1

Vectors and Matrices

In this chapter, we present basic definitions and results of complex vectors and matrices needed in the study of quantum information science and quantum computing using the linear algebra (Hilbert space) approach. Linear algebra and matrix theory are useful in many applied and research areas; one may see [2, 5, 6] for general background. We shall emphasize the roles of vectors and matrices in quantum mechanics.

1.1 Complex vectors, linearly independent sets, and bases

Let $\mathbb{C}^n$ be the set of column vectors with n complex entries $x_1, \ldots, x_n$. Occasionally, we will focus on the set of real vectors $\mathbb{R}^n$ in $\mathbb{C}^n$.* Using the **Dirac notation**, we denote a vector in $\mathbb{C}^n$ by

$$|x\rangle = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad x_1, \ldots, x_n \in \mathbb{C} \tag{1.1}$$

It is often written as a **transpose** of a row vector, as $|x\rangle = (x_1, x_2, \ldots, x_n)^t$, to save space. Recall that the complex conjugate of a complex number $z = a+ib$ is $z^* = a-ib$, and the modulus of z is $|z| = \sqrt{z^* z} = \sqrt{a^2 + b^2}$. In mathematics books, z^* is often written in the form $\bar{z}$. We will follow the physicist's notation.

An element $|x\rangle$ is also called a **ket vector** or simply a **ket**. The **bra vector**, or simple the **bra**, $\langle x|$ associated with $|x\rangle$ is the row vector

$$|x\rangle \mapsto \langle x| = (x_1^*, \ldots, x_n^*). \tag{1.2}$$

For example, if $|x\rangle = (1, i, 1-i)^t$, then $\langle x| = (1, -i, 1+i)$. Note that each component is complex-conjugated under this correspondence. The bra and ket vectors are the basic entities needed to model quantum systems.

*Readers may be more familiar with real vectors, but one needs to use complex vectors to model quantum systems as we will see.

For $|x\rangle = (x_1, \ldots, x_n)^t, |y\rangle = (y_1, \ldots, y_n)^t \in \mathbb{C}^n$ and $a \in \mathbb{C}$, define the addition of $|x\rangle$ and $|y\rangle$, and the scalar multiplication of a and $|x\rangle$ as

$$|x\rangle + |y\rangle = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix} \quad \text{and} \quad a|x\rangle = \begin{pmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_n \end{pmatrix}, \tag{1.3}$$

respectively. All the components of the **zero-vector** $|\mathbf{0}\rangle$ are zero. The zero-vector is also written as $\mathbf{0}$ in a less strict manner. Under the addition and scalar multiplication defined above, $\mathbb{C}^n$ forms a vector space with the zero element $|\mathbf{0}\rangle$. Readers can consult linear algebra books for general background.

A **linear combination** of $|v_1\rangle, \ldots, |v_k\rangle \in \mathbb{C}^n$ is a vector of the form

$$|v\rangle = c_1|v_1\rangle + \cdots + c_k|v_k\rangle \quad \text{with } c_1, \ldots, c_k \in \mathbb{C}.$$

The set of all linear combinations of vectors in a set S in $\mathbb{C}^n$ is denoted by $\operatorname{Span} S$, which is a linear subspace of $\mathbb{C}^n$.

A subset $S = \{|v_1\rangle, \ldots, |v_k\rangle\}$ of $\mathbb{C}^n$ is **linearly dependent** if there is $(c_1, \ldots, c_k) \neq (0, \ldots, 0)$ such that

$$c_1|v_1\rangle + \cdots + c_k|v_k\rangle = |\mathbf{0}\rangle. \tag{1.4}$$

If $(c_1, \ldots, c_k) = (0, \ldots, 0)$ is the only solution of Eq. (1.4), the set is said to be **linearly independent**. Construct the $n \times k$ matrix $A = (|v_1\rangle \cdots |v_k\rangle)$. Then the set S is linearly dependent if and only if there is a nonzero vector $|c\rangle = (c_1, \ldots, c_k)^t$ such that $A|c\rangle = |\mathbf{0}\rangle \in \mathbb{C}^{\mathbf{n}}$. Using the basic theory of linear equations, one readily deduces the following.

- If S is linearly independent, then $k \leq n$. Equivalently, the set S is linearly dependent if $k > n$.
- The set S is linearly dependent if one of the vectors is expressed as a linear combination of the other vectors.
- If $|\mathbf{0}\rangle \in S$, then S is linearly dependent.

A subset S of $\mathbb{C}^n$ is a **generating set** of $\mathbb{C}^n$ if every vector in $\mathbb{C}^n$ is a linear combination of vectors in S, i.e., $\operatorname{Span} S = \mathbb{C}^n$. A linearly independent generating set is a **basis** for $\mathbb{C}^n$. The vectors are called **basis vectors**. Again, one may use the basic theory of linear equations to deduce the following.

- If $S \subseteq \mathbb{C}^n$ has fewer than n elements, then S is not a generating set.
- A set $S \subseteq \mathbb{C}^n$ with n elements is a basis if any one of the following conditions holds:

  (a) S is a linearly independent set. (b) S is a generating set.

- Every basis of $\mathbb{C}^n$ has n elements. We say that $\mathbb{C}^n$ has dimension n.
- If $\{|v_1\rangle, \ldots, |v_n\rangle\}$ is a basis for $\mathbb{C}^n$, then every $|x\rangle \in \mathbb{C}^n$ admits a unique representation $|x\rangle = \sum_{j=1} c_j |v_j\rangle$. The n complex numbers $c_1, \ldots, c_n$ are called the **components** of $|x\rangle$ with respect to the basis $\{|v_1\rangle, \ldots, |v_n\rangle\}$.

1.2 Inner product, Gram-Schmidt orthonormalization

The **inner product** of $|x\rangle = (x_1, \ldots, x_n)^t, |y\rangle = (y_1, \ldots, y_n)^t$ in $\mathbb{C}^n$ is defined by

$$\langle x|y\rangle = \sum_{j=1}^{n} x_j^* y_j \tag{1.5}$$

This product is nothing but an ordinary matrix multiplication of a $1 \times n$ matrix and an $n \times 1$ matrix once the matrix multiplication is defined. In the mathematical literature, complex conjugation is taken rather with respect to the y_j. In the present book, we use the physics convention (1.5). Note the following sesquilinearity:*

$$\langle x|c_1 y_1 + c_2 y_2\rangle = c_1 \langle x|y_1\rangle + c_2 \langle x|y_2\rangle \tag{1.6}$$
$$\langle c_1 x_1 + c_2 x_2|y\rangle = c_1^* \langle x_1|y\rangle + c_2^* \langle x_2|y\rangle, \tag{1.7}$$

where $|c_1 y_1 + c_2 y_2\rangle \equiv c_1|y_1\rangle + c_2|y_2\rangle$.

The (inner) **norm** of a vector $|x\rangle = (x_1, \ldots, x_n)^t \in \mathbb{C}^n$ is defined by

$$\| |x\rangle \| = \sqrt{\langle x|x\rangle} = \left[\sum_{j=1}^{n} |x_j|^2\right]^{1/2} \geq 0. \tag{1.8}$$

If $|x\rangle$ has norm one, we say that $|x\rangle$ is a **unit vector**. Each nonzero vector $|y\rangle \in \mathbb{C}^n$ can be **normalized** to a unit vector as $|y\rangle / \| |y\rangle \|$.

A subset $S = \{|v_1\rangle, \ldots, |v_k\rangle\}$ of $\mathbb{C}^n$ is **orthogonal** if $\langle v_r|v_s\rangle = 0$ whenever $r \neq s$. If in addition that $\langle v_r|v_r\rangle = 1$ for all $r = 1, \ldots, k$, then S is an **orthonormal set**.

An orthonormal set $\{|v_1\rangle, \ldots, |v_k\rangle\}$ is always independent. To see this, suppose $\sum_{j=1}^{k} c_j |v_j\rangle = |\mathbf{0}\rangle$. Then $0 = \langle v_\ell|\mathbf{0}\rangle = \langle v_\ell| \sum_{j=1}^{k} c_j |v_j\rangle = c_\ell$ for $\ell = 1, \ldots, k$. As a result, an orthnormal set $\{|e_1\rangle, \ldots, |e_n\rangle\}$ in $\mathbb{C}^n$ with n

*sesqui = 1.5.

elements must be a basis, which is called an **orthonormal basis**. Clearly, $\{|e_1\rangle, \ldots, |e_n\rangle\} \subseteq \mathbb{C}^n$ is an orthonormal basis if and only if

$$\langle e_i|e_j\rangle = \delta_{ij}, \tag{1.9}$$

where

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & \text{otherwise} \end{cases}$$

is the Kronecker delta. Every vector can be written uniquely as a linear combination of the vectors in a basis. For an orthonormal basis $\{|e_1\rangle, \ldots, |e_n\rangle\}$, it is easy to express $|x\rangle \in \mathbb{C}^n$ as $|x\rangle = \sum_{j=1} c_j|e_j\rangle$, namely, $c_j = \langle e_j|x\rangle$ for $j = 1, \ldots, n$.

There are many orthonormal bases for $\mathbb{C}^n$, and it may be important to use a special one in a specific application. For instance, the orthonormal bases

$$\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \quad \text{and} \quad \left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$$

can be used to represent the vertical and horizontal polarizations of a photon, and the polarizations of a photon passing through polarization plate making a $\pm 45^\circ$ with the vertical axis, respectively.

One can always apply the following **Gram-Schmidt process** to construct an orthonormal set $\{|e_1\rangle, \ldots, |e_k\rangle\}$ from a given linearly independent set $\{|v_1\rangle, \ldots, |v_k\rangle\}$ in $\mathbb{C}^n$ such that $\text{Span}\{|v_1\rangle, \ldots, |v_\ell\rangle\} = \text{Span}\{|e_1\rangle, \ldots, |e_\ell\rangle\}$ for $\ell = 1, \ldots, k$. In case $k = n$, we get an orthonormal basis for $\mathbb{C}^n$.

The Gram-Schmidt orthonormalization Process.

Let $|e_1\rangle = |v_1\rangle/\||v_1\rangle\|$. For $j = 2, \ldots, k$, let $|e_j\rangle = |f_j\rangle/\||f_j\rangle\|$, where

$$|f_j\rangle = |v_j\rangle - \sum_{\ell=1}^{j-1} \langle e_\ell|v_j\rangle|e_\ell\rangle.$$

Note that $|f_j\rangle \neq |\mathbf{0}\rangle$; otherwise, $\{|v_1\rangle, \ldots, |v_j\rangle\}$ is linearly dependent, and so is $\{|v_1\rangle, \ldots, |v_k\rangle\}$. So, $|f_j\rangle/\||f_j\rangle\|$ is a well-defined unit vector.

