Quantum mechanics describes nature in the language of linear algebra. In this lecture, we will review concepts from linear algebra that will be essential for our study of quantum entanglement. Our focus here is on finite-dimensional vector spaces since all physical systems that we consider in this course are finite-dimensional.

Contents

1 Hilbert Space 1
   1.1 Vector Spaces, Kets and Bras .................................................. 1
   1.2 Computational Basis and Matrix Representations .......................... 4
   1.3 Direct Sum Decompositions of Hilbert Space ............................... 5

2 Linear Operators 6
   2.1 Outer Products, Adjoints, and Eigenvalues ................................ 6
   2.2 Important Types of Operators .................................................. 9
      2.2.1 Projections ................................................................. 9
      2.2.2 Normal Operators ....................................................... 10
      2.2.3 Unitary Operators ....................................................... 12
      2.2.4 Hermitian Operators .................................................... 13
      2.2.5 Positive Operators ..................................................... 13
   2.3 Functions of Normal Operators ............................................. 13
   2.4 The Polar and Singular Value Decompositions ......................... 14

3 Tensor Products 15
   3.1 Tensor Product Spaces and Vectors ......................................... 15

4 Exercises 19

1 Hilbert Space

1.1 Vector Spaces, Kets and Bras

In 1939, P.M. Dirac published a short paper whose opening lines express the adage that “good notation can be of great value in helping the development of a theory” [Dir39]. In that paper, Dirac introduced notation for linear algebra that is commonly referred today as “Dirac” or “bra-ket” notation.
To set the stage for explaining Dirac notation, let us assume that we are dealing with a \(d\)-dimensional complex vector space \(V\). Now \(V\) could be a set of anything - numbers, matrices, chairs, polarizations of a photon, etc. - so long as it satisfies the properties of a vector space. But regardless of what these objects are, we can represent them by kets, which are symbols of the form \(|\psi\rangle\). Thus on an abstract level, we can treat \(V\) as just being a collection of kets \({|\psi\rangle, |\varphi\rangle, |\alpha\rangle, |\beta\rangle, \ldots\}\). Since it is a vector space, that means that if \(|\psi\rangle\) and \(|\varphi\rangle\) are two element of \(V\), then so is any complex linear combination
\[
a|\psi\rangle + b|\varphi\rangle \in V,
\]
where \(a, b \in \mathbb{C}\).

For a vector space \(V\), a collection of vectors \(|\alpha\rangle, |\beta\rangle, \ldots\rangle\) is called linearly independent if \(0 = \sum_{i=1}^{r} c_i |\alpha_i\rangle\) implies that \(c_i = 0\) for all \(i\). In other words, the zero vector cannot be written as a nontrivial linear combination of elements in the set; otherwise the set is called linearly dependent. The vector space is finite-dimensional with dimension \(d\) if every element in \(V\) can be written as linear combination of vectors belonging to a linear independent set of \(d\) vectors. Such a set is called a basis for \(V\). In this course, we will be dealing exclusively with finite-dimensional vector spaces.

We will further assume that an inner product is defined on \(V\). Recall that an inner product is a complex valued function \(\omega : V \times V \to \mathbb{C}\) that satisfies the three properties for all \(|\alpha\rangle, |\beta\rangle, |\gamma\rangle \in V\)

1. Conjugate symmetry: \(\omega(|\alpha\rangle, |\beta\rangle) = \omega(|\beta\rangle, |\alpha\rangle)^*\);
2. Linearity in the second argument: \(\omega(|\gamma\rangle, a|\alpha\rangle + b|\beta\rangle) = a\omega(|\gamma\rangle, |\alpha\rangle) + b\omega(|\gamma\rangle, |\beta\rangle)\);
3. Positive-definiteness: \(\omega(|\alpha\rangle, |\alpha\rangle) \geq 0\) with equality holding if and only if \(|\alpha\rangle = 0\).

It is a simple exercise to show that properties one and two can be combined to show complex linearity in the first argument. That is,
\[
\omega(a|\alpha\rangle + b|\beta\rangle, |\gamma\rangle) = a^*\omega(|\alpha\rangle, |\gamma\rangle) + b^*\omega(|\beta\rangle, |\gamma\rangle).
\]

Two vectors \(|\alpha\rangle\) and \(|\beta\rangle\) are called orthogonal with respect to the inner product if \(\omega(|\alpha\rangle, |\beta\rangle) = 0\). As we will see in this course, the concept of orthogonality corresponds to certain physical properties in quantum information processing.

The second property above says that any inner product \(\omega : V \times V \to \mathbb{C}\) can be converted into a linear function by fixing a vector in the first argument of \(\omega\). That is, for any fixed \(|\gamma\rangle \in V\) and inner product \(\omega\), define the linear function \(\omega(|\gamma\rangle, \cdot) : V \to \mathbb{C}\) that maps any ket \(|\alpha\rangle\) to its inner product with \(|\gamma\rangle\). Functions like \(\omega(|\gamma\rangle, \cdot)\) are used so frequently in quantum mechanics that we denote them simply as \(\langle \gamma |\cdot \rangle\), and their action on an arbitrary \(|\alpha\rangle \in V\) is written as
\[
\langle \gamma |\alpha \rangle := \langle \gamma |(|\alpha\rangle) = \omega(|\gamma\rangle, |\alpha\rangle).
\]

The function \(\langle \gamma |\cdot \rangle\) is called the bra or dual vector of the ket \(|\gamma\rangle\). Thus, we think of bras as linear functions that act on kets and map them to set of complex numbers. The action of a bra on a ket

\(^1\)A complex vector space is a set \(V\) that is closed under two operations, vector addition \(+ : V \times V \to V\) and scalar multiplication \(\mathbb{C} \times V \to V\), and such that (i) vector addition is associative and commutative, (ii) there exists a zero vector \(0 \in V\) such that \(|\alpha\rangle + 0 = |\alpha\rangle\) for any \(|\alpha\rangle \in V\), (iii) there exists an additive inverse vector \(-|\alpha\rangle \in V\) for any \(|\alpha\rangle \in V\) such that \(|\alpha\rangle + (-|\alpha\rangle) = 0\), (iv) for any \(|\alpha\rangle, |\beta\rangle \in V\) and \(a, b \in \mathbb{C}\) it holds that \(a(|b|\alpha\rangle) = (ab)|\alpha\rangle\), \(1|\alpha\rangle = |\alpha\rangle\), \(a(|\alpha\rangle + |\beta\rangle) = a|\alpha\rangle + b|\beta\rangle\), and \((a + b)|\alpha\rangle = a|\alpha\rangle + b|\alpha\rangle\).
is sometimes called a **contraction** because it maps some \( d \)-dimensional vector space \((V)\) down to a 0-dimensional vector space \((C)\).

The collection of all bras acting on \( V \) itself forms a vector space called the **dual space** of \( V \). If \( \langle \alpha \rangle \) and \( \langle \beta \rangle \) are two bras, then any linear combination \( a\langle \alpha \rangle + b\langle \beta \rangle \) forms a new bra whose action on kets is defined by

\[
(a\langle \alpha \rangle + b\langle \beta \rangle) \langle \gamma \rangle := a\langle \alpha \gamma \rangle + b\langle \beta \gamma \rangle.
\]

This is the bra generated by a linear combination of other bras, but what about the bra corresponding to a linear combination of kets? Equation (1) shows how to obtain determine the latter. For example, if \( |\psi\rangle = a|\alpha\rangle + b|\beta\rangle \) and \( |\gamma\rangle \) is any other vector, then the bra of \( |\psi\rangle \) satisfies

\[
\langle \psi|\gamma \rangle = \omega(|\psi\rangle, |\gamma\rangle) = \omega(a|\alpha\rangle + b|\beta\rangle, |\gamma\rangle) = a^* \omega(|\alpha\rangle, |\gamma\rangle) + b^* \omega(|\beta\rangle, |\gamma\rangle) = a^* \langle \alpha|\gamma \rangle + b^* \langle \beta|\gamma \rangle = (a^* \langle \alpha \rangle + b^* \langle \beta \rangle) \langle \gamma \rangle.
\]

This shows that the bra of \( |\psi\rangle = a|\alpha\rangle + b|\beta\rangle \) is \( \langle \psi| = a^* \langle \alpha| + b^* \langle \beta| \). In general, if \( \{|\mu_i\rangle\}_{i=1}^n \) is any other set of complex coefficients, then the mapping between bras and kets is given by

\[
\text{ket: } \sum_{i=1}^n a_i |\mu_i\rangle \quad \leftrightarrow \quad \text{bra: } \sum_{i=1}^n a_i^* \langle \mu_i|.
\]

In particular, this implies that if \( \{|\nu_i\rangle\}_{i=1}^{n'} \) is another set of vectors and \( \{b_i\}_{i=1}^{n'} \) another set of complex coefficients, then the inner product between vectors \( |\psi\rangle = \sum_{i=1}^n c_i |\mu_i\rangle \) and \( |\varphi\rangle = \sum_{i=1}^{n'} d_i |\nu_i\rangle \) is given by

\[
\langle \psi|\varphi \rangle = \left( \sum_{i=1}^n a_i^* \langle \mu_i| \right) \left( \sum_{j=1}^{n'} b_j \langle \nu_j| \right) = \sum_{i=1}^n \sum_{j=1}^{n'} a_i^* b_j \langle \mu_i|\nu_j\rangle.
\]

A complex vector space with a defined inner product is called an inner product space. In finite-dimensions, inner product spaces are also known as **Hilbert spaces**. When working with infinite-dimensional vector spaces, Hilbert spaces require more structure than just being an inner product space. But we will not worry about such details in this course since we only consider finite dimensions, where the two types of spaces are equivalent. As is common practice, we will often denote the Hilbert space we are working in as \( \mathcal{H} \) (in place of \( V \)).

