoplus_{\substack{i=1\\ i\neq p}}^{r} \mathcal{I}_i.$$

Then as shown above there are isomorphisms $\mathbf{T}_1 : \mathcal{I}_p \to \mathcal{V}_1$ and $\mathbf{T}_2 : \mathcal{I}_p \to \mathcal{V}_2$ given by

$$\mathbf{T}_1(\mathbf{z}_p) = \rho_1(\mathbf{z}_p)|e_1\rangle, \quad \text{and} \quad \mathbf{T}_2(\mathbf{z}_p) = \rho_2(\mathbf{z}_p)|e_2\rangle$$

with

$$\rho_1(\mathcal{I}_p)|e_1\rangle = \mathcal{V}_1 \quad \text{and} \quad \rho_2(\mathcal{I}_p)|e_2\rangle = \mathcal{V}_2 \tag{4.21}$$

as in Eq. (4.20). The composite map $\mathbf{S} \equiv \mathbf{T}_2 \circ \mathbf{T}_1^{-1}$ maps $\mathcal{V}_1$ isomorphically onto $\mathcal{V}_2$. We now show that

$$\mathbf{S} \circ \rho_1(\mathbf{a}) = \rho_2(\mathbf{a}) \circ \mathbf{S} \quad \text{for all } \mathbf{a} \in \mathcal{A},$$

and hence that $\rho_1 \sim \rho_2$. Applying the right-hand side of this equation on a $|v_1\rangle \in \mathcal{V}_1$, and noting that by (4.21) $|v_1\rangle = \rho_1(\mathbf{z}_p)|e_1\rangle$ for some $\mathbf{z}_p \in \mathcal{I}_p$, we get

$$\begin{aligned}\rho_2(\mathbf{a}) \circ \mathbf{S}|v_1\rangle &= \rho_2(\mathbf{a}) \circ \big(\mathbf{T}_2 \circ \mathbf{T}_1^{-1}\big)\rho_1(\mathbf{z}_p)|e_1\rangle = \rho_2(\mathbf{a}) \circ \mathbf{T}_2\big(\mathbf{T}_1^{-1}\rho_1(\mathbf{z}_p)|e_1\rangle\big)\\ &= \big(\rho_2(\mathbf{a})\big) \circ \mathbf{T}_2(\mathbf{z}_p) = \rho_2(\mathbf{a})\rho_2(\mathbf{z}_p)|e_2\rangle = \rho_2(\mathbf{a}\mathbf{z}_p)|e_2\rangle,\end{aligned}$$

while the left-hand side gives

$$\begin{aligned}\mathbf{S} \circ \rho_1(\mathbf{a})|v_1\rangle &= \big(\mathbf{T}_2 \circ \mathbf{T}_1^{-1}\big)\rho_1(\mathbf{a})\rho_1(\mathbf{z}_p)|e_1\rangle = \big(\mathbf{T}_2 \circ \mathbf{T}_1^{-1}\big)\rho_1(\mathbf{a}\mathbf{z}_p)|e_1\rangle\\ &= \mathbf{T}_2\big(\mathbf{T}_1^{-1}\rho_1(\mathbf{a}\mathbf{z}_p)|e_1\rangle\big) = \mathbf{T}_2(\mathbf{a}\mathbf{z}_p) = \rho_2(\mathbf{a}\mathbf{z}_p)|e_2\rangle.\end{aligned}$$

We have shown that if two irreducible representations of a semi-simple algebra have the same kernel, then they are equivalent. The converse is much easier to prove (see Problem 4.36). □

4.6 Problems

4.1 Consider a linear operator $\mathbf{T}$ on a finite-dimensional vector space $\mathcal{V}$.

(a) Show that there exists a polynomial p such that $p(\mathbf{T}) = \mathbf{0}$. Hint: Take a basis $B = \{|a_i\rangle\}_{i=1}^N$ and consider the vectors $\{\mathbf{T}^k|a_1\rangle\}_{k=0}^M$ for large enough M and conclude that there exists a polynomial $p_1(\mathbf{T})$ such that $p_1(\mathbf{T})|a_1\rangle = 0$. Do the same for $|a_2\rangle$, etc. Now take the product of all such polynomials.