By construction, one sees that $\langle e_r|e_s\rangle = 0$ for all $1 \leq r < s \leq k$, and $\text{Span}\{|v_1\rangle, \ldots, |v_\ell\rangle\} = \text{Span}\{|e_1\rangle, \ldots, |e_\ell\rangle\}$ for $\ell = 1, \ldots, k$.

Denote by $\mathbf{M}_{r,s}$ the set of $r \times s$ complex matrices, and $\mathbf{M}_r = \mathbf{M}_{r,r}$. If $A \in \mathbf{M}_{n,k}$ has linearly independent columns $|v_1\rangle, \ldots, |v_k\rangle$, then the Gram-Schmidt procedure yields a matrix $U \in \mathbf{M}_{n,k}$ with columns $|e_1\rangle, \ldots, |e_k\rangle$ such that $U^\dagger U = I_k$ and $A = UR$ for an upper triangular matrix R. This is known as the QR decomposition of A in mathematics books.

EXAMPLE 1.2.1. *Let* $A = \begin{pmatrix} 1 & 2i \\ i & 0 \end{pmatrix}$. *Applying Gram-Schmidt process to the columns of* A, *we obtain*

$$|e_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}, \; |f_2\rangle = \begin{pmatrix} 2i \\ 0 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 \\ i \end{pmatrix}(1, -i)\begin{pmatrix} 2i \\ 0 \end{pmatrix} = \begin{pmatrix} i \\ 1 \end{pmatrix}, \; |e_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} i \\ 1 \end{pmatrix}.$$

Then $A = QR$ *with* $Q = (|e_1\rangle, |e_2\rangle) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$ *and* $R = \sqrt{2}\begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix}$.

1.3 Matrices, linear transformations, and dual space

General quantum states, quantum operations, measurement operations, etc. are modeled by matrices. Here we describe some basic properties of matrices needed in the subsequent discussion. Again, we will focus on the results relevant to quantum mechanics.

For $A = (a_{rs}), B = (b_{rs}) \in \mathbf{M}_{m,n}$ and $c \in \mathbb{C}$, define the addition and scalar multiplication by

$$C = A + B = (a_{rs} + b_{rs}), \qquad \text{and} \qquad cA = (ca_{rs}),$$

respectively. Then $\mathbf{M}_{m,n}$ is a linear space under these operations. If $A \in \mathbf{M}_{m,n}$ has rows $\langle u_1|, \ldots, \langle u_m|$ and $B = \mathbf{M}_{n,\ell}$ has columns $|v_1\rangle, \ldots, |v_\ell\rangle$, then the product of A and B is

$$C = AB = (c_{rs}) \in \mathbf{M}_{m,\ell} \qquad \text{with } c_{rs} = \langle u_r|v_s\rangle \;\text{ for } 1 \le r \le m, 1 \le s \le \ell.$$

The following observations about matrix products are useful.

- The matrix C has columns $A|v_1\rangle, \ldots, A|v_\ell\rangle$, and rows $\langle u_1|B, \ldots, \langle u_m|B$.
- (**Block multiplication**) If $A = (A_{pq}), B = (B_{rs})$ are block matrices with $1 \le p \le \hat{m}, 1 \le q, r \le \hat{n}, 1 \le s \le \hat{\ell}$ so that the block matrices $A_{pq}B_{rs}$ is defined, i.e., the number of columns A_{pq} and the number of the rows of B_{rs} are the same, whenever $q = r$, then $AB = (C_{uv})$ with $C_{uv} = \sum_{\ell=1}^{\hat{n}} A_{u\ell}B_{\ell v}$ for $1 \le u \le \hat{m}, 1 \le v \le \hat{n}$.
- In particular, if A has columns $|x_1\rangle, \ldots, |x_n\rangle$ and B has rows $\langle y_1|, \ldots, \langle y_n|$, then

  $$AB = |x_1\rangle\langle y_1| + \cdots + |x_n\rangle\langle y_n|,$$

  where $|x_j\rangle\langle y_j|$ has rank at most one for $j = 1, \ldots, n$.

- If $m = \ell$ and A is a diagonal matrix with diagonal entries $a_1, \ldots, a_m$, then AB has rows $a_1\langle y_1|, \ldots, a_m\langle y_m|$, i.e.,

$$AB = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_m \end{pmatrix} \begin{pmatrix} \langle y_1| \\ \vdots \\ \langle y_m| \end{pmatrix} = \begin{pmatrix} a_1\langle y_1| \\ \vdots \\ a_m\langle y_m| \end{pmatrix}.$$

- If $\ell = n$ and B is a diagonal matrix with diagonal entries $b_1, \ldots, b_n$, then AB has columns $b_1|x_1\rangle, \ldots, b_n|x_n\rangle$, i.e.,

$$AB = (|x_1\rangle \cdots |x_n\rangle) \begin{pmatrix} b_1 & & \\ & \ddots & \\ & & b_n \end{pmatrix} = \left(b_1|x_1\rangle \cdots b_n|x_n\rangle\right).$$

The product of $A, B \in \mathbf{M}_n$ is a matrix in $\mathbf{M}_n$. The set $\mathbf{M}_n$ is an algebra under the addition, scalar multiplication, and the matrix multiplication defined above.

EXAMPLE 1.3.1. *Let* $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 1 & i & 2 \\ 0 & 1 & -i \end{pmatrix}$, *and* $D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$. *Then*

$$AB = \begin{pmatrix} 1 \\ 3 \end{pmatrix} (1, i, 2) + \begin{pmatrix} 2 \\ 4 \end{pmatrix} (0, 1, -i) = \begin{pmatrix} 1 & 2+i & 2-2i \\ 3 & 4+3i & 6-4i \end{pmatrix},$$

and

$$ADB = 2\begin{pmatrix} 1 \\ 3 \end{pmatrix} (1, i, 2) + 3\begin{pmatrix} 2 \\ 4 \end{pmatrix} (0, 1, -i) = \begin{pmatrix} 8 & 8i & 16 \\ 18 & 18i & 36 \end{pmatrix}.$$

The **conjugate** of $A = (a_{rs}) \in \mathbf{M}_{m,n}$ is $A^* = (a^*_{rs}) \in \mathbf{M}_{m,n}$. The **Hermitian conjugate** of A is $A^\dagger = (A^*)^t$.[†] It is clear that the (r, s) entry of $A^\dagger$ is the complex conjugate of the (s, r) entry of A. This definition also applies to a ket vector $|x\rangle$. We have

$$|x\rangle^\dagger = (x_1^*, \ldots, x_n^*) = \langle x|.$$

Namely, the procedure to produce a bra vector from a ket vector is regarded as a Hermitian conjugation of the ket vector.

The following properties of the Hermitian conjugate are clear for $A \in \mathbf{M}_{m,n}$ and $B \in \mathbf{M}_{n,\ell}$:

$$(A^\dagger)^\dagger = A \quad \text{and} \quad (AB)^\dagger = B^\dagger A^\dagger. \tag{1.10}$$

Using the Hermitian conjugate, we can define the following special types of matrices, which are useful in the study of quantum science.

[†] Mathematicians use A^* to denote the Hermitian conjugate of A. We will follow the physics convention.

- A matrix $A \in \mathbf{M}_n$ is **Hermitian** if $A^\dagger = A$.

  Equivalently, the (r, s) entry of A equals the conjugate of the (s, r) entry of A, i.e., $a_{rs} = a_{sr}^*$ for all $1 \le r, s \le n$ if $A = (a_{rs})$. In particular, the diagonal entries of A are real.

- A matrix $A \in \mathbf{M}_n$ is **skew-Hermitian** if $A^\dagger = -A$.

  Equivalently, the (r, s) entry of A equals the negative conjugate of the (s, r) entry of A, i.e., $a_{rs} = -a_{sr}^*$ for all $1 \le r, s \le n$ if $A = (a_{rs})$. In particular, the diagonal entries of A are pure imaginary. It is also equivalent to the condition that iA is Hermitian.

- A matrix $A \in \mathbf{M}_n$ is **unitary** if $A^\dagger A = I_n$, i.e., $A^\dagger = A^{-1}$.

If A is real, then A is Hermitian means that A is symmetric; A is skew-Hermitian means that A is skew-symmetric with vanishing (zero) diagonal entries. Unlike Hermitian and Skew-Hermitian matrices, it is not easy to detect a unitary matrix by looking at its entries. Nevertheless, one easily verifies the following.

PROPOSITION 1.3.2. *A matrix $U \in \mathbf{M}_n$ is unitary if and only if the columns $|u_1\rangle, \dots, |u_n\rangle$ form an orthonormal basis for $\mathbb{C}^n$.*

Note also that U is unitary if and only if $U^{-1} = U^\dagger$. Consequently, $I_n = UU^\dagger$ and we obtain the **completeness relation**

$$I_n = UU^\dagger = \sum_{j=1}^{n} |u_j\rangle\langle u_j|. \tag{1.11}$$

The completeness relation is actually a property for any orthonormal basis $\{|e_1\rangle, \dots, |e_n\rangle\}$ of $\mathbb{C}^n$ and is quite useful. The *matrix*

$$P_k \equiv |e_k\rangle\langle e_k| \tag{1.12}$$

introduced above is the **projection operator** in the direction defined by $|e_k\rangle$. This projects a vector $|v\rangle$ to a vector parallel to $|e_k\rangle$ in such a way that $|v\rangle - P_k|v\rangle$ is orthogonal to $|e_k\rangle$ (see Fig. 1.1). The set $\{P_1, \dots, P_n\}$ satisfies the conditions

$$\text{(i)} \qquad P_k^2 = P_k, \tag{1.13}$$

$$\text{(ii)} \qquad P_k P_j = 0 \quad (k \neq j), \tag{1.14}$$

$$\text{(iii)} \qquad \sum_k P_k = I_n \quad \text{(completeness relation)}. \tag{1.15}$$

The conditions (i) and (ii) are obvious from the orthonormality $\langle e_j | e_k\rangle = \delta_{jk}$.

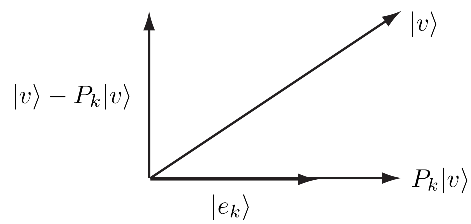

FIGURE 1.1
A vector $|v\rangle$ is projected to the direction defined by a unit vector $|e_k\rangle$ by the action of $P_k = |e_k\rangle\langle e_k|$. The difference $|v\rangle - P_k|v\rangle$ is orthogonal to $|e_k\rangle$.