In a Hilbert space, the **norm** of a vector \(|\psi\rangle\) is defined as

\[
|||\psi||| := \sqrt{\langle \psi|\psi \rangle}.
\]

A vector \(|\psi\rangle\) can always be multiplied by a complex scalar so that its norm is one; such a vector is called “normalized.” Notice that this scalar multiplier is not unique. Specifically, if \(|\psi\rangle\) is an arbitrary vector, then \(\frac{|\psi\rangle}{\sqrt{\langle \psi|\psi \rangle}}\) is a normalized vector for any \( \theta \in \mathbb{R} \). The factor \( e^{i\theta} \) is called an **overall phase**, but as we will see shortly, overall phase factors have no physical meaning in quantum theory. Therefore, when dividing \(|\psi\rangle\) by its norm to obtain a normalized vector, without loss of generality we can always choose the overall phase to be +1 (i.e. take \( \theta = 0 \).
1.2 Computational Basis and Matrix Representations

Every $d$-dimensional Hilbert space possesses an orthonormal basis $\{ |\epsilon_1\rangle, |\epsilon_2\rangle, \ldots, |\epsilon_d\rangle \}$ which is a set of vectors that satisfy $\langle \epsilon_i | \epsilon_j \rangle = \delta_{ij}$. The existence of an orthonormal basis for any vector space $V$ follows from the so-called Gram-Schmidt orthogonalization process, which we will not review here. Any normalized vector $|\varphi\rangle \in \mathcal{H}$ can be expressed as a linear combination of the basis vectors $|\varphi\rangle = \sum_{k=1}^{d} b_k |\epsilon_k\rangle$. By contracting both sides of the first equality by $\langle \epsilon_i |$, we see that

$$\langle \epsilon_j | \varphi \rangle = \langle \epsilon_j | \left( \sum_{k=1}^{d} b_k |\epsilon_k\rangle \right) = \sum_{k=1}^{d} b_k \langle \epsilon_j | \epsilon_k \rangle = \sum_{k=1}^{d} b_k \delta_{jk} = b_j.$$  

Therefore, for an orthonormal basis $\{|\epsilon_k\rangle\}_{k=1}^{d}$, an arbitrary vector $|\varphi\rangle$ in $\mathcal{H}$ is a linear combination of the form

$$|\varphi\rangle = \sum_{k=1}^{d} \langle \epsilon_k | \varphi \rangle |\epsilon_k\rangle. \quad (8)$$

There are an infinite number of orthonormal bases for a vector space. However, when performing calculations, it is customary to choose one particular basis and relabel the kets simply as $\{|1\rangle, |2\rangle, \ldots, |d\rangle \}$. This basis is then often referred to as the **computational basis**. It should be emphasized that the computational basis is not something intrinsic to the vector space, but rather it is chosen by the individual for mathematical convenience or because the elements of a particular basis correspond to physically important objects.

Once a computational basis is chosen, the entire vector space can be represented using column matrices. Namely, one first makes the column vector identifications and then extends by linearity:

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad |2\rangle = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots \quad |d\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \quad \Rightarrow \quad \sum_{k=1}^{d} b_k |k\rangle = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_d \end{pmatrix}. \quad (9)$$

Customarily, the notation changes slightly when the Hilbert space has dimension $d = 2$. Two-dimensional Hilbert spaces play an important role in quantum information theory as they represent quantum systems known as **qubits**. We will learn quite a bit more about qubits in the upcoming chapters. But for now, we just remark that when dealing with two-dimensional spaces, the computational basis vectors are denoted by $|0\rangle$ and $|1\rangle$ (instead of $|1\rangle$ and $|2\rangle$) with corresponding column vector representations

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (10)$$

Two vector spaces are called **isomorphic** to one another if there exists an invertible linear map between the two spaces. The mapping in Eq. (9) shows that every $d$-dimensional Hilbert space $\mathcal{H}$ is isomorphic to $\mathbb{C}^d$, a relationship we denote by $\mathcal{H} \cong \mathbb{C}^d$. Throughout this book we will also use the notation $\cong$ to identify abstract elements in an arbitrary vector space $\mathcal{H}$ with their coordinate representations in $\mathbb{C}^d$, as in Eqs. (9) and (10).

The computational basis is an orthonormal basis meaning that the inner product between any two basis vectors satisfies $\langle l | k \rangle = \delta_{lk}$. For two vectors $|\psi\rangle = \sum_{k=1}^{d} a_k |k\rangle$ and $|\varphi\rangle = \sum_{k=1}^{d} b_k |k\rangle$, their inner product is computed as in Eq. (6):

$$\langle \psi | \varphi \rangle = \left( \sum_{l=1}^{d} a_l^* \langle l | \right) \left( \sum_{k=1}^{d} b_k |k\rangle \right) = \sum_{l,k=1}^{d} a_l^* b_k \langle l | k \rangle = \sum_{l,k=1}^{d} a_l^* b_k \delta_{lk} = \sum_{k=1}^{d} a_k^* b_k. \quad (11)$$
Figure 1: A direct sum decomposition of $\mathcal{H}$ into four orthogonal subspaces. The only vector common to each of the subspaces is the zero vector $0$. Any $|\psi\rangle \in \mathcal{H}$ can be uniquely expressed as a linear combination of vectors $|\phi_i\rangle \in V_i$.

This equation shows that if bras are represented by row vectors, then the inner product can be computed by standard matrix multiplication. Namely, for a ket $|\psi\rangle = \sum_k a_k |k\rangle$, one obtains the row vector representation of its bra according to $x\psi|\phi\rangle = (a_1^* a_2^* \cdots a_d^*) \cdot \left( \begin{array}{c} b_1 \\ b_2 \\ \vdots \\ b_d \end{array} \right) = \sum_{k=1}^d a_k^* b_k$. \hfill (12)

### 1.3 Direct Sum Decompositions of Hilbert Space

The notion of vector orthogonality can be extended to subspaces. We say that subspace $V_1$ is orthogonal to subspace $V_2$ if $\langle \alpha_1 | \alpha_2 \rangle = 0$ for every $|\alpha_1\rangle \in V_1$ and $|\alpha_2\rangle \in V_2$. An orthogonal direct sum decomposition of a finite-dimensional Hilbert space $\mathcal{H}$ is a collection of orthogonal subspaces $\{V_i\}_{i=1}^n$ whose union contains all of $\mathcal{H}$ (see Fig. 1.3).\footnote{A subspace of Hilbert space $\mathcal{H}$ is any subset $V \subset \mathcal{H}$ that is itself a vector space}

When $\{V_i\}_{i=1}^n$ provides a direct sum decomposition of $\mathcal{H}$, we write

$$\mathcal{H} = V_1 \oplus V_2 \oplus \cdots \oplus V_n. \hfill (13)$$

Since $\mathcal{H} = \bigcup_{i=1}^n V_i$, every $|\psi\rangle \in \mathcal{H}$ can be uniquely written as $|\psi\rangle = \sum_i a_i |\phi_i\rangle$ with $|\phi_i\rangle \in V_i$. For each subspace $V_i$ an orthonormal basis $\{|e_{i,k}\rangle\}_{k=1}^{d_i}$ can be chosen, and the full set $\bigcup_{i=1}^n \{|e_{i,k}\rangle\}_{k=1}^{d_i}$ constitutes an orthonormal basis for $\mathcal{H}$. Here the dimension of $V_i$ is $d_i$ and obviously $\sum_{i=1}^t d_i = d$, where $d$ is the dimension of $\mathcal{H}$.

For any subspace $V \subset \mathcal{H}$, its orthogonal complement, denoted by $V^\perp$, is the collection of all $|\varphi\rangle$ such that $\langle \psi | \varphi \rangle = 0$ for every $|\psi\rangle \in \mathcal{H}$. It is easy to verify that $V^\perp$ is a vector space; indeed if $\langle \psi | \varphi \rangle = \langle \psi | \varphi' \rangle = 0$ for any $|\psi\rangle \in \mathcal{H}$, then $\langle \psi | (|\varphi\rangle + |\varphi'\rangle) \rangle = 0$. A natural direct sum decomposition of $\mathcal{H}$ is then given by

$$\mathcal{H} = V \oplus V^\perp. \hfill (14)$$

From the definitions of vector space dimension and direct sum decomposition, we obtain the following proposition.