(b) From (a) conclude that for large enough n, $\mathbf{T}^n$ can be written as a linear combination of smaller powers of $\mathbf{T}$.

(c) Now conclude that any infinite series in $\mathbf{T}$ collapses to a polynomial in $\mathbf{T}$.

4.2 Use mathematical induction to show that $[\mathbf{A}, \mathbf{A}^m] = \mathbf{0}$.

4.3 For $\mathbf{D}$ and $\mathbf{T}$ defined in Example 2.3.5:

(a) Show that $[\mathbf{D}, \mathbf{T}] = \mathbf{1}$.

(b) Calculate the linear transformations $\mathbf{D}^3\mathbf{T}^3$ and $\mathbf{T}^3\mathbf{D}^3$.

4.4 Consider three linear operators $\mathbf{L}_1$, $\mathbf{L}_2$, and $\mathbf{L}_3$ satisfying the commutation relations $[\mathbf{L}_1, \mathbf{L}_2] = i\mathbf{L}_3$, $[\mathbf{L}_3, \mathbf{L}_1] = i\mathbf{L}_2$, $[\mathbf{L}_2, \mathbf{L}_3] = i\mathbf{L}_1$, and define the new operators $\mathbf{L}_\pm = \mathbf{L}_1 \pm i\mathbf{L}_2$.

(a) Show that the operator $\mathbf{L}^2 \equiv \mathbf{L}_1^2 + \mathbf{L}_2^2 + \mathbf{L}_3^2$ commutes with $\mathbf{L}_k$, $k = 1, 2, 3$.
(b) Show that the set $\{\mathbf{L}_+, \mathbf{L}_-, \mathbf{L}_3\}$ *is closed under commutation*, i.e., the commutator of any two of them can be written as a linear combination of the set. Determine these commutators.
(c) Write $\mathbf{L}^2$ in terms of $\mathbf{L}_+$, $\mathbf{L}_-$, and $\mathbf{L}_3$.

4.5 Prove the rest of Proposition 4.1.8.

4.6 Show that if $[[\mathbf{A}, \mathbf{B}], \mathbf{A}] = \mathbf{0}$, then for every positive integer k,

$$\left[\mathbf{A}^k, \mathbf{B}\right] = k\mathbf{A}^{k-1}[\mathbf{A}, \mathbf{B}].$$

Hint: First prove the relation for low values of k; then use mathematical induction.

4.7 Show that for $\mathbf{D}$ and $\mathbf{T}$ defined in Example 2.3.5,

$$\left[\mathbf{D}^k, \mathbf{T}\right] = k\mathbf{D}^{k-1} \quad \text{and} \quad \left[\mathbf{T}^k, \mathbf{D}\right] = -k\mathbf{T}^{k-1}.$$

4.8 Evaluate the derivative of $\mathbf{H}^{-1}(t)$ in terms of the derivative of $\mathbf{H}(t)$ by differentiating their product.

4.9 Show that for any $\alpha, \beta \in \mathbb{R}$ and any $\mathbf{H} \in \mathrm{End}(\mathcal{V})$, we have

$$e^{\alpha\mathbf{H}}e^{\beta\mathbf{H}} = e^{(\alpha+\beta)\mathbf{H}}.$$

4.10 Show that $(\mathbf{U} + \mathbf{T})(\mathbf{U} - \mathbf{T}) = \mathbf{U}^2 - \mathbf{T}^2$ if and only if $[\mathbf{U}, \mathbf{T}] = \mathbf{0}$.

4.11 Prove that if $\mathbf{A}$ and $\mathbf{B}$ are hermitian, then $i[\mathbf{A}, \mathbf{B}]$ is also hermitian.

4.12 Find the solution to the operator differential equation

$$\frac{d\mathbf{U}}{dt} = t\mathbf{H}\mathbf{U}(t).$$

Hint: Make the change of variable $y = t^2$ and use the result of Example 4.2.3.