EXAMPLE 1.3.3. *Let* $\theta \in \mathbb{R}$*,*

$$|e_1\rangle = \begin{pmatrix} \cos\theta \\ e^{i\phi}\sin\theta \end{pmatrix} \quad \textit{and} \quad |e_2\rangle = \begin{pmatrix} -\sin\theta \\ e^{i\phi}\cos\theta \end{pmatrix}.$$

Then one readily checks that $\{|e_1\rangle, |e_2\rangle\}$ *is an orthonormal basis. The corresponding projection operators* $P_1 = |e_1\rangle\langle e_1|$ *and* $P_2 = |e_2\rangle\langle e_2|$ *are*

$$P_1 = \begin{pmatrix} \cos^2\theta & e^{-i\phi}\cos\theta\sin\theta \\ e^{i\phi}\cos\theta\sin\theta & \sin^2\theta \end{pmatrix}, \quad P_2 = \begin{pmatrix} \sin^2\theta & -e^{-i\phi}\cos\theta\sin\theta \\ -e^{i\phi}\cos\theta\sin\theta & \cos^2\theta \end{pmatrix}.$$

They satisfy the completeness relation

$$\sum_k P_k = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$$

and the orthogonality condition

$$P_1 P_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

One readily verifies that $P_k^2 = P_k$*.*

Unitary matrices are important tools for transforming quantum states represented as vectors. In general, we consider a **linear transformation** (also known as linear operator, linear function, or linear map), $T : \mathbb{C}^n \to \mathbb{C}^m$, i.e., a function satisfying

$$T(|x\rangle + |y\rangle) = T(|x\rangle) + T(|y\rangle) \quad \text{and} \quad T(c|x\rangle) = cT(|x\rangle) \tag{1.16}$$

for arbitrary $|x\rangle, |y\rangle \in \mathbb{C}^n$ and $c \in \mathbb{C}$. Note that every linear operator $T : \mathbb{C}^n \to \mathbb{C}^m$ has the form

$$T(|v\rangle) = A|v\rangle \qquad \text{for all } |v\rangle \in \mathbb{C}^n, \tag{1.17}$$

for a matrix $A \in \mathbf{M}_{m,n}$. It is not hard to check that a function of the form (1.17) is linear. Conversely, suppose $T : \mathbb{C}^n \to \mathbb{C}^m$ is a linear operator. Let $\{|e_1\rangle, \ldots, |e_n\rangle\}$ be the standard basis for $\mathbb{C}^n$, i.e., $|e_j\rangle$ equals the jth column of the identity matrix I_n for $j = 1, \ldots, n$. Similarly, let $\{|f_1\rangle, \ldots, |f_m\rangle\}$ be the standard basis for $\mathbb{C}^m$. Set $A \in \mathbf{M}_{m,n}$ with columns $T(|e_1\rangle), \ldots, T(|e_n\rangle)$. Then for any $|x\rangle = \sum_{j=1}^n x_j|e_j\rangle$, we have

$$T(|x\rangle) = \sum_{j=1} x_j T(|e_j\rangle) = A|x\rangle.$$

EXAMPLE 1.3.4. *Let $T : \mathbb{C}^3 \to \mathbb{C}^2$ be a linear map, and $\{|e_1\rangle, |e_2\rangle, |e_3\rangle\}$ be the standard bases for $\mathbb{C}^3$. If*

$$T(|e_1\rangle) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \; T(|e_2\rangle) = \begin{pmatrix} 3 \\ -i \end{pmatrix}, \quad \textit{and} \;\; T(|e_3\rangle) = \begin{pmatrix} -i \\ 1 \end{pmatrix},$$

then

$$T(|x\rangle) = \begin{pmatrix} 1 & 3 & -i \\ 2 & -i & 1 \end{pmatrix} |x\rangle \qquad \textit{for all} \;\; |x\rangle \in \mathbb{C}^3.$$

As we shall see, quantum operations $T : \mathbb{C}^n \to \mathbb{C}^n$ transforming quantum states represented as vectors in $\mathbb{C}^n$ must be of the form

$$T(|x\rangle) = U|x\rangle \quad \text{ for all } |x\rangle \in \mathbb{C}^n$$

for a unitary matrix U.

For readers with group theory background, it is interesting to note the following. The set of unitary matrices form a group called the **unitary group**, which is denoted by $\mathrm{U}(n)$. Note that for a unitary $U \in \mathbf{M}_n$,

$$1 = \det(I_n) = \det(U^\dagger U) = |\det(U^\dagger)\det(U)| = |\det(U)|^2.$$

Hence, $\det(U) = e^{i\alpha}$ for some $\alpha \in \mathbb{R}$. A special unitary matrix is a unitary matrix with determinant 1. The set of special unitary matrices is a group called the **special unitary group**, which is denoted by $\mathrm{SU}(n)$. A real unitary matrix A is called an orthogonal matrix, which satisfies $\det(A) = \pm 1$. If $\det A = 1$, it is called a special orthogonal matrix. The set of orthogonal (special orthogonal) matrices is a group called the **orthogonal group** (**special orthogonal group**) and denoted by $\mathrm{O}(n)$ ($\mathrm{SO}(n)$).

As mentioned before, there are different orthonormal bases of $\mathbb{C}^n$, which may be useful for specific problems. For instance, suppose $T : \mathbb{C}^n \to \mathbb{C}^m$ is a linear operator such that $T(|x\rangle) = A|x\rangle$ for all $|x\rangle \in \mathbb{C}^n$ with $A \in \mathbf{M}_{m,n}$. Let $\mathcal{B}_1 = \{|v_1\rangle, \ldots, |v_n\rangle\} \subseteq \mathbb{C}^n$ and $\mathcal{B}_2 = \{|u_1\rangle, \ldots, |u_m\rangle\} \subseteq \mathbb{C}^m$ be orthonormal bases so that $V = (|v_1\rangle \;\cdots\; |v_n\rangle) \in \mathbf{M}_n$ and $U = (|u_1\rangle \;\cdots\; |u_m\rangle) \in \mathbf{M}_m$ are unitary. Suppose $B = (b_{rs}) \in \mathbf{M}_{m,n}$ with columns $|b_1\rangle, \ldots, |b_n\rangle$ satisfies

$$T(|v_j\rangle) = \sum_{\ell=1}^m b_{\ell j}|u_\ell\rangle = U|b_j\rangle, \quad j = 1, \ldots, n.$$

Then

$$AV = A(|v_1\rangle\ \dots\ |v_n\rangle) = (U|b_1\rangle\ \cdots\ U|b_n\rangle) = UB,$$

and hence $B = U^\dagger AV$. We will show in Section 1.6 that orthonormal bases $\mathcal{B}_1$ and $\mathcal{B}_2$ can be chosen such that B only has nonzero entries in the (j, j) position for $j = 1, \dots, k$ with $k \le \min\{m, n\}$. To do this, we will need the concept of eigenvalues and eigenvectors for $A \in \mathbf{M}_n$ treated in the next section.

We conclude this section by considering the set of linear functions $f : \mathbb{C}^n \to \mathbb{C}$, which is often called a linear functional. By the previous discussion, there is an $1 \times n$ matrix (a row vector) $(\alpha_1, \dots, \alpha_n)$, which can be regarded as the bra vector $\langle y|$ of $|y\rangle = (\alpha_1^*, \dots, \alpha_n^*)^t \in \mathbb{C}^n$, such that

$$f(|x\rangle) = (\alpha_1, \dots, \alpha_n)|x\rangle = \langle y|x\rangle \quad \text{for all} \quad |x\rangle \in \mathbb{C}^n.$$

In general, the set of linear functional on a vector space $\mathbf{V}$ ($\mathbb{C}^n$ in the present case) is called the **dual vector space**, or simply the **dual space**, of $\mathbf{V}$ and denoted by $\mathbf{V}^*$.[‡] So, we may identify $\mathbb{C}^{n*}$ with the set of bra vectors, viz.,

$$\mathbb{C}^{n*} = \{(\alpha_1, \dots, \alpha_n) : \alpha_1, \dots, \alpha_n \in \mathbb{C}\} = \{\langle y| : |y\rangle \in \mathbb{C}^n\}. \tag{1.18}$$

1.4 Eigenvalues, eigenvectors, and Schur Triangularization

Let $A \in \mathbf{M}_n$, and $|v\rangle$ be a nonzero vector in $\mathbb{C}^n$. It is usually not true that $A|v\rangle = \mu|v\rangle$, i.e., $A|v\rangle$ is not proportional to $|v\rangle$ in general. If, however, $|v\rangle$ is properly chosen, we may end up with $A|v\rangle$, which is a scalar multiple of $|v\rangle$;

$$A|v\rangle = \lambda|v\rangle, \quad \lambda \in \mathbb{C}. \tag{1.19}$$

Then λ is called an **eigenvalue** of A, while $|v\rangle$ is called the corresponding **eigenvector**. The above equation being a linear equation, the norm of the eigenvector cannot be fixed. Of course, it is always possible to normalize $|v\rangle$ such that $\||v\rangle\| = 1$. We often use the symbol $|\lambda\rangle$ for an eigenvector corresponding to an eigenvalue λ to save symbols.

Rewrite the equation (1.19) as $(A - \lambda I_n)|v\rangle = 0$. By the theory of linear equations, there is a nonzero vector $|v\rangle$ for some $\lambda \in \mathbb{C}$ if and only if $A - \lambda I_n$ is singular, equivalently,

$$D(\lambda) = \det(A - \lambda I_n) = 0. \tag{1.20}$$

[‡]The symbol $*$ here denotes the dual and should not be confused with complex conjugation.

This equation (1.20) is called the **characteristic equation** or the **eigen equation** of A. Note that $D(\lambda)$ can be written as the product of n linear factors, say,

$$D(\lambda) = \prod_{j=1}^{n}(\lambda_j - \lambda) = (-\lambda)^n + \sum_{j=1}^{n}(-\lambda)^{n-k}E_k(\lambda_1, \ldots, \lambda_n),$$

where $E_k(\lambda_1, \ldots, \lambda_n) = \sum_{1 \le j_1 < \cdots < j_k \le n} \lambda_{j_1} \cdots \lambda_{j_k}$ is known as the kth elementary symmetric function of $\lambda_1, \ldots, \lambda_n$ for $1 \le k \le n$. The characteristic equation always has n solutions $\lambda_1, \lambda_2, \ldots, \lambda_n$, counting the multiplicity.

One may expand $\det(A - \lambda I)$, say, by Laplace expansion, and compare coefficients with $\prod_{j=1}^{n}(\lambda_j - \lambda)$ and conclude that $E_k(\lambda_1, \ldots, \lambda_n)$ equals the sum of the determinants of the $k \times k$ submatrices of A lying in rows and columns indexed by $1 \le j_1 < \cdots < j_k \le n$. In particular, if $A = (a_{ij})$, then

$$\det(A) = E_n(\lambda_1, \ldots, \lambda_n) = \prod_{j=1}^{n} \lambda_j,$$

and

$$E_1(\lambda_1, \ldots, \lambda_n) = \sum_{j=1}^{n} \lambda_j = \sum_{j=1}^{n} a_{jj},$$

which is called the **trace** of A, and denoted as $\operatorname{Tr} A$.