**Proposition 1.** Let $V \subset \mathcal{H}$ be any subspace of a $d$-dimensional space $\mathcal{H}$. Then

$$\dim[\mathcal{H}] = \dim[V] + \dim[V^\perp], \hfill (14)$$

where $\dim[V]$ is the dimension of $V$ and likewise for $\dim[V^\perp]$.\footnote{A subspace of Hilbert space $\mathcal{H}$ is any subset $V \subset \mathcal{H}$ that is itself a vector space}
2 Linear Operators

2.1 Outer Products, Adjoints, and Eigenvalues

For two Hilbert spaces \( \mathcal{H}_1 \) and \( \mathcal{H}_2 \), a linear operator \( R \) is simply a linear function that maps vectors in \( \mathcal{H}_1 \) to \( \mathcal{H}_2 \). That is, for any \( |\alpha\rangle, |\beta\rangle \in \mathcal{H}_1 \) and any \( a, b \in \mathbb{C} \), the operator \( R \) has action

\[
R (a|\alpha\rangle + b|\beta\rangle) = aR|\alpha\rangle + bR|\beta\rangle \in \mathcal{H}_2.
\]

We denote the set of all linear operators mapping \( \mathcal{H}_1 \) to \( \mathcal{H}_2 \) by \( \text{L}(\mathcal{H}_1, \mathcal{H}_2) \). For example, the vector space of bras acting \( \mathcal{H} \) is precisely the set \( \text{L}(\mathcal{H}, \mathbb{C}) \). When \( \mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H} \), we denote \( \text{L}(\mathcal{H}, \mathcal{H}) \) simply by \( \text{L}(\mathcal{H}) \).

Equation (15) implies that every linear operator is defined by its action on a basis. Specifically if \( \mathcal{H}_1 \) is a \( d_1 \)-dimensional space with computational basis \( \{|i\rangle\}_{i=1}^{d_1} \), then an operator \( R \in \text{L}(\mathcal{H}_1, \mathcal{H}_2) \) acts on this basis according to

\[
R|i\rangle = \sum_{j=1}^{d_2} R_{ji}|j\rangle \quad \text{for } k = 1, \ldots, d_1,
\]

where \( \{|j\rangle\}_{j=1}^{d_2} \) is the computational basis in \( \mathcal{H}_2 \). By contracting both the RHS and LHS with \( \langle j | \) we see that

\[
R_{ji} = \langle j | R|i\rangle.
\]

Bra-ket notation provides a convenient way to write any linear operator. With respect to bases \( \{|i\rangle\}_{i=1}^{d_1} \) for \( \mathcal{H}_1 \) and \( \{|j\rangle\}_{j=1}^{d_2} \) for \( \mathcal{H}_2 \), any linear operator \( R \) can be written as

\[
R = \sum_{i=1}^{d_1} \sum_{j=1}^{d_2} R_{ji}|j\rangle \langle i|.
\]

When \( R \) acts on a state \( |\psi\rangle \), the transformed state is

\[
R|\psi\rangle = \left( \sum_{i=1}^{d_1} \sum_{j=1}^{d_2} R_{ji}|j\rangle \langle i| \right) |\psi\rangle = \sum_{i=1}^{d_1} \sum_{j=1}^{d_2} R_{ji} \langle i|\psi\rangle |j\rangle.
\]

In general, the “ket-bra” \( |\beta\rangle \langle \alpha| \) is called the outer product of kets \( |\alpha\rangle \in \mathcal{H}_1 \) and \( |\beta\rangle \in \mathcal{H}_2 \). It is the linear operator in \( \text{L}(\mathcal{H}_1, \mathcal{H}_2) \) that acts on an arbitrary ket \( |\psi\rangle \) by outputting the ket \( |\beta\rangle \) multiplied by the bra-ket (i.e. the inner product) \( \langle \alpha|\psi\rangle \):  

\[
|\psi\rangle \rightarrow (|\beta\rangle \langle \alpha|) |\psi\rangle = \langle \alpha|\psi\rangle |\beta\rangle.
\]

Eqns. (16) and (17) show how a linear operator can be represented as a matrix. Using bases \( \{|i\rangle\}_{i=1}^{d_1} \) and \( \{|j\rangle\}_{j=1}^{d_2} \) for \( \mathcal{H}_1 \) and \( \mathcal{H}_2 \) respectively, an operator \( R \) can have the matrix representation

\[
R = \begin{pmatrix}
\langle 1|R|1\rangle & \langle 1|R|2\rangle & \cdots & \langle 1|R|d_1\rangle \\
\langle 2|R|1\rangle & \langle 2|R|2\rangle & \cdots & \langle 2|R|d_1\rangle \\
\vdots & \vdots & \ddots & \vdots \\
\langle d_2|R|1\rangle & \langle d_2|R|2\rangle & \cdots & \langle d_2|R|d_1\rangle
\end{pmatrix}.
\]
An alternative way of forming this matrix can be obtained by taking a linear combination of outer products. For a kets \( |a\rangle = \sum_{i=1}^{d_1} a_i |i\rangle \) and \( |\beta\rangle = \sum_{j=1}^{d_2} b_j |j\rangle \), their outer product is computed in matrix form as

\[
|\beta\rangle \langle a| = \begin{pmatrix}
    b_1 \\
    b_2 \\
    \vdots \\
    b_{d_2}
\end{pmatrix}
\begin{pmatrix}
    a_1^* & a_2^* & \cdots & a_{d_1}^*
\end{pmatrix} =
\begin{pmatrix}
    b_1 a_1^* & b_1 a_2^* & \cdots & b_1 a_{d_1}^* \\
    b_2 a_1^* & b_2 a_2^* & \cdots & b_2 a_{d_1}^* \\
    \vdots & \vdots & \ddots & \vdots \\
    b_{d_2} a_1^* & b_{d_2} a_2^* & \cdots & b_{d_2} a_{d_1}^*
\end{pmatrix}.
\tag{21}
\]

In particular, the outer product \( |j\rangle \langle i| \) is simply the \( d_2 \times d_1 \) matrix that has zeros everywhere except a 1 in the \( j \)th row and \( i \)th column. From Eq. (17), the matrix for \( R \) can be formed by a linear combination of the \( |j\rangle \langle i| \) with the coefficients being \( R_{ji} \).

For every operator \( R \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2) \), there exists a unique operator \( R^\dagger \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_1) \) called the adjoint of \( R \) that is defined by the condition that

\[
\langle \beta | R^\dagger | \alpha \rangle = \langle \alpha | R | \beta \rangle^* \quad \forall |\alpha\rangle \in \mathcal{H}_1, \; |\beta\rangle \in \mathcal{H}_2.
\tag{22}
\]

From this equation we can easily determine the bra for vector \( R |\beta\rangle \). Indeed, let \( |\gamma\rangle = R |\beta\rangle \). Then Eq. (22) says that \( \langle \alpha | \gamma \rangle^* = \langle \beta | R^\dagger | \alpha \rangle \). But we also know that \( \langle \gamma | \alpha \rangle = \langle \alpha | \gamma \rangle^* \), and so \( \langle \gamma | \alpha \rangle = \langle \beta | R^\dagger | \alpha \rangle \). Since this holds for any \( |\alpha\rangle \), we obtain the general rule

\[
\text{ket: } R |\beta\rangle \quad \Longleftrightarrow \quad \text{bra: } \langle \beta | R^\dagger.
\tag{23}
\]

In terms of its matrix representation, we can write

\[
\langle j | R^\dagger | i \rangle = \langle i | R | j \rangle^* = R_{ij}^*
\tag{24}
\]

for the orthonormal bases \( \{|i\rangle\}_{i=1}^{d_1} \) and \( \{|j\rangle\}_{j=1}^{d_2} \). We therefore see that

\[
R^\dagger = \sum_{i=1}^{d_1} \sum_{j=1}^{d_2} R_{ij}^* |j\rangle \langle i|.
\tag{25}
\]

The matrix representation for \( R^\dagger \) is obtained by taking the complex conjugate of each element in \( R \) and then taking the matrix transpose.

Finally, we recall that the eigenvalue of a linear operator \( R \) is a complex number \( \lambda \) such that \( R |\phi\rangle = \lambda |\phi\rangle \) for some vector \( |\phi\rangle \). Note that \( |\phi\rangle \) will not be the unique vector satisfying this equation. In fact, it is easy to see that for a fixed \( \lambda \), the set \( V_\lambda = \{|\phi\rangle : R |\phi\rangle = \lambda |\phi\rangle \} \) forms a vector space. This vector space \( V_\lambda \) is called the eigenspace associated with \( \lambda \) (or the \( \lambda \)-eigenspace), and any vector from this space is an eigenvector of \( R \) with eigenvalue \( \lambda \). The kernel (or null space) of an operator \( R \) is the eigenspace associated with the eigenvalue 0, and it is denoted by \( \ker(R) \).

**Lemma 1.** For a nonempty Hilbert space \( \mathcal{H} \), every operator \( R \in \mathcal{L}(\mathcal{H}) \) has at least one eigenvalue.