4.13 Verify that

$$\frac{d}{dt}\mathbf{H}^3 = \left(\frac{d\mathbf{H}}{dt}\right)\mathbf{H}^2 + \mathbf{H}\left(\frac{d\mathbf{H}}{dt}\right)\mathbf{H} + \mathbf{H}^2\left(\frac{d\mathbf{H}}{dt}\right).$$

4.14 Show that if $\mathbf{A}$ and $\mathbf{B}$ commute, and f and g are arbitrary functions, then $f(\mathbf{A})$ and $g(\mathbf{B})$ also commute.

4.15 Assuming that $[[\mathbf{S}, \mathbf{T}], \mathbf{T}] = \mathbf{0} = [[\mathbf{S}, \mathbf{T}], \mathbf{S}]$, show that

$$\left[\mathbf{S}, \exp(t\mathbf{T})\right] = t[\mathbf{S}, \mathbf{T}]\exp(t\mathbf{T}).$$

Hint: Expand the exponential and use Problem 4.6.

4.16 Prove that

$$\exp(\mathbf{H}_1 + \mathbf{H}_2 + \mathbf{H}_3) = \exp(\mathbf{H}_1)\exp(\mathbf{H}_2)\exp(\mathbf{H}_3) \times \exp\left\{-\frac{1}{2}\big([\mathbf{H}_1, \mathbf{H}_2] + [\mathbf{H}_1, \mathbf{H}_3] + [\mathbf{H}_2, \mathbf{H}_3]\big)\right\}$$

provided that $\mathbf{H}_1, \mathbf{H}_2$, and $\mathbf{H}_3$ commute with all the commutators. What is the generalization to $\mathbf{H}_1 + \mathbf{H}_2 + \cdots + \mathbf{H}_n$?

4.17 Denoting the derivative of $\mathbf{A}(t)$ by $\dot{\mathbf{A}}$, show that

$$\frac{d}{dt}[\mathbf{A}, \mathbf{B}] = [\dot{\mathbf{A}}, \mathbf{B}] + [\mathbf{A}, \dot{\mathbf{B}}].$$

4.18 Prove Theorem 4.3.2. Hint: Use Eq. (4.11) and Theorem 2.3.7.

4.19 Let $\mathbf{A}(t) \equiv \exp(t\mathbf{H})\mathbf{A}_0 \exp(-t\mathbf{H})$, where $\mathbf{H}$ and $\mathbf{A}_0$ are constant operators. Show that $d\mathbf{A}/dt = [\mathbf{H}, \mathbf{A}(t)]$. What happens when $\mathbf{H}$ commutes with $\mathbf{A}(t)$?

4.20 Let $|f\rangle, |g\rangle \in \mathbb{C}(a, b)$ with the additional property that

$$f(a) = g(a) = f(b) = g(b) = 0.$$

Show that for such functions, the derivative operator $\mathbf{D}$ is anti-hermitian. The inner product is defined as usual:

$$\langle f|g\rangle \equiv \int_a^b f^*(t)g(t)\,dt.$$

4.21 In this problem, you will go through the steps of proving the rigorous statement of the **Heisenberg uncertainty principle**. Denote the expectation (average) value of an operator $\mathbf{A}$ in a state $|\Psi\rangle$ by A_{avg}. Thus, $A_{\text{avg}} = \langle A\rangle = \langle\Psi|\mathbf{A}|\Psi\rangle$. The *uncertainty* (deviation from the mean) in the normalized state $|\Psi\rangle$ of the operator $\mathbf{A}$ is given by

$$\Delta A = \sqrt{\big\langle (A - A_{\text{avg}})^2\big\rangle} = \sqrt{\langle\Psi|(\mathbf{A} - A_{\text{avg}}\mathbf{1})^2|\Psi\rangle}.$$

(a) Show that for any two *hermitian* operators $\mathbf{A}$ and $\mathbf{B}$, we have

$$\big|\langle\Psi|\mathbf{AB}|\Psi\rangle\big|^2 \le \langle\Psi|\mathbf{A}^2|\Psi\rangle\langle\Psi|\mathbf{B}^2|\Psi\rangle.$$

Hint: Apply the Schwarz inequality to an appropriate pair of vectors.