If $A \in \mathbf{M}_n$ has n linearly independent eigenvectors $|v_1\rangle, \ldots, |v_n\rangle$, which will form a basis for $\mathbb{C}^n$, corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$, then every $|x\rangle$ in $\mathbb{C}^n$ can be written as $\sum_{j=1}^{n} c_j|v_j\rangle$ so that $A|x\rangle = \sum_{j=1}^{n} \lambda_j c_j |v_j\rangle$. However, not every $A \in \mathbf{M}_n$ has n linearly independent eigenvaectors. For example, let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Then eigenvectors of A must be a multiple of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

In general, we have the following **Schur Triangularization Theorem**.

THEOREM 1.4.1. *Let $A \in \mathbf{M}_n$. There is a unitary U such that $U^\dagger AU$ is in upper triangular form, i.e., the (r, s) entry of $U^\dagger AU$ is zero whenever $r > s$.*

Proof. By induction on n. When $n = 1$, the result trivially holds. Suppose the result holds for matrices of size at most $n - 1$. For $A \in \mathbf{M}_n$, one can solve $\det(A - \lambda I) = 0$ and get a unit vector $|x\rangle$ such that $A|x\rangle = \lambda|x\rangle$ for an eigenvalue λ. Let $U_1 \in \mathbf{M}_n$ with $|x\rangle$ as the first column. Then $U_1^\dagger AU_1 = \begin{pmatrix} \lambda & \star \\ 0 & A_1 \end{pmatrix}$. By induction assumption, there is unitary $U_2 \in \mathbf{M}_{n-1}$ such that $U_2^\dagger A_1 U_2 = T$ is in triangular form. If $U = U_1 \begin{pmatrix} 1 & \\ & U_2 \end{pmatrix}$, then $U^\dagger AU = \begin{pmatrix} \lambda & \star \\ 0 & T \end{pmatrix}$ is in triangular form. ∎

EXAMPLE 1.4.2. *Let* $A = \begin{pmatrix} 2 & 2 \\ -1 & -1 \end{pmatrix}$. *Then* $\det(A - \lambda I) = \lambda^2 - \lambda$ *so that* A *has eigenvalues* $\{0, 1\}$. *If we let* $|u_1\rangle = (1, -1)^t/\sqrt{2}$ *be a unit eigenvector for the eigenvalue 0, then* $|u_1\rangle$ *and* $|u_2\rangle = (1, 1)^t/\sqrt{2}$ *form an orthonormal basis, and* $U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ *satisfies* $U^\dagger AU = \begin{pmatrix} 0 & 3 \\ 0 & 1 \end{pmatrix}$.

1.5 Normal matrices and spectral decomposition

If $A \in \mathbf{M}_n$ has an orthonormal set of eigenvectors $\{|\lambda_1\rangle, \ldots, |\lambda_n\rangle\}$ corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$, and $U \in \mathbf{M}_n$ has columns $|\lambda_1\rangle, \ldots, |\lambda_n\rangle$, then $U^\dagger AU = D$ is a diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$. That is, the triangular matrix in Theorem 1.4.1 become a diagonal matrix. However, even if $A \in \mathbf{M}_n$ has n linearly independent eigenvectors, they may not form an orthnormal set in general.

EXAMPLE 1.5.1. *Let* $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$. *The unit eigenvectors have the form*

$$|v_1\rangle = \gamma_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad |v_2\rangle = \frac{\gamma_2}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

with $|\gamma_1| = |\gamma_2| = 1$. *So,* $|\langle v_1|v_2\rangle| = \frac{1}{\sqrt{2}} \neq 0$.

A matrix $A \in \mathbf{M}_n$ is **normal** if $AA^\dagger = A^\dagger A$. It is immediate from the definitions that Hermitian, skew-Hermitian, and unitary matrices are normal. It turns out that normal matrices are precisely the matrices with an orthonormal basis of eigenvectors as shown in the next theorem. It is interesting that a generic matrix in $\mathbf{M}_n$ is non-normal, but most matrices which are useful in quantum information science are normal matrices.

THEOREM 1.5.2. *A matrix* $A \in \mathbf{M}_n$ *is normal if and only if there are* $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ *and a unitary* $U \in \mathbf{M}_n$ *such that* $U^\dagger AU$ *is a diagonal matrix with diagonal entries* $\lambda_1, \ldots, \lambda_n$. *Equivalently,*

$$A = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^\dagger = \lambda_1|\lambda_1\rangle\langle\lambda_1| + \cdots + \lambda_n|\lambda_n\rangle\langle\lambda_n|, \tag{1.21}$$

where $|\lambda_1\rangle, \ldots, |\lambda_n\rangle$ *are the columns of* U.

Proof. Suppose $A = UDU^\dagger$ for a diagonal matrix D. Then

$$AA^\dagger = UDD^\dagger U^\dagger = UD^\dagger DU^\dagger = A^\dagger A.$$

Thus, A is normal. Conversely, suppose $AA^\dagger = A^\dagger A$, and suppose $U \in M_n$ is unitary such that $UAU^\dagger = T$ is in upper triangular form. Then

$$TT^\dagger = UAA^\dagger U^\dagger = UA^\dagger AU^\dagger = T^\dagger T.$$

For $j = 1, \ldots, n$, the $(1,1)$ entry of $TT^\dagger$ is

$$(t_{11}, \ldots, t_{1n})(t_{11}^*, \ldots, t_{1n}^*)^t = |t_{11}|^2 + \cdots + |t_{1n}|^2,$$

and the $(1,1)$ entry of $T^\dagger T$ is

$$(t_{11}^*, 0, \ldots, 0)(t_{11}, 0, \ldots, 0)^t = |t_{11}|^2.$$

So, $TT^\dagger = T^\dagger T$ implies that $t_{12} = \cdots = t_{1n} = 0$. Now, compare the $(2,2)$ entries of $TT^\dagger$ and $T^\dagger T$, we see that $t_{23} = \cdots = t_{2n} = 0$. Repeating this argument, we see that T is a diagonal matrix. ∎

As mentioned before, physicists use $|\lambda_j\rangle$ to represent a unit vector of A corresponding to the eigenvalue λ_j. Then the representation of a normal matrix $A \in \mathbf{M}_n$ in (1.21) becomes

$$A = \sum_{j=1}^{n} \lambda_j |\lambda_j\rangle\langle\lambda_j|,$$

which is called the **spectral decomposition** of A.

We have the following corollary concering Hermitian, skew-Hermitian, and unitary matrices are normal matrices. The proof is left as an exercise.

COROLLARY 1.5.3. *Let $A \in \mathbf{M}_n$ be a normal matrix.*

1. *The matrix A is Hermitian if and only if all its eigenvalues are real.*
2. *The matrix A is skew-Hermitian if and only if all its eigenvalues are pure imaginary.*
3. *The matrix A is unitary if and only if all its eigenvalues are unimodular, i.e., having modulus one.*

The **Pauli matrices**, also known as the spin matrices, are:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

They are also denoted by σ_1, σ_2 and σ_3, respectively. They are Hermitian and unitary matrices in $\mathbf{M}_2$, and are useful in quantum information science. We will use them to illustrate properties of normal matrices. Their additional properties are given in Subsection 1.8.3.

EXAMPLE 1.5.4. *The Pauli matrix*

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

is Hermitian. Let us find its eigenvalues and corresponding eigenvectors. From

$$\det(\sigma_y - \lambda I) = \lambda^2 - 1 = 0,$$

we find the eigenvalues $\lambda_1 = 1$ *and* $\lambda_2 = -1$*. For the eigenvalue* λ_1 *we solve*

$$(\sigma_y - I)|v_1\rangle = \begin{pmatrix} -1 & -i \\ i & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and get $|v_1\rangle = (x, y)^t = (1, i)^t$*. After normalization, we may let* $|\lambda_1\rangle = \frac{1}{\sqrt{2}}(1, i)^t$*. Similarly, for the eigenvalue* λ_2 *we solve* $(\sigma_y + I)|v_2\rangle = |\mathbf{0}\rangle$ *and get* $|v_2\rangle = (i, 1)^t$*. After normalization, we may let* $|\lambda_2\rangle = \frac{1}{\sqrt{2}}(i, 1)^t$*. We have*

$$\langle\lambda_1|\lambda_2\rangle = \frac{1}{2}(1, -i) \begin{pmatrix} i \\ 1 \end{pmatrix} = 0.$$

so that $\{|\lambda_1\rangle, |\lambda_2\rangle\}$ *is an orthonormal set. If*

$$P_1 = |\lambda_1\rangle\langle\lambda_1| = \frac{1}{2}\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}, \quad \text{and} \quad P_2 = |\lambda_2\rangle\langle\lambda_2| = \frac{1}{2}\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix},$$

then $\sigma_y = \sum_k \lambda_k |\lambda_k\rangle\langle\lambda_k| = P_1 - P_2$ *and* $I = P_1 + P_2$*. If* $U = (|\lambda_1\rangle, |\lambda_2\rangle) = \frac{1}{\sqrt{2}}\begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix}$*, then* U *is unitary and*

$$U^\dagger \sigma_y U = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

EXAMPLE 1.5.5. *(1) The eigenvalues and the corresponding eigenvectors of* σ_x *are found in a similar way as the above example as* $\lambda_1 = 1, \lambda_2 = -1$ *and*

$$|\lambda_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad |\lambda_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

(2) Let us consider the eigenvalue problem of a matrix

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} I_2 & \\ & \sigma_x \end{pmatrix} = I_2 \oplus \sigma_x.$$

Note that this matrix is block diagonal with diagonal blocks I_2 *and* σ_x*. Thus, the eigenvalues are those from* I_2 *and* σ_x*, i.e.,* $1, 1, 1$ *and* -1*. The corresponding eigenvectors can be extended from those of* I_2 *and* σ_x*, which has* *been obtained in (1), to get*

$$\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\ \begin{pmatrix}0\\1\\0\\0\end{pmatrix},\ \frac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\1\\1\end{pmatrix},\ \frac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\1\\-1\end{pmatrix}.$$

(3) Let us consider the eigenvalue problem of a matrix

$$B = \begin{pmatrix}0&0&0&1\\0&1&0&0\\0&0&1&0\\1&0&0&0\end{pmatrix}.$$

So, I_2 *acts on* $\mathrm{Span}\,\{|e_2\rangle, |e_3\rangle\}$ *and* σ_x *acts on* $\mathrm{Span}\,\{|e_1\rangle, |e_4\rangle\}$*. Although this matrix is not block diagonal, change of the order of basis vectors from* $|e_1\rangle, |e_2\rangle, |e_3\rangle, |e_4\rangle$ *to* $|e_3\rangle, |e_2\rangle, |e_1\rangle, |e_4\rangle$ *maps the matrix* B *to* A *in (2). Therefore the eivenvalues of* B *are the same as those of* A*. (Note that the characteristic equation is left unchanged under a permutation of basis vectors.) By putting back the order of the basis vectors, the eigenvectors of* A *are mapped to those of* B *as*

$$\begin{pmatrix}0\\0\\1\\0\end{pmatrix},\ \begin{pmatrix}0\\1\\0\\0\end{pmatrix},\ \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\0\\1\end{pmatrix},\ \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\0\\-1\end{pmatrix}.$$

By the spectral decomposition, we have the following **spectral theorem** for normal matrices.