**Proof.** The proof of this lemma is borrowed from Axler’s textbook [Axl97], and it does not rely on determinants. Suppose \( \mathcal{H} \) is \( d \)-dimensional, and let \( |\phi\rangle \) be any element of \( \mathcal{H} \). Consider the set of \( d + 1 \) vectors \( \{|\phi\rangle, R |\phi\rangle, R^2 |\phi\rangle, \ldots, R^d |\phi\rangle \} \). As the dimension of \( \mathcal{H} \) is \( d \), these vectors must be linearly dependent. Hence there exists constants \( c_i \) not all equaling zero such that

\[
0 = \left( \sum_{k=0}^{d} c_k R^k \right) |\phi\rangle.
\]
Now consider the polynomial \( p(z) = \sum_{k=0}^{d} c_k z^k \) for \( z \in \mathbb{C} \). If instead of a complex number \( z \) we substitute the operator \( R \) and multiply additive constants by the identity, then we obtain \( p(R) := \sum_{k=0}^{d} c_k R^k \), which is precisely the term in parentheses in the above equation. The fundamental theorem of algebra says that every degree \( d \) polynomial has \( d \) roots \( r_1, r_2, \ldots, r_d \) (including multiplicities), and so \( r(z) \) can be factored as \( p(z) = c \prod_{k=1}^{d} (z - r_k) \). Applying the same factorization to \( p(R) \), we obtain
\[
0 = p(R)|\phi\rangle = c \prod_{k=1}^{d} (R - r_k I)|\phi\rangle.
\]
(26)

This implies that there must exist some \( r_k \) such that \( 0 = (R - r_k I)|\phi_k\rangle \). \( \square \)

We finally describe an important relationship between the kernel of \( R \) and its range. The **range** (or **image**) of a linear operator \( R \in \text{L}(\mathcal{H}_1, \mathcal{H}_2) \) is the vector space defined by
\[
\text{Rng}(R) = \left\{ |\phi\rangle \in \mathcal{H}_2 \mid |\phi\rangle = R|\psi\rangle \text{ for some } |\psi\rangle \in \mathcal{H}_1 \right\}.
\]
The image of an operator is itself a vector space, a fact that can be easily verified. The dimension of \( \text{Rng}(R) \) is called the **rank** of \( R \) (denoted by \( \text{rk}(R) \)), and the following fundamental theorem relates \( \text{rk}(R) \) and \( \dim[\ker(R)] \).

**Theorem 1** (Rank-Nullity Theorem). For any operator \( R \in \text{L}(\mathcal{H}_1, \mathcal{H}_2) \),
\[
dim[\mathcal{H}_1] = \text{rk}(R) + \dim[\ker(R)].
\]
(27)

**Proof.** By Prop. 1 we must show that \( \text{rk}(R) = \dim[\ker(R)] \). Expand \( R \) in the computational basis as
\[
R = \sum_{i=1}^{d_1} \sum_{j=1}^{d_2} R_{ji}|i\rangle\langle j| = \sum_{j=1}^{d_2} \sum_{i=1}^{d_1} |j\rangle\langle i|, \quad \text{where} \quad \langle j| = \sum_{i=1}^{d_1} R_{ji}|i|.
\]

From this we see that \( |\phi\rangle \in \ker(R) \iff \langle j|\phi\rangle = 0 \) for all \( j \). In other words, \( \{|\phi_j\rangle\}_{j=1}^{d_1} \) belong to \( \ker(R) \). Let us take \( \{|\psi_i\rangle\}_{i=1}^{r} \) as an orthonormal basis for \( \ker(R)^\perp \) so that we can write \( |\phi_j\rangle = \sum_{k=1}^{r} c_{jk}|\psi_k\rangle \) for every \( |\phi_j\rangle \). Then
\[
R = \sum_{j=1}^{d_2} |j\rangle\langle j| = \sum_{j=1}^{d_2} \sum_{k=1}^{r} c_{jk}^* |j\rangle\langle k| = \sum_{k=1}^{r} |v_k\rangle\langle v_k|.
\]
(28)

where \( |v_k\rangle = \sum_{j=1}^{d_2} c_{jk}^* |j\rangle \). The \( \{|v_k\rangle\}_{k=1}^{r} \) must be linearly independent. Indeed, suppose that \( \sum_{k=1}^{r} a_k |v_k\rangle = 0 \) for \( a_k \) not all zero. Then \( |\psi\rangle = \sum_{i=1}^{r} a_i^* |\psi_i\rangle \) satisfies \( R|\psi\rangle = 0 \), which is impossible since all the \( |\psi_i\rangle \) belong to \( \ker(R)^\perp \). Hence, the \( \{|v_k\rangle\}_{k=1}^{r} \) are linearly independent, and they clearly span \( \text{Rng}(R) \). This proves that \( r = \dim[\ker(R)^\perp] = \dim[\text{Rng}(R)] \). \( \square \)

Typically, the rank of an operator \( R \) is defined as the dimension of the column space (or row space) when \( R \) is represented as a matrix. Eq. (28) shows that this is equivalent to the definition given here of \( \text{rk}(R) = \dim[\text{Rng}(R)] \). Indeed, the \( \{|v_k\rangle\}_{k=1}^{r} \) are precisely the column vectors when expressing \( R \) in an \( \{|\psi_k\rangle\}_{k=1}^{r} \) basis of \( \mathcal{H}_1 \).
2.2 Important Types of Operators

In quantum mechanics certain types of operators are vitally important to the theory as they correspond to certain physical processes or objects. This will be discussed in the next chapter. For now we stay on the level of mathematics and describe five different classes of linear operators. Their inclusion structure is depicted in Fig. 2.2.

2.2.1 Projections

We begin with projectors since, as the Spectral Theorem shows below (Thm. 2), they are the basic building block for all normal operators. A projector on a d-dimensional space $\mathcal{H}$ is any operator of the form

$$P_V = \sum_{i=1}^{s} |\epsilon_i\rangle\langle\epsilon_i|,$$  \hspace{1cm} (29)

where the $\{|\epsilon_i\rangle\}_{i=1}^{s}$ is an orthonormal set of vectors for some integer $1 \leq s \leq d$. These vectors form a basis for some $s$-dimensional vector subspace $V \subset \mathcal{H}$, and $P_V$ is called the subspace projector onto $V$. To understand why it has this name, first note that $P_V|\psi\rangle = |\psi\rangle$ for any $|\psi\rangle \in V$. This equality can be verified by writing $|\psi\rangle = \sum_i \langle\epsilon_i|\psi\rangle|\epsilon_i\rangle$ (which is possible since $|\psi\rangle \in V$) and then applying $P_V$ to both sides:

$$P_V|\psi\rangle = \left(\sum_{i=1}^{s} |\epsilon_i\rangle\langle\epsilon_i|\right)|\psi\rangle = \sum_{i=1}^{s} \langle\epsilon_i|\psi\rangle|\epsilon_i\rangle = |\psi\rangle.$$

Next, let $V^\perp$ be the orthogonal complement of $V$ and consider the direct sum decomposition $\mathcal{H} = V \oplus V^\perp$. Any $|\psi\rangle \in \mathcal{H}$ can be uniquely written as $|\psi\rangle = a|\varphi_1\rangle + b|\varphi_2\rangle$ with $|\varphi_1\rangle \in V$ and $|\varphi_2\rangle \in V^\perp$. Then since $P_V|\varphi_1\rangle = |\varphi_1\rangle$ and $P_V|\varphi_2\rangle = 0$, we have $P_V|\psi\rangle = a|\varphi_1\rangle$. In other words, when $P_V$ acts on arbitrary vector, it “projects” the vector onto its component that lies in $V$.

Notice that if $V_1$ and $V_2$ are orthogonal subspaces of $\mathcal{H}$, then it is easy to verify that

$$P_{V_i}P_{V_j} = \delta_{ij}P_{V_i}, \quad i, j \in \{1, 2\}.$$  \hspace{1cm} (30)

In particular, every projector satisfies $P_V^2 = P_V$ for any subspace $V$. Operators having the property $A^2 = A$ are often called idempotent.
For $s > 1$, there are an infinite number of orthonormal bases for a given $s$-dimensional subspace $V$, and each of these provides a different way to represent the subspace projector. In other words, if $\{|e_i\rangle\}_{i=1}^s$ and $\{|\delta_i\rangle\}_{i=1}^s$ are both orthonormal bases for $V$, then

$$\sum_{i=1}^s |e_i\rangle\langle e_i| = \sum_{i=1}^s |\delta_i\rangle\langle \delta_i|.$$  

An extreme case is when $V$ is the entire vector space $\mathcal{H}$. In this case, the subspace projector is simply the identity operator $I$, i.e. the unique operator in $L(\mathcal{H})$ satisfying $I|\psi\rangle = |\psi\rangle$ for all $|\psi\rangle \in \mathcal{H}$. Then if $\{|e_i\rangle\}_{i=1}^d$ is any orthonormal basis for $\mathcal{H}$, we can express the identity as

$$I = \sum_{i=1}^d |e_i\rangle\langle e_i|.$$  

Throughout this course we will work heavily with the eigenspaces of various operators. For an operator $R$ with eigenvalue $\lambda$, we define its $\lambda$-eigenprojector to be the projector $P_\lambda$ that projects onto the $\lambda$-eigenspace $V_\lambda$.

### 2.2.2 Normal Operators

An operator $N \in L(\mathcal{H})$ is called normal if $NN^\dagger = N^\dagger N$. Note, each normal operator is represented by a square matrix since it maps a given vector space $\mathcal{H}$ onto itself. In general, every normal operator has very nice mathematical properties, two of which are the following.

**Proposition 2.** If $N \in L(\mathcal{H})$ is normal with eigenvalue $\lambda$ and eigenspace $V_\lambda$, then $\lambda^*$ is an eigenvalue $N^\dagger$ with eigenspace $V_\lambda$.