(b) Using the above and the triangle inequality for complex numbers, show that

$$\big|\langle\Psi|[\mathbf{A}, \mathbf{B}]|\Psi\rangle\big|^2 \le 4\langle\Psi|\mathbf{A}^2|\Psi\rangle\langle\Psi|\mathbf{B}^2|\Psi\rangle.$$

(c) Define the operators $\mathbf{A}' = \mathbf{A} - \alpha\mathbf{1}$, $\mathbf{B}' = \mathbf{B} - \beta\mathbf{1}$, where α and β are *real* numbers. Show that $\mathbf{A}'$ and $\mathbf{B}'$ are hermitian and $[\mathbf{A}', \mathbf{B}'] = [\mathbf{A}, \mathbf{B}]$.

(d) Now use all the results above to show the celebrated uncertainty relation

Heisenberg uncertainty principle

$$(\Delta A)(\Delta B) \geq \frac{1}{2}\left|\langle\Psi|[\mathbf{A},\mathbf{B}]|\Psi\rangle\right|.$$

What does this reduce to for position operator $\mathbf{x}$ and momentum operator $\mathbf{p}$ if $[\mathbf{x},\mathbf{p}] = i\hbar$?

4.22 Show that $\mathbf{U} = \exp\mathbf{A}$ is unitary if $\mathbf{A}$ is anti-hermitian. Furthermore, if $\mathbf{A}$ commutes with $\mathbf{A}^\dagger$, then $\exp\mathbf{A}$ is unitary. Hint: Use Proposition 4.2.4 on $\mathbf{U}\mathbf{U}^\dagger = \mathbf{1}$ *and* $\mathbf{U}^\dagger\mathbf{U} = \mathbf{1}$

4.23 Find $\mathbf{T}^\dagger$ for each of the following linear operators.

(a) $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$\mathbf{T}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ x-y \end{pmatrix}.$$

(b) $\mathbf{T} : \mathbb{R}^3 \to \mathbb{R}^3$ given by

$$\mathbf{T}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+2y-z \\ 3x-y+2z \\ -x+2y+3z \end{pmatrix}.$$

(c) $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$\mathbf{T}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix},$$

where θ is a real number. What is $\mathbf{T}^\dagger\mathbf{T}$?

(d) $\mathbf{T} : \mathbb{C}^2 \to \mathbb{C}^2$ given by

$$\mathbf{T}\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} \alpha_1 - i\alpha_2 \\ i\alpha_1 + \alpha_2 \end{pmatrix}.$$

(e) $\mathbf{T} : \mathbb{C}^3 \to \mathbb{C}^3$ given by

$$\mathbf{T}\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} \alpha_1 + i\alpha_2 - 2i\alpha_3 \\ -2i\alpha_1 + \alpha_2 + i\alpha_3 \\ i\alpha_1 - 2i\alpha_2 + \alpha_3 \end{pmatrix}.$$

4.24 Show that if $\mathbf{P}$ is a (hermitian) projection operator, so are $\mathbf{1} - \mathbf{P}$ and $\mathbf{U}^\dagger\mathbf{P}\mathbf{U}$ for any unitary operator $\mathbf{U}$.

4.25 For the vector

$$|a\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix},$$

(a) find the associated projection matrix, $\mathbf{P}_a$.
(b) Verify that $\mathbf{P}_a$ does project an arbitrary vector in $\mathbb{C}^4$ along $|a\rangle$.
(c) Verify directly that the matrix $\mathbf{1} - \mathbf{P}_a$ is also a projection operator.