THEOREM 1.5.6. *Suppose* $A \in \mathbf{M}_n$ *is normal satisfying* (1.21)*. For any positive integer* m*,*

$$A^m = U \begin{pmatrix}\lambda_1^m & & \\ & \ddots & \\ & & \lambda_n^m\end{pmatrix} U^\dagger = \lambda_1^m|\lambda_1\rangle\langle\lambda_1| + \cdots + \lambda_n^m|\lambda_n\rangle\langle\lambda_n|. \tag{1.22}$$

(1) If A *is invertible, then (1.22) holds for negative integers* m *as well.*

(2) If A *has nonnegative eigenvalues, then (1.22) holds for all positive real numbers* m*.*

(3) If A *has positive eigenvalues, then (1.22) holds for real numbers* m*.*

(4) We may replace the power function $f(z) = z^m$ *in (1.22) by a polynomial function* $f(z) = a_0 + a_1 z + \cdots + a_N z^N$*, or analytic functions* $f(t)$*, which admits a power series expansion, so that*

$$f(A) = \sum_{j=1}^{n} f(\lambda_j)|\lambda_j\rangle\langle\lambda_j|.$$

Proof. Since $A = UDU^\dagger$, where D is the diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$,

$$A^m = (UDU^\dagger)^m = (UDU^\dagger)\cdots(UDU^\dagger) = UD^mU^\dagger = \sum_{j=1}^{n} \lambda_j^m |\lambda_j\rangle\langle\lambda_j|.$$

One can apply a similar argument to prove (1) – (4). ∎

EXAMPLE 1.5.7. *Consider σ_y again. By Example 1.5.4, $\sigma_y = P_1 - P_2$. Hence, for $\alpha \in \mathbb{R}$,*

$$\exp(i\alpha\sigma_y) \equiv \sum_{k=0}^{\infty} \frac{(i\alpha\sigma_y)^k}{k!} = e^{i\alpha}P_1 + e^{-i\alpha}P_2 = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}.$$

Even when $f(x)$ does not admit a series expansion, we may still formally define $f(A)$ by Eq. (1.25). Let $f(x) = \sqrt{x}$ and $A = \sigma_y$, for example. Then we obtain from Example 1.5.4 that

$$\sqrt{\sigma_y} = (\pm 1)P_1 + (\pm i)P_2.$$

It is easy to show that the RHS squares to σ_y. However, there are four possible $\sqrt{\sigma_y}$ depending on the choice of $\pm$ for each eigenvalue. Therefore the spectral decomposition is not unique in this case. Of course this ambiguity originates in the choice of the branch in the definition of $\sqrt{x}$.

We prove a formula, which will be useful in our future discussion, extending Example 1.5.4 and Example 1.5.7.

PROPOSITION 1.5.8. *Let $\hat{\boldsymbol{n}} = (n_x, n_y, n_z) \in \mathbb{R}^3$ be a unit vector, $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$, and*

$$A = \hat{\boldsymbol{n}} \cdot \boldsymbol{\sigma} = \begin{pmatrix} n_z & n_x - in_y \\ n_x + in_y & -n_z \end{pmatrix}.$$

Then $\det(A - \lambda I_2) = 1 - \lambda^2$ so that A has eigenvalues $(\lambda_1, \lambda_2) = (1, -1)$, and spectral decomposition

$$A = |\lambda_1\rangle\langle\lambda_1| - |\lambda_2\rangle\langle\lambda_2|$$

with $|\lambda_1\rangle\langle\lambda_1| = (A + I)/2$ and $|\lambda_2\rangle\langle\lambda_2| = (I - A)/2$. If $\alpha \in \mathbb{R}$, then

$$\exp(i\alpha\hat{\boldsymbol{n}} \cdot \boldsymbol{\sigma}) = \cos\alpha I + i\sin\alpha(\hat{\boldsymbol{n}} \cdot \boldsymbol{\sigma}). \tag{1.23}$$

Proof. Note that $\det(A - \lambda I_2) = \lambda^2 - 1$. Hence A has eigenvalues $\lambda_1 = +1$ and $\lambda_2 = -1$. A has spectral decomposition $A = |\lambda_1\rangle\langle\lambda_1| - |\lambda_2\rangle\langle\lambda_2| = P_1 - P_2$ such that $I = P_1 + P_2$. It follows that

$$A + I = 2P_1 \quad \text{and} \quad I - A = 2P_2;$$

hence

$$P_1 = \frac{(A+I)}{2} = \frac{1}{2}\begin{pmatrix} 1+n_z & n_x - in_y \\ n_x + in_y & 1-n_z \end{pmatrix},$$
$$P_2 = \frac{(A-I)}{-2} = \frac{1}{2}\begin{pmatrix} 1-n_z & -n_x + in_y \\ -n_x - in_y & 1+n_z \end{pmatrix}.$$

As a result,

$$\begin{aligned} e^{i\alpha A} &= \frac{e^{i\alpha}}{2}\begin{pmatrix} 1+n_z & n_x - in_y \\ n_x + in_y & 1-n_z \end{pmatrix} + \frac{e^{-i\alpha}}{2}\begin{pmatrix} 1-n_z & -n_x + in_y \\ -n_x - in_y & 1+n_z \end{pmatrix} \\ &= \cos\alpha I + i\sin\alpha(\hat{\boldsymbol{n}} \cdot \boldsymbol{\sigma}). \end{aligned}$$ ∎

If $\hat{\boldsymbol{n}} = (0,1,0)$, we see that the spectral projections of σ_y are $P_1 = (\sigma_y + I)/2$ and $P_2 = (I - \sigma_y)/2$ as shown in Example 1.5.4, and $\exp(i\alpha\sigma_y) = \cos\alpha I + i\sin\alpha\sigma_y$ as shown in Example 1.5.7.

If $A \in \mathbf{M}_n$ is normal having the form (1.21), then $P_j = |\lambda_j\rangle\langle\lambda_j|$ is an eigenprojection of A corresponding to the eigenvalue λ_j. Suppose A has k distinct eigenvalues $\mu_1, \dots, \mu_k$. If we add the eigenprojections P_j corresponding to the same eigenvalues μ_ℓ to get Q_ℓ for $\ell = 1, \dots, k$. Then

$$Q_\ell^2 = Q_\ell, \quad Q_r Q_s = 0 \quad \text{whenever } r \neq s, \quad \text{and} \quad A = \sum_{\ell=1}^{k} \mu_\ell Q_\ell.$$

Theorem 1.5.6 can be stated as:

$$A^m = \sum_{\ell=1}^{k} \mu_\ell^m Q_\ell. \tag{1.24}$$

(1) If A is invertible then (1.24) holds for any integer m.

(2) If A has nonnegative eigenvalues, then (1.24) holds for any positive real number m.

(3) if A has positive eigenvalues, then (1.24) holds for any real number m.

(4) For any analytic functions $f(z)$ we have

$$f(A) = \sum_{\ell=1}^{k} f(\mu_\ell) Q_\ell. \tag{1.25}$$

1.6 Singular Value Decomposition (SVD)

Let $T : \mathbb{C}^n \to C^m$ be a linear operator of the form $T(|x\rangle) = A|x\rangle$ for all $|x\rangle \in \mathbb{C}^n$. In the following, we will show that there are orthonormal bases of

$\mathbb{C}^n$ and $\mathbb{C}^m$ such that the matrix of the linear operator T has simple form. Equivalently, there are unitary matrices $U \in \mathbf{M}_m$ and $V \in \mathbf{M}_n$ such that $B = U^\dagger AV$ has simple form. It can be viewed as a generalization of the eigenvalue problem to arbitrary matrices. The result is useful in studying quantum states in a **bipartite** quantum system, i.e., a system composed of two subsystems.

THEOREM 1.6.1. *Let $A \in \mathbf{M}_{m,n}$. Then there exist $U \in \mathrm{U}(m)$ with columns $|u_1\rangle, \ldots, |u_m\rangle$, $V \in \mathrm{U}(n)$ with columns $|v_1\rangle, \ldots, |v_n\rangle$, and a matrix $\Sigma \in \mathbf{M}_{m,n}$ with (j,j) entry equal to s_j for $j \le k \le \min\{m,n\}$ and all other entries equal to zero, such that $s_1 \ge \cdots \ge s_k > 0$ and*

$$A = U\Sigma V^\dagger = U\begin{pmatrix} D & 0_{k,n-k} \\ 0_{m-k,k} & 0_{m-k,n-k} \end{pmatrix} V^\dagger = \sum_{j=1}^{k} s_j |u_j\rangle\langle v_j|, \tag{1.26}$$

where $D \in \mathbf{M}_k$ is the diagonal matrix with diagonal entries $s_1, \ldots, s_k$.

Proof. Assume that $n = \min\{m,n\}$. By (1.10), if $T = A^\dagger A$, then $T^\dagger = (A^\dagger A)^\dagger = (A^\dagger)(A^\dagger)^\dagger = A^\dagger A = T$. So, T is Hermitian. Moreover, if λ is an eigenvalue of T, then $\lambda \in \mathbb{R}$ by Corollary 1.5.3. In fact, we have

$$\lambda = \langle\lambda|T|\lambda\rangle = \langle\lambda|A^\dagger A|\lambda\rangle = \|A|\lambda\rangle\|^2 \ge 0.$$

By Theorem 1.5.2, there is a unitary $V \in \mathbf{M}_n$ such that $V^\dagger A^\dagger AV = D$ with diagonal entries $\lambda_1 \ge \cdots \ge \lambda_n \ge 0$. Let $s_j = \sqrt{\lambda_j}$ for $j = 1, \ldots, n$. If AV has columns $|y_1\rangle, \ldots, |y_n\rangle \in \mathbb{C}^m$, then $V^\dagger A^\dagger AV = D$ implies that $\||y_j\rangle\|^2 = \lambda_j$ and $\langle y_r|y_s\rangle = 0$ for $r \ne s$. Suppose $d_k > 0 = d_{k+1}$. It is possible that $k = n$. Let $|u_j\rangle = |y_j\rangle/s_j$ for $j = 1, \ldots, k$, and extend $\{|u_1\rangle, \ldots, |u_k\rangle\}$ to an orthonormal basis.[§] One can use $s_1, \ldots, s_k$ to construct Σ stated in the theorem. Then the matrices U, V, Σ will satisfy $AV = U\Sigma$, and the desired form will follow.