**Proof.** First, let $|\psi\rangle$ be any element in $V_\lambda$. By assumption we have that $(N - \lambda I)|\psi\rangle = 0$. It is easy to see that if $N$ is normal, then so is $N - \lambda I$. Therefore,

$$0 = \langle \psi|(N - \lambda I)^\dagger(N - \lambda I)|\psi\rangle$$

$$= \langle \psi|(N - \lambda I)(N - \lambda I)^\dagger|\psi\rangle. \tag{31}$$

The last line implies that $(N^\dagger - \lambda^* I)|\psi\rangle = 0$, which is what we set out to prove. $\square$

**Proposition 3.** If $N \in L(\mathcal{H})$ is normal with distinct eigenvalues $\lambda_1$ and $\lambda_2$, then the associated eigenspaces $V_{\lambda_1}$ and $V_{\lambda_2}$ are orthogonal. Furthermore, $\mathcal{H} = \bigcup_{\lambda_i} V_{\lambda_i}$, where the union is taken over all distinct eigenvalues of $N$.

**Proof.** Let $|\psi_1\rangle \in V_{\lambda_1}$ and $|\psi_2\rangle \in V_{\lambda_2}$ be arbitrary. Then (i) $N|\psi_1\rangle = \lambda_1|\psi_1\rangle$, and (ii) $N^\dagger|\psi_2\rangle = \lambda_2^*|\psi_2\rangle$. Hence

$$\langle \psi_2\rangle N|\psi_1\rangle = \langle \psi_2\rangle (N|\psi_1\rangle) = \lambda_1 \langle \psi_2\rangle|\psi_1\rangle$$

$$= \langle \langle \psi_2\rangle N|\psi_1\rangle = \lambda_2 \langle \psi_2\rangle|\psi_1\rangle, \tag{32}$$

where the second line follows from (ii). Subtracting these gives $\langle \psi_2|\psi_1\rangle = 0$.

To prove the second claim of the proposition let $V = \bigcup_{\lambda_i} V_{\lambda_i}$. We want to show that $V^\perp = \emptyset$. To this end, let $\hat{N}$ be the operator $N$ with domain restricted to $V^\perp$; that is, $\hat{N} : V^\perp \to \mathcal{H}$ with action defined by $\hat{N}|\beta\rangle = N|\beta\rangle$ for all $|\beta\rangle \in V^\perp$. Notice that for any $|\alpha\rangle \in V$ and $|\beta\rangle \in V^\perp$ we have

$$\langle \alpha|\hat{N}|\beta\rangle = \langle \alpha|N|\beta\rangle = \langle \beta|N^\dagger|\alpha\rangle^* = 0,$$
where we have used Prop. 2 to conclude that $N|\alpha\rangle \in V$. This shows that $\hat{N}$ maps $V^\perp$ onto itself; i.e. $\hat{N} \in L(V^\perp)$. Therefore, we can use Lem. 1 which says that if $V^\perp$ is nonempty, then $\hat{N}$ has at least one eigenvalue and eigenvector. But this would be a contradiction since any eigenvector of $\hat{N}$ is also an eigenvector $N$, and all eigenspaces of $N$ are contained in $V$. Hence $V^\perp$ must by the empty set. 

Proposition 3 shows that every normal operator $N \in L(\mathcal{H})$ is uniquely associated with a direct sum decomposition

$$\mathcal{H} = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_n}$$

(33)

where the $V_{\lambda_i}$ are the eigenspaces of $N$. This directly leads to the spectral decomposition of $N$, which is one of the most important mathematical theorems in quantum mechanics.

**Theorem 2 (Spectral Decomposition).** Every normal operator $N \in L(\mathcal{H})$ can be uniquely written as

$$N = \sum_{i=1}^{n} \lambda_i P_{\lambda_i}$$

(34)

where the $\{\lambda_i\}_{i=1}^{n}$ are the distinct eigenvalues of $N$ and the $\{P_{\lambda_i}\}_{i=1}^{n}$ are the corresponding eigenspace projectors.

**Proof.** By the direct sum decomposition of Eq. (33), any vector $|\psi\rangle \in \mathcal{H}$ can be written as $|\psi\rangle = \sum_{i=1}^{n} |\alpha_i\rangle$ with $|\alpha_i\rangle \in V_{\lambda_i}$. Thus,

$$(N - \sum_{i=1}^{n} \lambda_i P_{\lambda_i})|\psi\rangle = N\sum_{i=1}^{n} |\alpha_i\rangle - \sum_{i=1}^{n} \lambda_i P_{\lambda_i}\sum_{j=1}^{n} |\alpha_j\rangle = \sum_{i=1}^{n} \lambda_i |\alpha_i\rangle - \sum_{i=1}^{n} \lambda_i |\alpha_i\rangle = 0.$$ 

Since this holds for all $|\psi\rangle \in \mathcal{H}$, we must have $(N - \sum_{i=1}^{n} \lambda_i P_{\lambda_i}) = 0$, or $N = \sum_{i=1}^{n} \lambda_i P_{\lambda_i}$. 

If a normal operator has a zero eigenvalue, then the associated eigenspace is $\text{ker}(N)$. The vector space

$$\text{supp}(N) := \text{ker}(N)^\perp = \bigcup_{\lambda_k \not= 0} V_{\lambda_k}$$

(35)

is called the support of $N$, and it is denoted by $\text{supp}(N)$. From the spectral decomposition, we can immediately see that the support of $N$ is equivalent to its range; i.e. $\text{supp}(N) = \text{rng}(N)$.

Since every projector is a normal operator, it also has a spectral decomposition. In fact, this is precisely what is given in Eq. (29). Then if $\{|\lambda_{i,j}\rangle\}_{j=1}^{d_i}$ is an orthonormal basis for the eigenspace $V_{\lambda_i}$ of a normal operator $N$, we can write the spectral decomposition of $N$ as

$$N = \sum_{i=1}^{n} \lambda_i P_{\lambda_i} = \sum_{i=1}^{n} \lambda_i \sum_{j=1}^{d_i} |\lambda_{i,j}\rangle \langle \lambda_{i,j}| = \sum_{i=1}^{n} \sum_{j=1}^{d_i} \lambda_i |\lambda_{i,j}\rangle \langle \lambda_{i,j}|.$$ 

(36)

Customarily, we combine the two sums over $(i,j)$ into a single sum where the same eigenvalue is counted multiple times. More precisely, the eigenvalue $\lambda_i$ is counted $\text{dim}[V_{\lambda_i}]$ times. Since $\sum_{i=1}^{n} d_i = d = \text{dim}[\mathcal{H}]$ we can write

$$N = \sum_{k=1}^{d} \lambda_k |\lambda_k\rangle \langle \lambda_k|.$$ 

(37)

As a matter of convention, whenever the $\lambda_k$ are enumerated from $k = 1, \cdots, d$ we will assume that the set $\{\lambda_k\}_{k=1}^{d}$ may contain repeated eigenvalues. The full collection $\{|\lambda_k\rangle\}_{k=1}^{d}$ spans all of $\mathcal{H}$, and it is sometimes called a complete eigenbasis.
2.2.3 Unitary Operators

An element $U \in L(H)$ is called unitary if $U^\dagger U = UU^\dagger = I$. In fact, it is easy to see that if $U^\dagger U = I$, then we must also have $UU^\dagger = I$. Indeed, suppose that $U^\dagger U = I$, and let $\{|\psi_i\rangle\}_{i=1}^d$ be an orthonormal basis for $H$. Clearly $U|\psi_i\rangle$ is also an orthonormal basis. But then $U^\dagger U = I$ implies that $(UU^\dagger - I)|\psi_i\rangle = 0$ for all $i$, which is only possible if $UU^\dagger - I = 0$.

An important property of unitary operators is that they preserve inner products when transforming vectors: $\langle \alpha | \beta \rangle = \langle U | U^\dagger | \alpha \rangle \langle \beta | U^\dagger | \rangle$ for all $|\alpha\rangle, |\beta\rangle \in H$. In particular, if $\{|\psi_i\rangle\}_{i=1}^d$ is an orthonormal basis for $H$, then so will be $\{|U|\psi_i\rangle\}_{i=1}^d$ for any unitary operator $U$. If we denote $|\delta_i\rangle = U|\psi_i\rangle$, then the unitary $U$ can be written as

$$U = \sum_{i=1}^d |\delta_i\rangle \langle \psi_i|.$$  (38)

Conversely, for any two orthonormal bases, there always exists a unitary operator like Eq. (38) that transforms one basis to other.

An important class of unitary operators are permutations, which are operators that act invariantly on the set of computational basis vectors $\{|\delta\rangle, |1\rangle, \cdots, |d\rangle\}$. Any permutation has the form

$$\Pi = \sum_{i=1}^d |\pi(i)\rangle \langle i|,$$  (39)

where $\pi$ is a bijection acting on the set of integers $\{1, \cdots, d\}$. It is easy to verify that $\Pi^\dagger \Pi = \Pi \Pi^\dagger = I$.