4.26 Prove Proposition 4.4.6

4.27 Let $|a_1\rangle \equiv \mathbf{a}_1 = (1, 1, -1)$ and $|a_2\rangle \equiv \mathbf{a}_2 = (-2, 1, -1)$.

(a) Construct (in the form of a matrix) the projection operators $\mathbf{P}_1$ and $\mathbf{P}_2$ that project onto the directions of $|a_1\rangle$ and $|a_2\rangle$, respectively. Verify that they are indeed projection operators.
(b) Construct (in the form of a matrix) the operator $\mathbf{P} = \mathbf{P}_1 + \mathbf{P}_2$ and verify directly that it is a projection operator.
(c) Let $\mathbf{P}$ act on an arbitrary vector (x, y, z). What is the dot product of the resulting vector with the vector $\mathbf{a}_1 \times \mathbf{a}_2$? What can you say about $\mathbf{P}$ and your conclusion in (b)?

4.28 Let $\mathbf{P}^{(m)} = \sum_{i=1}^{m} |e_i\rangle\langle e_i|$ be a projection operator constructed out of the first m orthonormal vectors of the basis $B = \{|e_i\rangle\}_{i=1}^{N}$ of $\mathcal{V}$. Show that $\mathbf{P}^{(m)}$ projects into the subspace spanned by the first m vectors in B.

4.29 What is the length of the projection of the vector $(3, 4, -4)$ onto a line whose parametric equation is $x = 2t + 1$, $y = -t + 3$, $z = t - 1$? Hint: Find a unit vector in the direction of the line and construct its projection operator.

4.30 The parametric equation of a line L in a coordinate system with origin O is

$$x = 2t + 1, \qquad y = t + 1, \qquad z = -2t + 2.$$

A point P has coordinates $(3, -2, 1)$.

(a) Using the projection operators, find the length of the projection of $\overline{OP}$ on the line L.
(b) Find the vector whose beginning is P and ends perpendicularly on L.
(c) From this vector calculate the distance from P to L.

4.31 Let the operator $\mathbf{U} : \mathbb{C}^2 \to \mathbb{C}^2$ be given by

$$\mathbf{U}\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} i\frac{\alpha_1}{\sqrt{2}} - i\frac{\alpha_2}{\sqrt{2}} \\ \frac{\alpha_1}{\sqrt{2}} + \frac{\alpha_2}{\sqrt{2}} \end{pmatrix}.$$

Find $\mathbf{U}^\dagger$ and test if $\mathbf{U}$ is unitary.

4.32 Show that the product of two unitary operators is always unitary, but the product of two hermitian operators is hermitian if and only if they commute.

4.33 Let $\mathbf{S}$ be an operator that is both unitary and hermitian. Show that

(a) $\mathbf{S}$ is involutive (i.e., $\mathbf{S}^2 = \mathbf{1}$), and

(b) $\mathsf{S} = \mathsf{P}^+ - \mathsf{P}^-$, where P^+ and P^- are hermitian projection operators.

4.34 Show that if a representation $\rho : \mathcal{A} \to \mathcal{L}(\mathcal{V})$ is surjective, then it is irreducible. Hint: The operator $|a\rangle\langle a|$ is in $\mathcal{L}(\mathcal{V})$ for any $|a\rangle \in \mathcal{V}$.

4.35 Show that $\rho(\mathbf{e}_i \mathbf{e}_j) = \rho(\mathbf{e}_i)\rho(\mathbf{e}_j)$ for $i, j = 0, 1, 2, 3$ in Example 4.5.3.

4.36 Show that *any* two equivalent representations of *any* algebra have the same kernel.

4.37 To prove Proposition 4.5.5, first show that $\rho(\mathcal{A})|v\rangle$ is a subspace. Then prove that $\rho(\mathcal{A})\mathcal{W} \subset \mathcal{W}$. For the "only if" part of an irreducible representation, take $|v\rangle$ to be in any subspace of $\mathcal{V}$.