If $m = \min\{m,n\}$, apply the argument to $A^\dagger$ to get the conclusion. ∎

The decomposition (1.26) is called the **singular value decomposition** and is often abbreviated as SVD. The numbers $s_1, \cdots, s_k$ are the nonzero **singular values** of A, and the matrix Σ is called the **singular value matrix**. Clearly, k is the rank of the matrix A.

Note that the proof of Theorem 1.6.1 actually provides the steps of computing U, V and Σ.

EXAMPLE 1.6.2. *Let*

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ i & i \end{pmatrix} \qquad \text{so that} \qquad A^\dagger A = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}.$$

[§] This can be done by first extending $\{|u_1\rangle, \ldots, |u_k\rangle\}$ to a basis, i.e., a linearly independent set with n vectors, and apply the Gram-Schmidt process.

The eigenvalues of $A^\dagger A$ *are* $\lambda_1 = 4$ *and* $\lambda_2 = 0$ *with the corresponding eigenvectors*

$$|\lambda_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \quad |\lambda_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\end{pmatrix}.$$

From these, we can construct unitary matrix V *and the singular value matrix* Σ *as*

$$V = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\1 & 1\end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix}2 & 0\\0 & 0\\0 & 0\end{pmatrix}.$$

To construct U*, we need*

$$|\mu_1\rangle = \frac{1}{2}A|\lambda_1\rangle = \frac{1}{\sqrt{2}}(1, 0, i)^t$$

and two other vectors orthogonal to $|\mu_1\rangle$*. By inspection, we find*

$$|\mu_2\rangle = (0, 1, 0)^t \quad \text{and} \quad |\mu_3\rangle = \frac{1}{\sqrt{2}}(i, 0, 1)^t,$$

for example. From these vectors we construct U *as*

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0 & i\\0 & \sqrt{2} & 0\\i & 0 & 1\end{pmatrix}.$$

One can verify that $U\Sigma V^\dagger$ *really reproduces* A*.*

1.7 Tensor Product (Kronecker Product)

In this section, we present the definition and some results on the **tensor product** (**Kronecker product)** of two matrices of any sizes. These are extremely important in the study of quantum systems.

Let A be an $m \times n$ matrix and let B be a $p \times q$ matrix. Then

$$A \otimes B = \begin{pmatrix} a_{11}B, a_{12}B, \ldots, a_{1n}B \\ a_{21}B, a_{22}B, \ldots, a_{2n}B \\ \ldots \\ a_{m1}B, a_{m2}B, \ldots, a_{mn}B \end{pmatrix} \tag{1.27}$$

is an $(mp) \times (nq)$ matrix called the **tensor product** (**Kronecker product)** of A and B.

It should be noted that not all $(mp) \times (nq)$ matrices are tensor products of an $m \times n$ matrix and a $p \times q$ matrix. It is easy to see that for

$$T = \begin{pmatrix} T_{11} & \cdots & T_{1n} \\ \vdots & \ddots & \vdots \\ T_{m1} & \cdots & T_{mn} \end{pmatrix} \quad \text{with } T_{rs} \in \mathbf{M}_{p,q}, \tag{1.28}$$

$T = A \otimes B$ with $A \in \mathbf{M}_{m,n}, B \in \mathbf{M}_{p,q}$ if and only if T_{rs} are multiple of each others for all r, s.

In general, an $(mp) \times (np)$ matrix has $mnpq$ degrees of freedom, while $m \times n$ and $p \times q$ matrices have $mn + pq$ in total. Observe that $mnpq \gg mn + pq$ for large enough m, n, p and q. This fact is ultimately related to the power of quantum computing compared to its classical counterpart.

EXAMPLE 1.7.1.

$$\sigma_x \otimes \sigma_z = \begin{pmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

EXAMPLE 1.7.2. *We can also apply the tensor product to vectors as a special case. Let*

$$|u\rangle = \begin{pmatrix} a \\ b \end{pmatrix}, \quad |v\rangle = \begin{pmatrix} c \\ d \end{pmatrix}.$$

Then we obtain

$$|u\rangle \otimes |v\rangle = \begin{pmatrix} a|v\rangle \\ b|v\rangle \end{pmatrix} = \begin{pmatrix} ac \\ ad \\ bc \\ bd \end{pmatrix}.$$

The tensor product $|u\rangle \otimes |v\rangle$ is often abbreviated as $|u\rangle|v\rangle$ or $|uv\rangle$ when it does not cause confusion.

From the definition, one readily checks that

(1) $A \otimes (B + C) = A \otimes B + A \otimes C$ if B and C have the same size, and

(2) $(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger$.

One can also show that

(3) For matrices A, B, C, D such that AC and BD are defined,

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD).$$

To see this, assume that $A = (a_{pq}), C = (c_{rs})$ and $AC = (t_{uv}) \in \mathbf{M}_{mn}$. By block multiplication,

$$(A \otimes B)(C \otimes D) = (a_{pq}B)(c_{rs}D) = (t_{uv}BD) = (AC) \otimes (BD).$$

Similarly, we have

$$(A_1 \otimes B_1)(A_2 \otimes B_2)(A_3 \otimes B_3) = (A_1 A_2 A_3) \otimes (B_1 B_2 B_3),$$

and its generalizations whenever the dimensions of the matrices match so that the products make sense.

By (1), (2), (3), we can deduce the following.

(4) Suppose $A \in \mathbf{M}_m$ and $B \in \mathbf{M}_n$.

(4.a) If $A|u\rangle = \lambda|u\rangle$ and $B|v\rangle = \mu|v\rangle$ for nonzero vectors $|u\rangle, |v\rangle$, then $(A \otimes B)|uv\rangle = (\lambda\mu)|uv\rangle$. That is, $\lambda\mu$ is an eigenvalue of $A \otimes B$ with eigenvector $|uv\rangle = |u\rangle \otimes |v\rangle$.

(4.b) If A and B are invertible, then $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.

The assertion (4.a) can be verified readily.
For (4.b), note that $(A \otimes B)(A^{-1} \otimes B^{-1}) = I_m \otimes I_n = I_{mn}$.

By the above properties one can check that the tensor product of two unitary matrices is also unitary, and the tensor product of two Hermitian matrices is also Hermitian.[¶] Also, we have the following.

(5) For any matrices A, B, if $R_1 A S_1 = T_1, R_2 B S_2 = T_2$, then

$$(R_1 \otimes R_2)(A \otimes B)(S_1 \otimes S_2) = T_1 \otimes T_2.$$

(5.a) Suppose $A \in \mathbf{M}_m$ and $B \in \mathbf{M}_n$, and R_1, S_1, R_2, S_2 are unitary such that $R_1 = S_1^\dagger$ and $R_2 = S_2^\dagger$. If T_1, T_2 are in triangular (diagonal) form, then $U = S_1 \otimes S_2$ is unitary such that

$$U^\dagger (A \otimes B) U = T_1 \otimes T_2$$

is in triangular (diagonal) form.

As a result, if $A \in \mathbf{M}_m$ and $B \in \mathbf{M}_n$ are normal with spectral decomposition $A = \sum_{r=1}^m \lambda_r |\lambda_r\rangle\langle\lambda_r|$ and $B = \sum_{s=1}^n \mu_s |\mu_s\rangle\langle\mu_s|$, then $A \otimes B$ is normal with spectral decomposition $A \otimes B = \sum_{r,s} \lambda_r \mu_s |\lambda_r \mu_s\rangle\langle\lambda_r \mu_s|$.

(5.b) If A, B are rectangular matrices with singular decomposition

$$A = \sum_{i=1}^r a_i |u_i\rangle\langle v_i| \quad \text{and} \quad B = \sum_{j=1}^s b_j |x_j\rangle\langle y_j|,$$

then

$$A \otimes B = \sum_{r,s} a_i b_j |u_i x_j\rangle\langle v_i y_j|$$

is the singular value decomposition of $A \otimes B$.

[¶]Note that the usual product of two Hermitian matrices may not be Hermitian.

EXAMPLE 1.7.3. *Let* $U^\dagger AU = V^\dagger BV = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$*, where* $U, V \in \mathbf{M}_2$ *are unitary matrices with columns* $|u_1\rangle, |u_2\rangle$ *and* $|v_1\rangle, |v_2\rangle$*, respectively. Then*

$$(U \otimes V)^\dagger (A \otimes B)(U \otimes V) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus, $|u_1\rangle \otimes |v_1\rangle, |u_1\rangle \otimes |v_2\rangle, |v_2\rangle \otimes |v_1\rangle$ *are eigegenvectors of* $A \otimes B$ *corresponding to the eigenvalue 0. Note that one can only get the eigenvector* $|u_1\rangle \otimes |v_1\rangle$ *for the eigenvalue 0 using* 4(a).

1.8 Additional topics

1.8.1 Factorization and Norm Properties of matrices

As mentioned before, by the Gram-Schmidt process, we have the following QR **factorization** of matrices.

PROPOSITION 1.8.1. *Let* $A \in \mathbf{M}_{m,n}$ *have linearly independent columns. Then* $A = QR$*, where* $Q \in \mathbf{M}_m$ *is unitary, and* $R \in \mathbf{M}_{m,n}$ *with zero* (r, s) *entry whenever* $s > r$*.*

Proof. Since A has linearly independent columns, we have $m \geq n$. Applying the Gram-Schmidt process to the columns of A, we get orthonormal vectors $|u_1\rangle, \ldots, |u_n\rangle$. Extend these vector to an orthonormal basis and use them as the columns of Q, and let $R = Q^\dagger Q$. Then $A = QR$ as asserted. ∎

Using the singular value decomposition, we have the **polar decomposition** of square matrices.

PROPOSITION 1.8.2. *Let* $A \in \mathbf{M}_n$*. Then there are unitary* $U \in \mathbf{M}_n$ *and positive semi-definite matrices* $P, Q \in \mathbf{M}_n$ *such that* $A = UP = QU$*.*

Proof. Suppose A has a singular value decomposition $A = X\Sigma Y$. Let $U = XY, P = Y^\dagger \Sigma Y, Q = X\Sigma X^\dagger$. The conclusion holds. ∎

The definitions of inner product can be extended to $\mathbf{M}_{m,n}$ defined by $(A, B) = \mathrm{Tr}\,(A^* B)$[‖] One readily verifies that for any $A, B, C \in \mathbf{M}_{m,n}$.