Using unitary operators, the spectral decomposition can be expressed in an alternative form. Let $\{|\lambda_k\rangle\}_{k=1}^d$ be a complete eigenbasis of some normal operator $N$. Define the unitary operator

$$U = \sum_{k=1}^d |\lambda_k\rangle \langle k|,$$

which rotates the computational basis $\{|k\rangle\}_{k=1}^d$ into the eigenbasis of $N$. Then the spectral decomposition of $N$ can be written as

$$N = U \Lambda U^\dagger,$$  (40)

where $\Lambda$ is a $d \times d$ matrix diagonal in the computational basis of the form

$$\Lambda = \begin{pmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \lambda_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \lambda_d
\end{pmatrix},$$

with $\lambda_k$ an eigenvalue associated with eigenvector $|\lambda_k\rangle$. When applying a function $f$ to the matrix $N$, we have

$$\hat{f}(N) = U f(\Lambda) U^\dagger,$$  (41)

where $f(\Lambda)$ is the diagonal matrix obtained by applying $f$ to each of the eigenvalues of $N$ on the diagonal.
2.2.4 Hermitian Operators

A special type of normal operator are those satisfying $A = A^\dagger$. Such operators are called hermitian, and we will let $\text{Herm}(\mathcal{H}) \subseteq L(\mathcal{H})$ denote the space of all hermitian operators acting on $\mathcal{H}$. Proposition 2 says that $\lambda^*$ is an eigenvalue of $A^\dagger$ with eigenspace $V_\lambda$ whenever $\lambda$ is an eigenvalue of $A$ with the same eigenspace $V_\lambda$. In other words, if $|\psi\rangle \in V_\lambda$ is a nonzero vector, then $A|\psi\rangle = \lambda|\psi\rangle$ and $A^\dagger|\psi\rangle = \lambda^*|\psi\rangle$. Since $A = A^\dagger$ for hermitian operators, this immediately implies that $\lambda = \lambda^*$. We have thus proven an important fact.

**Proposition 4.** Every hermitian operator has real eigenvalues.

Less frequently encountered objects in quantum mechanics are anti-hermitian operators. These are normal operators satisfying $A = -A^\dagger$. In contrast to hermitian operators, all the eigenvalues of an anti-hermitian operator are purely imaginary.

2.2.5 Positive Operators

A normal operator $A$ is called positive (or positive semi-definite) if all its eigenvalues are nonnegative. It is called positive definite if all its eigenvalues are positive. An alternative characterization of positive operators is the following.

**Proposition 5.** A normal operator $A$ is positive if and only if 

$$\langle \psi | A | \psi \rangle \geq 0 \quad \text{for all} \quad |\psi\rangle \in \mathcal{H}.$$

If $R \in L(H_1, H_2)$ is an arbitrary operator, then $R^\dagger R \in L(H_1)$ and $RR^\dagger \in L(H_2)$ are both positive since $\langle \psi | R^\dagger R | \psi \rangle \geq 0$ for all $|\psi\rangle$. We will let $\text{Pos}(\mathcal{H}) \subseteq L(\mathcal{H})$ denote the set of all positive operators acting on $\mathcal{H}$.

2.3 Functions of Normal Operators

Let $f : \mathcal{X} \to \mathbb{C}$ be a complex-valued function whose domain is $\mathcal{X} \subseteq \mathbb{C}$. We would like to extend this function to be a mapping of normal operators. If $N$ has eigenvalues lying in $\mathcal{X}$, then we can define a new operator

$$\hat{f}(N) = \sum_{i=1}^{n} f(\lambda_i) P_{\lambda_i},$$

where $\sum_{i=1}^{n} \lambda_i P_{\lambda_i}$ is the spectral decomposition of $N$. For example, if $f(x) = x^2$, then $\hat{f}(N) = \sum_{i=1}^{N} \lambda_i^2 P_{\lambda_i}$. Furthermore, since

$$N^2 = \sum_{i=1}^{n} \lambda_i P_{\lambda_i} \sum_{j=1}^{n} \lambda_j P_{\lambda_j} = \sum_{i=1}^{n} \lambda_i^2 P_{\lambda_i},$$

we see that $\hat{f}(N) = N^2$. In other words, $\hat{f}$ has the same functional form as $f$ except applied to normal matrices having eigenvalues in the domain of $f$.

In many cases, we will be interested in studying functions whose domain does not include zero. In this case, we can still consider the operator obtained by applying the function to the nonzero eigenvalues of $N$. This is often described as restricting the function to the support of $N$. For example, the logarithm $f(x) = \log(x)$ is only defined in the interval $(0, +\infty)$. For a positive operator $X = \sum_{i=1}^{n} \lambda_i P_{\lambda_i}$, we define

$$\log(X) := \sum_{\lambda_i \neq 0} \log(\lambda_i) P_{\lambda_i},$$

13
Likewise, as a function on $C$, the multiplicative inverse is given by $f(z) = z^{-1} = \frac{1}{z}$. For any normal operator $N = \sum_{\lambda_i \neq 0} \lambda_i P_{\lambda_i}$, we can define
\[
N^{-1} := \sum_{\lambda_i \neq 0} \lambda_i^{-1} P_{\lambda_i}.
\] (45)

Note that
\[
N^{-1} N = NN^{-1} = \sum_{\lambda_i \neq 0} \lambda_i P_{\lambda_i} \sum_{\lambda_j \neq 0} \lambda_j^{-1} P_{\lambda_j} = \sum_{\lambda_i \neq 0} P_{\lambda_i} = P_{\text{supp}(N)}.
\] (46)

Hence, when $N$ has a zero eigenvalue, $N^{-1} N$ is not the identity on $\mathcal{H}$ but rather the projector onto $\text{supp}(N) \subset \mathcal{H}$.

### 2.4 The Polar and Singular Value Decompositions

We next describe two useful decompositions of a general linear operator $R \in \text{L(\mathcal{H})}$ acting on a Hilbert space $\mathcal{H}$.

**Lemma 2.** Every operator $R \in \text{L(\mathcal{H})}$ can be written as
\[
R = U \sqrt{R^\dagger R},
\] (47)

where $U$ is a unitary operator acting on $\mathcal{H}$.

**Proof.** Let us take a spectral decomposition of $R^\dagger R = \sum_{k=1}^n \lambda_k P_{\lambda_k}$. Observe that $R |\psi\rangle = 0$ iff $R^\dagger R |\psi\rangle = 0$. Consider first the case when $\ker(R) = \emptyset$, i.e. $R$ does not have any zero eigenvalue. Then we can define the operator
\[
U = \sum_{k=1}^n \lambda_k^{-1/2} A P_{\lambda_k}.
\] (48)

Since $\sqrt{R^\dagger R} = \sum_{k=1}^n \lambda_k^{1/2} P_{\lambda_k}$, we clearly have $R = U \sqrt{R^\dagger R}$. Furthermore, direct calculation shows that after $U^\dagger U = \mathbb{I}$, and so $U$ is unitary. Now consider the case when $R$ has a zero eigenvalue, and let $P_0$ be the projector onto $\ker(R)$. By Thm. 1, there exists an invertible map $T$ such that $T|\varepsilon_i\rangle = |\delta_i\rangle$, where $\{|\varepsilon_i\rangle\}_{i=r+1}^{\dim(\mathcal{H})}$ is an orthonormal basis of $\ker(R)$ and $\{|\delta_i\rangle\}_{i=r+1}^{\dim(\mathcal{H})}$ is an orthonormal basis of $\text{rng}(R)\perp$. Then we modify Eq. (48) be defining
\[
U = \sum_{\lambda_k \neq 0} \lambda_k^{-1/2} A P_{\lambda_k} + TP_0.
\] (49)

Since $P_0^\dagger A = A^\dagger T P_0 = 0$, we have that $U^\dagger U = \sum_{\lambda_k \neq 0} P_{\lambda_k} + P_0 = \mathbb{I}$. As before, we have $R = U \sqrt{R^\dagger R}$. \hfill \Box

For an operator $R$ with $R^\dagger R = \sum_{k=1}^n \lambda_k P_{\lambda_k}$ having eigenvalues $\lambda_k$, the non-negative numbers
\[
\sigma_k := \sqrt{\lambda_k} \quad k = 1, \cdots, n
\] (50)

are called the **singular values** of $R$. Using the polar decomposition, we obtain the singular value decomposition of $R$, which is one of the most useful tools in the study of quantum entanglement.
**Theorem 3** (Singular Value Decomposition). Every operator \( R \in \mathcal{L}(\mathcal{H}) \) can be written as

\[
R = \sum_{k=1}^{n} \sigma_k U P_{c_k},
\]

where \( \{ \sigma_k \}_{k=1}^{n} \) are the distinct singular values of \( R \), \( U \) is a unitary operator, and the \( \{ P_{c_k} \}_{k=1}^{n} \) are projections on the eigenspaces of \( R^\dagger R \). Equivalently, \( R \) can be written as

\[
R = V \Lambda_v W^\dagger,
\]

where \( V \) and \( W \) are unitaries, and \( \Lambda_v \) is a diagonal matrix in the computational basis with diagonal elements being the singular values of \( R \).

**Proof.** Eq. (51) follows immediately from the polar decomposition of \( R \) by writing \( \sqrt{R^\dagger R} = \sum_{k=1}^{n} \lambda_k^{1/2} P_{d_k} \). Likewise, Eq. (52) is obtained by writing the spectral decomposition \( R^\dagger R = W \Lambda W^\dagger \) so that \( \sqrt{R^\dagger R} = W \sqrt{\Lambda} W^\dagger = \Lambda_v W^\dagger \). Then Eq. (52) follows by setting \( V = U W \).

\[\square\]

3 **Tensor Products**

3.1 **Tensor Product Spaces and Vectors**

The tensor product is a construction that combines different Hilbert spaces to form one large Hilbert space. As we will see, tensor products are used to describe quantum systems composed of multiple subsystems. In anticipation of this application, we will henceforth label different Hilbert spaces by superscript capital letters (i.e. \( \mathcal{H}^A, \mathcal{H}^B, \) etc.) instead of by subscript numbers (i.e. \( \mathcal{H}_1, \mathcal{H}_2, \) etc.) The reason is that later we will associate Hilbert space \( \mathcal{H}^A \) with Alice, \( \mathcal{H}^B \) with Bob, etc.