(1) $(A + B, C) = (A, C) + (B, C),\ (A, \mu B) = \mu(A, B),\ (A, B) = (B, A)^*$,

[‖] In mathematics literature, it is defined by $(A, B) = \mathrm{Tr}\,(AB^*)$ so that the function is linear in the first component.

(2) $(A, A) \geq 0$, where the equality holds if and only if $A = 0_{m,n}$.

One can define the corresponding inner product norm by

$$\|A\| = (A, A)^{1/2} \qquad \text{for any } A \in \mathbf{M}_{m,n},$$

which satisfies the following norm properties for any $A, B \in \mathbf{M}_{m,n}$ and $c \in \mathbb{C}$.

(a) $\|A\| \geq 0$, where the equality holds if and only if $A = 0_{m,n}$.

(b) $\|A + B\| \leq \|A\| + \|B\|$.

(c) $\|cA\| = |c|\|A\|$.

To verify (b), we need the following. **Cauchy-Schwartz inequality**.

PROPOSITION 1.8.3. *Let $A, B \in \mathbf{M}_{m,n}$. Then $|(A, B)|^2 \leq (A, A)(B, B)$, where the equality holds if and only if $\{A, B\}$ is linearly dependent.*

Proof. Let $a = (A, A), b = (B, B)$, and $c = |(A, B)| = e^{i\alpha}(A, B)$, with $\alpha \in [0, 2\pi)$. Then for any $t \in \mathbb{R}$,

$$0 \leq (tA + e^{i\alpha}B, tA + e^{i\alpha}B) = at^2 + 2ct + b.$$

By the theory of quadratic equation, $4|(A, B)|^2 = 4c^2 \leq 4ab = 4(A, A)(B, B)$, where the equality holds if and only if $\|tA - e^{i\alpha}B\| = 0$ with $t \in \mathbb{R}$. The conclusion follows. ∎

Clearly, the above proof works for $\mathbb{C}^n = \mathbf{M}_{n,1}$. In fact, the same proof works for a general inner product space with an inner product $(\cdot, \cdot)$ satisfying (1) and (2).

Besides the inner product norm, one can define other norms on $\mathbb{C}^n$ and $\mathbf{M}_{m,n}$. For instance, for $1 \leq p$, one can define the ℓ_p-norm on $\mathbb{C}^n$ by

$$\ell_p(|x\rangle) = (\sum_{j=1}^{n} |x_j|^p)^{1/p} \quad \text{for } |x\rangle = (x_1, \dots, x_n)^t$$

and the Schatten p-norm of $A \in \mathbf{M}_{m,n}$ by

$$S_p(A) = (\sum_{j=1}^{p} s_j(A)^p)^{1/p} \quad \text{for } A \in \mathbf{M}_{m,n},$$

where $s_1(A) \geq s_2(A) \geq \cdots$ are the singular values of A. Taking limit $p \to \infty$, we have $\ell_\infty(x) = \max\{|x_j| : 1 \leq j \leq n\}$ and $S_\infty(A) = s_1(A)$.

1.8.2 Construction of eigenprojections without eigenvectors

Let us recall that $P_i = |\lambda_i\rangle\langle\lambda_i|$ is a projection operator onto the direction of the eigenvector $|\lambda_i\rangle$ of a normal matrix A. Then the spectral decomposition claims that the operation of A in the one-dimensional subspace spanned by $|\lambda_i\rangle$ is equivalent with a multiplication by a scalar λ_i. This observation reveals a neat way to obtain the spectral decomposition of a normal matrix. Let A be a normal matrix and let $\{\lambda_\alpha\}$ and $\{|\lambda_{\alpha,p}\rangle\ (1 \le p \le g_\alpha)\}$ be the sets of eigenvalues and eigenvectors, respectively. Here we use subscripts $\alpha, \beta, \ldots$ to denote distinct eigenvalues, while g_α denotes the degeneracy (multiplicity) of the eigenvalue λ_α, namely λ_α has g_α linearly independent eigenvectors, which are indexed by p. Therefore we have

$$\sum_\alpha 1 \le n, \quad \sum_\alpha g_\alpha = \sum_i 1 = n.$$

Now consider the following expression:

$$P_\alpha = \frac{\prod_{\beta\neq\alpha}(A - \lambda_\beta I)}{\prod_{\gamma\neq\alpha}(\lambda_\alpha - \lambda_\gamma)}. \tag{1.29}$$

This is a projection operator onto the g_α-dimensional space corresponding to the eigenvalue λ_α. In fact, it is straightforward to verify that

$$P_\alpha|\lambda_{\alpha,p}\rangle = \frac{\prod_{\beta\neq\alpha}(\lambda_\alpha - \lambda_\beta)}{\prod_{\gamma\neq\alpha}(\lambda_\alpha - \lambda_\gamma)}|\lambda_{\alpha,p}\rangle = |\lambda_{\alpha,p}\rangle \quad (1 \le p \le g_\alpha)$$

and

$$P_\alpha|\lambda_{\delta,q}\rangle = \frac{\prod_{\beta\neq\alpha}(\lambda_\delta - \lambda_\beta)}{\prod_{\gamma\neq\alpha}(\lambda_\alpha - \lambda_\gamma)}|\lambda_{\delta,q}\rangle = 0 \quad (\delta \neq \alpha,\ 1 \le q \le g_\delta)$$

since one of $\beta(\neq \alpha)$ is equal to $\delta(\neq \alpha)$ in the numerator. Therefore, we conclude that P_α is a projection operator

$$P_\alpha = \sum_{p=1}^{g_\alpha} |\lambda_{\alpha,p}\rangle\langle\lambda_{\alpha,p}| \tag{1.30}$$

onto the g_α-dimensional subspace corresponding to the eigenvalue λ_α. It follows from Eq. (1.30) that rank $P_\alpha = g_\alpha$. Note also that

$$AP_\alpha = \lambda_\alpha P_\alpha. \tag{1.31}$$

The above method is particularly suitable when the eigenvalues are degenerate. It is also useful when eigenvectors are difficult to obtain or unnecessary.

Using this method, one can again deduce that the projection operators of σ_y are $P_1 = (I + \sigma_y)/2$ and $P_2 = (I - \sigma_y)/2$ as shown in Example 1.5.4.

1.8.3 Pauli Matrices

Let us consider spin 1/2 particles, such as an electron or a proton. These particles have an internal degree of freedom: the spin-up and spin-down states. (To be more precise, these are expressions that are relevant when the z-component of an angular momentum S_z is diagonalized. If S_x is diagonalized, for example, these two quantum states can be either "spin-right" or "spin-left.") Since the spin-up and spin-down states are orthogonal, we can take their components to be

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{1.32}$$

Then they are eigenvectors of σ_z satisfying $\sigma_z|\uparrow\rangle = |\uparrow\rangle$ and $\sigma_z|\downarrow\rangle = -|\downarrow\rangle$. In quantum information, we often use the notations $|0\rangle = |\uparrow\rangle$ and $|1\rangle = |\downarrow\rangle$. Moreover, the states $|0\rangle$ and $|1\rangle$ are not necessarily associated with spins. They may represent any two mutually orthogonal states, such as horizontally and vertically polarized photons. Thus we are free from any physical system, even though the terminology of spin algebra may be employed.

For electrons and protons, the spin angular momentum operator is conveniently expressed in terms of the Pauli matrices σ_k as $S_k = (\hbar/2)\sigma_k$. We often employ natural units in which $\hbar = 1$. Note the tracelessness property $\operatorname{Tr}\sigma_k = 0$ and the Hermiticity $\sigma_k^\dagger = \sigma_k$.** In addition to the Pauli matrices, we introduce the unit matrix I_2 in the algebra, which amounts to expanding the Lie algebra $\mathfrak{su}(2)$ to $\mathfrak{u}(2)$.

Let $A, B \in \mathbf{M}_n$. Their **anticommutator**, or **anticommutation relation**, is $\{A, B\} \equiv AB + BA$; their **commutator**, or **commutation relation**, is $[A, B] \equiv AB - BA$.

The Pauli matrices satisfy the anticommutation relations

$$\{\sigma_j, \sigma_k\} = \sigma_j\sigma_k + \sigma_k\sigma_j = 2\delta_{jk} I. \tag{1.33}$$

For $\ell = 1, 2, 3$, the eigenvalues of σ_ℓ are ± 1.

The commutation relations between the Pauli matrices are

$$[\sigma_j, \sigma_k] = \sigma_j\sigma_k - \sigma_k\sigma_j = 2i \sum_\ell \varepsilon_{jk\ell}\sigma_\ell, \tag{1.34}$$

where $\varepsilon_{jk\ell}$ is the totally antisymmetric tensor of rank 3, also known as the **Levi-Civita symbol**,

$$\varepsilon_{jk\ell} = \begin{cases} 1, & (j,k,\ell) = (1,2,3), (2,3,1), (3,1,2) \\ -1 & (j,k,\ell) = (2,1,3), (1,3,2), (3,2,1) \\ 0 & \text{otherwise.} \end{cases}$$

**Mathematically speaking, these two properties imply that $i\sigma_k$ are generators of the $\mathfrak{su}(2)$ Lie algebra associated with the Lie group SU(2).

The commutation relations, together with the anticommutation relations, yield

$$\sigma_j\sigma_k = i\sum_{\ell=1}^{3}\varepsilon_{jk\ell}\sigma_\ell + \delta_{jk}I\,. \tag{1.35}$$

The spin-flip ("ladder") operators are defined by

$$\sigma_+ = \frac{1}{2}(\sigma_x + i\sigma_y) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \sigma_- = \frac{1}{2}(\sigma_x - i\sigma_y) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{1.36}$$

Verify that $\sigma_+|{\uparrow}\rangle = \sigma_-|{\downarrow}\rangle = 0, \quad \sigma_+|{\downarrow}\rangle = |{\uparrow}\rangle, \quad \sigma_-|{\uparrow}\rangle = |{\downarrow}\rangle$. The projection operators to the eigenspaces of σ_z with the eigenvalues ± 1 are

$$\begin{aligned} P_+ &= |\uparrow\rangle\langle\uparrow| = \tfrac{1}{2}(I + \sigma_z) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \\ P_- &= |\downarrow\rangle\langle\downarrow| = \tfrac{1}{2}(I - \sigma_z) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{1.37}$$

In fact, it is straightforward to show

$$P_+|{\uparrow}\rangle = |{\uparrow}\rangle,\;\; P_+|{\downarrow}\rangle = 0,\;\; P_-|{\uparrow}\rangle = 0,\;\; P_-|{\downarrow}\rangle = |{\downarrow}\rangle\;\; .$$

Finally, we note the following identities:

$$\sigma_\pm^2 = 0, \quad P_\pm^2 = P_\pm, \quad P_+P_- = 0. \tag{1.38}$$

1.9 Notes and Open problems

In this chapter, we presented a short review of linear algebra needed for our discussion. A common question for beginners is: Why do we need to use complex vectors and matrices? The short answer is: Only complex vectors and matrices can model quantum systems satisfactorily. In fact, "Matrix Mechanics" was a formulation of quantum mechanics introduced by Werner Heisenberg, Max Born, and Pascual Jordan (1925). John von Neumann formalized the mathematical framework, and used the Hilbert space approach to understand some basic quantum phenomena; see [7].

Even before going deep into the applications of linear algebra in quantum information science, we can list some open problems in matrix theory related to quantum information science that one may attempt.

1. **Mutually unbiased bases (MUB).** Determine the maximum number r for the existence of unitary matrices $U_0 = I_n, U_1, \ldots, U_r \in \mathbf{M}_n$ such that every entry of $U_j^*U_k$ has modulus $1/\sqrt{n}$.

It is known that $r \leq n+1$, and the equality holds if n is a prime power. The problem is open for $n = 6$.

One may see [1] and its references for more background and results.

2. **Orthonormal basis for symmetric matrices.** Construct or show the existence of symmetric unitary matrices $U_1, \ldots, U_N \in M_n$ with $N = n(n+1)/2$ such that $\operatorname{Tr}(U_j^\dagger U_k) = 0$ for all $j \neq k$.

   The construction of the cases when n is even or $n = 3$ is known. There are numerical evidence that the existence of $A_1, \ldots, A_N$ if $n = 5, 7$. There is no general proof for the construction.

   One can ask a similar question for the space of skew-symmetric matrices $A \in \mathbf{M}_n$, i.e., $A = -A^t$. In such a case, one would like find unitary $V_1, \ldots, V_N$ with $N = n(n-1)/2$ such that $\operatorname{Tr}(U_j^\dagger U_k) = 0$ for all $j \neq k$. Such a construction is known if n is even, and it is known that the construction is impossible if n is odd.

   This question is related to the operator sum representation of the **Werner-Holevo channel**. One may see [3] for more background.

Exercises for Chapter 1

EXERCISE 1.1. *Find the condition under which two vectors*

$$|v_1\rangle = \begin{pmatrix} x \\ y \\ 3 \end{pmatrix}, \ |v_2\rangle = \begin{pmatrix} 2i \\ x-y \\ 1 \end{pmatrix} \in \mathbb{C}^3$$

are linearly independent.

EXERCISE 1.2. *Show that a set of vectors*

$$|v_1\rangle = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad |v_2\rangle = \begin{pmatrix} 1 \\ 0 \\ i \end{pmatrix}, \quad |v_3\rangle = \begin{pmatrix} 1 \\ -1 \\ -1+i \end{pmatrix}$$

is a basis of $\mathbb{C}^3$.

EXERCISE 1.3. *Let*

$$|x\rangle = \begin{pmatrix} 1 \\ i \\ 2+i \end{pmatrix}, \quad |y\rangle = \begin{pmatrix} 2-i \\ 1 \\ 2+i \end{pmatrix}.$$

Find $\||x\rangle\|, \langle x|y\rangle$ *and* $\langle y|x\rangle$.

EXERCISE 1.4. *Prove that*

$$\langle x|y\rangle = \langle y|x\rangle^*. \tag{1.39}$$

EXERCISE 1.5. *Let* $\{|e_k\rangle\}$ *be as in Example 1.3.3 and let*

$$|v\rangle = \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \sum c_k |e_k\rangle.$$

Find the coefficients c_1 *and* c_2*.*

EXERCISE 1.6. *(1) Use the Gram-Schmidt process to find an orthonormal basis* $\{|e_k\rangle\}$ *from a linearly independent set of vectors*

$$|v_1\rangle = (-1, 2, 2)^t, \quad |v_2\rangle = (2, -1, 2)^t, \quad |v_3\rangle = (3, 0, -3)^t.$$

(2) Let

$$|u\rangle = (1, -2, 7)^t = \sum_k c_k |e_k\rangle.$$

Find the coefficients c_k*.*

EXERCISE 1.7. *Let*

$$|v_1\rangle = (1, i, 1)^t, \quad |v_2\rangle = (3, 1, i)^t.$$

Find an orthonormal basis for $\mathrm{Span}\{|v_1\rangle, |v_2\rangle\}$*.*

EXERCISE 1.8. *Show that the Gram-Schmidt process on a linearly independent set* $\{|v_1\rangle, \ldots, |v_k\rangle\}$ *using the projection operators as follows.*

Let $|e_1\rangle = |v_1\rangle/\||v_1\rangle\|$*. For* $j = 2, \ldots, k$, $|e_j\rangle = |f_j\rangle/\||f_j\rangle\|$*, where*

$$|f_j\rangle = |v_j\rangle - P_1|v_j\rangle - \cdots - P_{j-1}|v_j\rangle,$$

where $P_\ell = |e_\ell\rangle\langle e_\ell|$ *for* $\ell = 1, \ldots, j-1$*.*

EXERCISE 1.9. *Let* A *and* B *be* $n \times n$ *matrices and* $c \in \mathbb{C}$*. Show that*

$$(cA)^\dagger = c^* A^\dagger, \quad (A+B)^\dagger = A^\dagger + B^\dagger, \quad (AB)^\dagger = B^\dagger A^\dagger. \tag{1.40}$$

EXERCISE 1.10. *Let*

$$A = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1+i \\ 1-i & 0 \end{pmatrix}.$$

Find the eigenvalues and the corresponding normalized eigenvectors. Show that the eigenvectors are mutually orthogonal and that they satisfy the completeness relation. Find a unitary matrix which diagonalizes A*.*

EXERCISE 1.11. *Prove Corollary 1.5.3.*

EXERCISE 1.12. *A matrix $A \in \mathbf{M}_n$ is called positive-semidefinite if $\langle\psi|A|\psi\rangle \geq 0$ for any $|\psi\rangle$ in the relevant Hilbert space $\mathcal{H}$. Show that $A \in \mathbf{M}_n$ is positive semi-definite if and only if it is Hermian with non-negative eigenvalues.*

EXERCISE 1.13. *Show that*

$$U = \begin{pmatrix} 0 & 0 & i \\ 0 & i & 0 \\ i & 0 & 0 \end{pmatrix}.$$

is unitary, and find the eigenvalues (without calculation if possible) and the corresponding eigenvectors.

EXERCISE 1.14. *Let $H \in \mathbf{M}_n$ be a Hermitian matrix. Show that*

$$U = (I + iH)(I - iH)^{-1}$$

is unitary. (This transformation is called the **Cayley transformation***.) Show that if H has eigenvalues λ_j for $j = 1, \ldots, n$, then U has eigenvalues $\frac{1+i\lambda_j}{1-i\lambda_j}$ for $j = 1, \ldots, n$.*

EXERCISE 1.15. *In Example 1.4.2, show that there is a unitary matrix V such that $V^\dagger AV$ is in upper triangular form with $(1,1)$ entry equal to 1.*

EXERCISE 1.16. *Suppose $A \in \mathbf{M}_2$ has eigenvalues $-1, 3$ and the corresponding eigenvectors*

$$|e_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ i \end{pmatrix}, \ \ |e_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix},$$

respectively. Find A.

EXERCISE 1.17. *Let*

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

(1) Find the eigenvalues and the corresponding normalized eigenvectors of A.
(2) Write down the spectral decomposition of A.
(3) Find $\exp(i\alpha A)$.

EXERCISE 1.18. *Let*

$$A = \begin{pmatrix} 1 & i & -1 \\ -i & 1 & -i \\ -1 & i & 1 \end{pmatrix}.$$

(1) Find the eigenvalues and the corresponding eigenvectors of A.
(2) Find the spectral decomposition of A.
(3) Find the inverse of A by making use of the spectral decomposition.

EXERCISE 1.19. *Let $f : \mathbb{C} \to \mathbb{C}$ be an analytic function. Let $\hat{\boldsymbol{n}}$ be a real three-dimensional unit vector and α be a real number. Show that*

$$f(\alpha\hat{\boldsymbol{n}} \cdot \boldsymbol{\sigma}) = \frac{f(\alpha) + f(-\alpha)}{2} I + \frac{f(\alpha) - f(-\alpha)}{2} \hat{\boldsymbol{n}} \cdot \boldsymbol{\sigma}. \tag{1.41}$$

(c.f., Proposition 1.5.8.)

EXERCISE 1.20. *Find the SVD of*

$$A = \begin{pmatrix} 1 & 0 & i \\ i & 0 & 1 \end{pmatrix}.$$

EXERCISE 1.21. *Let A and B be an $m \times m$ matrix and a $p \times p$ matrix, respectively. Show that*

$$\mathrm{tr}(A \otimes B) = (\mathrm{tr}A)(\mathrm{tr}B) \quad \text{and} \quad \det(A \otimes B) = (\det A)^p (\det B)^m.$$

EXERCISE 1.22. *Let $A \in \mathbf{M}_m$ and $B \in \mathbf{M}_p$ with eigenvectors $|u_1\rangle, \ldots, |u_n\rangle$ and $|v_1\rangle, \ldots, |v_p\rangle$ corresponding to eigenvalues $\lambda_1, \ldots, \lambda_m$ and $\mu_1, \ldots, \mu_p$. Show that $A \otimes I_p + I_m \otimes B$ has the eigenvalues $\{\lambda_j + \mu_k\}$ with the corresponding eigenvectors $\{|u_j v_k\rangle\}$, where I_p is the $p \times p$ unit matrix.*

References

[1] I. Bengtsson, Three Ways to Look at Mutually Unbiased Bases, AIP Conference Proceedings 889, 40–51 (2007).

[2] R. Bhatia, *Matrix Analysis*, Springer (1997).

[3] M. Girard, D. Leung, J. Levick, C.K. Li, V. Paulsen, Y.T. Poon, and John Watrous1, On the mixed-unitary rank of quantum channels, https://arxiv.org/pdf/2003.14405.pdf

[4] Otfried Gühne, Geza Toth, Entanglement detection, Physics Reports 474, 1 (2009).

[5] R.A. Horn and R.C. Johnson, *Matrix Analysis* (2nd ed.), Cambridge University Press (2012).

[6] Peter D. Lax, Linear Algebra and Its Applications, Wiley-Interscience (2007).

[7] J. von Neumann, Mathematical Foundations of Quantum Mechanics, Princeton University Press, Princeton, 1996.