If \( \mathcal{H}^A \) is a Hilbert space with computational basis \( \{ |i\rangle^A \}_{i=1}^{d_A} \) and \( \mathcal{H}^B \) is a Hilbert space with computational basis \( \{ |j\rangle^B \}_{j=1}^{d_B} \), then their **tensor product space** is a Hilbert space \( \mathcal{H}^A \otimes \mathcal{H}^B \) with orthonormal basis given by \( \{ |i\rangle^A \otimes |j\rangle^B \}_{i=1}^{d_A} \_{j=1}^{d_B} \), called a **tensor product basis**. Every vector \( |\psi\rangle^{AB} \in \mathcal{H}^A \otimes \mathcal{H}^B \) can be written as a linear combination of the basis vectors:

\[
|\psi\rangle^{AB} = \sum_{i=1}^{d_A} \sum_{j=1}^{d_B} c_{ij} |i\rangle^A \otimes |j\rangle^B.
\]

In order to equip the tensor product space with more structure, we need a precise way to relate elements in \( \mathcal{H}^A \) and \( \mathcal{H}^B \) to elements in \( \mathcal{H}^A \otimes \mathcal{H}^B \). The **tensor product** is a map \( \otimes : \mathcal{H}^A \times \mathcal{H}^B \rightarrow \mathcal{H}^A \otimes \mathcal{H}^B \) that generates the **tensor product vector** \( |\psi\rangle^A \otimes |\phi\rangle^B \in \mathcal{H}^A \otimes \mathcal{H}^B \) for \( |\psi\rangle^A \in \mathcal{H}^A \) and \( |\phi\rangle^B \in \mathcal{H}^B \), such that

1. \( (c|\psi\rangle^A) \otimes |\phi\rangle^B = |\psi\rangle^A \otimes (c|\phi\rangle^B) = c(|\psi\rangle^A \otimes |\phi\rangle^B) \quad \forall |\psi\rangle^A \in \mathcal{H}^A, |\phi\rangle^B \in \mathcal{H}^B, c \in \mathbb{C}, \)
2. \( (|\psi\rangle^A + |\omega\rangle^A) \otimes |\phi\rangle^B = |\psi\rangle^A \otimes |\phi\rangle^B + |\omega\rangle^A \otimes |\phi\rangle^B \quad \forall |\psi\rangle^A, |\omega\rangle^A \in \mathcal{H}^A, |\phi\rangle^B \in \mathcal{H}^B, \)
3. \( |\psi\rangle^A \otimes (|\phi\rangle^B + |\omega\rangle^B) = |\psi\rangle^A \otimes |\phi\rangle^B + |\psi\rangle^A \otimes |\omega\rangle^B \quad \forall |\psi\rangle^A \in \mathcal{H}^A, |\phi\rangle^B, |\omega\rangle^B \in \mathcal{H}^B. \)

These three properties are known as “bilinearity.” Hence, the tensor product is a bilinear map from \( \mathcal{H}^A \times \mathcal{H}^B \) to \( \mathcal{H}^A \otimes \mathcal{H}^B \).

One important consequence of bilinearity is that it allows us to form a tensor product basis for \( \mathcal{H}^A \otimes \mathcal{H}^B \) using any pair of bases for \( \mathcal{H}^A \) and \( \mathcal{H}^B \). Indeed suppose that \( \{ |\tilde{i}\rangle^A \}_{i=1}^{d_A} \) and \( \{ |\tilde{j}\rangle^B \}_{j=1}^{d_B} \) are
arbitrary orthonormal bases for $\mathcal{H}^A$ and $\mathcal{H}^B$ respectively. Then we can write the computational basis vectors in these new bases by

$$| i \rangle^A = \sum_{j=1}^{d_A} a_{ij} | j \rangle^A \quad \text{for } i = 1, \ldots, d_A$$

$$| i \rangle^B = \sum_{j=1}^{d_B} b_{ij} | j \rangle^B \quad \text{for } i = 1, \ldots, d_B.$$ 

Substituting these into Eq. (53) and using bilinearity gives

$$| \psi \rangle = \sum_{i=1}^{d_A} \sum_{j=1}^{d_B} c_{ij} | i \rangle^A \otimes | j \rangle^B,$$

where $c_{ij} = \sum_{k=1}^{d_A} \sum_{l=1}^{d_B} a_{ik} b_{lj} c_{kl}$. Hence, $\{| i \rangle^A \otimes | j \rangle^B \}_{i=1,j=1}^{d_A,d_B}$ also provides a basis for $\mathcal{H}^A \otimes \mathcal{H}^B$.

For notation simplicity, we will drop the superscript labels $A$ and $B$ on the kets when the associated Hilbert spaces are clear. Also, the tensor product of two vectors $| \psi \rangle \otimes | \phi \rangle$ is often written as $| \psi \rangle | \phi \rangle$, or even just $| \psi \phi \rangle$.

It should be stressed that the tensor product space $\mathcal{H}^A \otimes \mathcal{H}^B$ is much larger than the set of tensor product vectors. The latter is given by $S = \{| \alpha \rangle \otimes | \beta \rangle : | \alpha \rangle \in \mathcal{H}^A, | \beta \rangle \in \mathcal{H}^B \}$ while the former consists of all linear combinations of vectors in $S$. We have $S \subset \mathcal{H}^A \otimes \mathcal{H}^B$, but $S \neq \mathcal{H}^A \otimes \mathcal{H}^B$ since not every $| \psi \rangle \in \mathcal{H}^A \otimes \mathcal{H}^B$ is a tensor product of two vectors. That is, $| \psi \rangle \notin | \alpha \rangle \otimes | \beta \rangle$ for most $| \psi \rangle \in \mathcal{H}^A \otimes \mathcal{H}^B$. We will later see that such vectors $| \psi \rangle$ correspond to entangled states in quantum states.

We next discuss the structure of linear operators acting on $\mathcal{H}^A \otimes \mathcal{H}^B$. If $A \in L(\mathcal{H}^A, \mathcal{H}^A')$ and $B \in L(\mathcal{H}^B, \mathcal{H}^B')$, then their tensor product $A \otimes B$ is a linear operator $\mathcal{H}^A \otimes \mathcal{H}^B \to \mathcal{H}^A \otimes \mathcal{H}^B$ with action defined by

$$A \otimes B \left( \sum_{i,j} c_{ij} | i \rangle \otimes | j \rangle \right) = \sum_{i,j} c_{ij} A | i \rangle \otimes B | j \rangle.$$  \hspace{1cm} (55)

This action is linear such that if $A, C \in L(\mathcal{H}^A, \mathcal{H}^A')$ and $B, D \in L(\mathcal{H}^B, \mathcal{H}^B')$, then $A \otimes B + C \otimes D \in L(\mathcal{H}^A \otimes \mathcal{H}^B, \mathcal{H}^A' \otimes \mathcal{H}^B')$ with action

$$(A \otimes B + C \otimes D) | \psi \rangle = A \otimes B | \psi \rangle + C \otimes D | \psi \rangle$$  \hspace{1cm} (56)

for all $| \psi \rangle \in \mathcal{H}^A \otimes \mathcal{H}^B$. Furthermore, from Eq. (55) we see that the product of tensor product operators satisfies

$$(A \otimes B)(C \otimes D) = AC \otimes BD.$$  \hspace{1cm} (57)

One important example of Eq. (55) is known as a partial contraction. If $\mathcal{H}^A' = C$, then $A \in L(\mathcal{H}^A, C)$ is a bra acting on $\mathcal{H}^A$; i.e. $A = \langle \phi |$ for some $| \phi \rangle \in \mathcal{H}^A$. Letting $\mathcal{H}^B = \mathcal{H}^B'$, then the partial contraction of $| \phi \rangle$ on system $A$ is the operator $\langle \phi |^A \otimes I^B \in L(\mathcal{H}^A \otimes \mathcal{H}^B, \mathcal{H}^A) \otimes \mathcal{H}^B)$ such that for any $| \psi \rangle^{AB} = \sum_{i,j} c_{ij} | i \rangle^A | j \rangle^B$,

$$\langle \phi |^A \otimes I^B (| \psi \rangle^{AB}) = \langle \phi |^A \otimes I^B \left( \sum_{i,j} c_{ij} | i \rangle^A | j \rangle^B \right) = \sum_{ij} c_{ij} \langle \phi |^A \otimes I^B | j \rangle^B.$$  \hspace{1cm} (58)
This is called a partial contraction since it is essentially contracting the two spaces $\mathcal{H}^A$ and $\mathcal{H}^B$ down to the single space $\mathcal{H}^l$. Below in Sect. ?? we will encounter another type of partial contraction that is performed on linear operators themselves. Also note that if $|\gamma\rangle^A = \sum_{i,j} a_{ij}^a |i\rangle^A \otimes |j\rangle^B$ and $|\omega\rangle^A_B = \sum_{i,j} b_{ij}^{|i\rangle \otimes |j\rangle}$ are two vectors in $\mathcal{H}^A \otimes \mathcal{H}^B$, then their inner product represents a full contraction:

$$\langle \gamma | \omega \rangle = \left( \sum_{i,j} a_{ij}^a \langle i |^A \otimes \langle j |^B \right) \left( \sum_{k,l} b_{kl}^{|k\rangle^A \otimes |l\rangle^B} \right) = \sum_{i,j,k,l} a_{ij}^a b_{kl}^{|i\rangle^A \langle k |^B \otimes |l\rangle^B} = \sum_{i,j} a_{ij}^a b_{ij}.$$ \hspace{1cm} (59)

Next, recall that any operator $A \in \mathcal{L}(\mathcal{H}^A, \mathcal{H}^A')$ and any operator $B \in \mathcal{L}(\mathcal{H}^B, \mathcal{H}^B')$ can be expressed as

$$A = \sum_{i=1}^{d_A} \sum_{k=1}^{d_A} a_{ik}^{|i\rangle \otimes |k\rangle}, \quad B = \sum_{j=1}^{d_B} \sum_{l=1}^{d_B} b_{jl}^{|j\rangle \otimes |l\rangle}.$$ 

Likewise, any operator $K \in \mathcal{L}(\mathcal{H}^A \otimes \mathcal{H}^B, \mathcal{H}^A' \otimes \mathcal{H}^B')$ can be expressed as

$$K = \sum_{i=1}^{d_A} \sum_{k=1}^{d_A} \sum_{j=1}^{d_B} \sum_{l=1}^{d_B} c_{ijkl}^{|i\rangle \otimes |k\rangle \otimes |j\rangle \otimes |l\rangle}.$$ \hspace{1cm} (60)

Analogous to the case of vectors in $\mathcal{H}^A \otimes \mathcal{H}^B$, not every $K \in \mathcal{L}(\mathcal{H}^A \otimes \mathcal{H}^B, \mathcal{H}^A' \otimes \mathcal{H}^B')$ is a tensor product of operators from $\mathcal{L}(\mathcal{H}^A, \mathcal{H}^A')$ and $\mathcal{L}(\mathcal{H}^B, \mathcal{H}^B')$.

The matrix representation of tensor products is given by the Kronecker product of matrices. That is, if

$$A \doteq \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1d_A} \\ a_{21} & a_{22} & \cdots & a_{2d_A} \\ \vdots \\ a_{d_A} & a_{d_A} & \cdots & a_{d_Ad_A} \end{pmatrix}, \quad B \doteq \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1d_B} \\ b_{21} & b_{22} & \cdots & b_{2d_B} \\ \vdots \\ b_{d_B} & b_{d_B} & \cdots & b_{d_Bd_B} \end{pmatrix}$$ \hspace{1cm} (61)

are matrix representations for $A \in \mathcal{L}(\mathcal{H}^A, \mathcal{H}^A')$ and $B \in \mathcal{L}(\mathcal{H}^B, \mathcal{H}^B)$, then

$$A \otimes B \doteq \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & \cdots & a_{11}b_{1d_B} & a_{12}b_{11} & a_{12}b_{12} & \cdots & a_{12}b_{1d_B} & a_{1d_A}b_{11} & a_{1d_A}b_{12} & \cdots & a_{1d_A}b_{1d_B} \\ a_{11}b_{21} & a_{11}b_{22} & \cdots & a_{11}b_{2d_B} & a_{12}b_{21} & a_{12}b_{22} & \cdots & a_{12}b_{2d_B} & a_{1d_A}b_{21} & a_{1d_A}b_{22} & \cdots & a_{1d_A}b_{2d_B} \\ \vdots \\ a_{d_A}b_{11} & a_{d_A}b_{12} & \cdots & a_{d_A}b_{1d_B} & a_{d_A}b_{21} & a_{d_A}b_{22} & \cdots & a_{d_A}b_{2d_B} & a_{d_A}b_{d_B} & a_{d_A}b_{d_B} & \cdots & a_{d_A}b_{d_Bd_B} \end{pmatrix}.$$ \hspace{1cm} (62)

An easy way to remember this construction is in terms of block matrices:

$$A \otimes B \doteq \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1d_A}B \\ a_{21}B & a_{22}B & \cdots & a_{2d_A}B \\ \vdots \\ a_{d_A}B & \cdots & \cdots & a_{d_A}B \end{pmatrix}.$$ \hspace{1cm} (63)
Here, each element $a_{ij}B$ is a $d_2 \times d_2$ submatrix in which $a_{ij}$ is being multiplied to every element in $B$. For example, if

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

then

$$\sigma_x \otimes \sigma_z = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad \sigma_z \otimes \sigma_x = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}. \quad (65)$$

This way of representing tensor products also applies to kets and bras. For example,

$$|a\rangle = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{d_A} \end{pmatrix}, \quad |b\rangle = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{d_B} \end{pmatrix}, \quad \Rightarrow \quad |a\rangle \otimes |b\rangle = \begin{pmatrix} a_1 b_1 \\ a_1 b_2 \\ \vdots \\ a_{d_A} b_{d_B} \end{pmatrix}. \quad (66)$$

For a general vector $|\psi\rangle$ that is not necessarily a tensor product of two vectors, we have

$$|\psi\rangle = \sum_{i=1}^{d_A} \sum_{j=1}^{d_B} c_{ij} |i\rangle |j\rangle \Rightarrow |\psi\rangle = \begin{pmatrix} c_{11} \\ c_{12} \\ \vdots \\ c_{1d_B} \\ c_{21} \\ \vdots \\ c_{d_A 1} \end{pmatrix}. \quad (67)$$

The case of $\mathcal{H}^A \otimes \mathcal{H}^B = \mathbb{C}^2 \otimes \mathbb{C}^2$ is so important that we explicitly describe it here. From Eq. (10), we see that the tensor product basis vectors have the matrix representations

$$|00\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |01\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad |10\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad |11\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad (68)$$

Then

$$|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle \Rightarrow |\psi\rangle = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}. \quad (69)$$
4 Exercises

Exercise 1
Suppose that $\mathcal{H} = \mathbb{C}^3$ and let $V$ be the one-dimensional subspace spanned by $|\psi\rangle = \sqrt{1/3}(|0\rangle + e^{i/3}|1\rangle - |2\rangle)$. Provide an orthonormal basis for $V^\perp$.

Exercise 2
Compute the eigenvalues and eigenvectors of the three matrices:
\[
\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\]

Exercise 3
For a $d$-dimensional Hilbert space, compute the eigenvalues and dimensions of the associated eigenspaces for the operators $R_1 = |\beta\rangle\langle\alpha|$ and $R_2 = I - |\beta\rangle\langle\alpha|$ when
(a) $|\alpha\rangle = |\beta\rangle$;
(b) $|\alpha\rangle$ and $|\beta\rangle$ are orthogonal.
In both cases assume that $|\alpha\rangle$ and $|\beta\rangle$ are normalized.

Exercise 4
Suppose that $\{|e_i\rangle\}_{i=1}^s$ and $\{|\delta_i\rangle\}_{i=1}^s$ are both orthonormal bases for a subspace $V \in \mathcal{H}$. Show that
\[
\sum_{i=1}^s |e_i\rangle\langle e_i| = \sum_{i=1}^s |\delta_i\rangle\langle \delta_i| = P_S.
\]

Exercise 5
(a) Give an example of a normal operator that is not hermitian.
(b) Give an example of a unitary operator that is not hermitian.
(c) Give an example of a hermitian operator that is not positive.

Exercise 6
Prove Prop. 5. That is, show that a normal operator $A$ is positive iff $\langle \psi | A | \psi \rangle \geq 0$ for all $|\psi\rangle \in \mathcal{H}$.

Exercise 7
Consider the operator $R : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ given by
\[
R = |0\rangle\langle 0| + |1\rangle\langle 1|,
\]
where $|+\rangle = \sqrt{1/2}(|0\rangle + |1\rangle)$. Find the singular value decomposition of $R$. 
Exercise 8

For $\mathcal{H}^A \otimes \mathcal{H}^B \cong \mathbb{C}^2 \otimes \mathbb{C}^2$, consider the operator $Y \in \mathcal{L}(\mathcal{H}^A \otimes \mathcal{H}^B)$ given by

$$Y = I \otimes \sigma_y - \sigma_y \otimes I,$$

where $\sigma_y$ is defined in Exercise 2.

(a) What is the matrix representation of $Y$ in the computational basis?

(b) Compute $Y|\psi\rangle$ with $|\psi\rangle$ having the general form $|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle$.

(c) Find a tensor product vector that is an eigenvector of $Y$ (Hint: Use your answer from Exercise 2).

Exercise 9

For operators $X$ and $Y$, compute their two partial traces $\text{Tr}_A$ and $\text{Tr}_B$.

(a) $X = |\Phi^+_2\rangle\langle \Phi^+_2| + |0\rangle\langle 0| \otimes |1\rangle\langle 1|$, where $|\Phi^+_2\rangle = \sqrt{1/2}(|00\rangle + |11\rangle)$;

(b) $Y = F_d$, where $F = \sum_{i,j=1}^d |ij\rangle\langle ji|$ is the $d$-dimensional SWAP operator.

References
