Lecture 2: State space and vector spaces

Rajat Mittal

IIT Kanpur

We will start by looking at the first question? How do we represent data/information stored inside a quantum computer/system. In other words, how do we mathematically capture the state of a quantum computer at any instance of the computation.

Exercise 1. How do we capture the state for a classical computer?

Simply, the state of a computer can be represented by a sequence of 0's and 1's, where each such number is called a *bit*.

For the quantum case, the answer is given by postulates of quantum mechanics (first and fourth to be particular). Intuitively, they express the state of a quantum system in terms of linear algebra. Actually the state space of probabilistic computation can naturally be viewed in terms of linear algebra too. As a bonus, the states of classical computation (sequence of 0 and 1) can be expressed in the same terms with more restrictions.

We will give a brief introduction to the basic linear algebraic concepts, though it is assumed that the reader is familiar with the concept of vector space, basis and linear independence. Strang's book [2] is a good source to brush up these concepts.

Dirac's notation is used widely in quantum computing to represent these linear algebraic quantities, because it simplifies the understanding of quantum mechanical concepts. We will switch between the standard vector notation and Dirac notation in these notes.

Exercise 2. Read about vector space, basis, and linear independence if you are not comfortable with these words.

1 Background for linear algebra

1.1 Vector spaces

One of the most fundamental concept in linear algebra is that of a *vector space*. You must have seen vector spaces over real numbers. The vector space $\mathbb{R}^n$ (dimension n) consists of vectors with n coordinates such that each coordinate is a real number.

You have seen the vector space $\mathbb{R}^n$, the vector space of dimension n over real numbers. For example, $\mathbb{R}^2$ is a vector space of dimension 2. All elements in this vector space can be written as a linear combination of two standard basis vectors,

$$e_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \;\; \& \;\; e_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

So, every vector $v \in \mathbb{R}^2$ can be written as $\alpha_0 e_0 + \alpha_1 e_1$, where α_0, α_1 are two complex numbers.

Suppose $\{v_1, v_2, \cdots, v_n\}$ is the basis of the vector space, then any vector v can be written as

$$v = a_1 v_1 + \cdots + a_n v_n,$$

where all a_i's are real numbers.

You have also seen the regular inner product defined for $\mathbb{R}^n$, i.e. $v^T w = \sum_{i \in [n]} v_i w_i$. What is a unit vector?

For a vector space $V \in \mathbb{R}^n$, its orthogonal complement $V^\perp$ is defined as,

$$V^\perp := \{w \in \mathbb{R}^n : \; v^T w = 0 \;\; \forall v \in V\}.$$

1.2 Operators

A quantum computer will not be very useful if it remains in the same state. It needs to move to desired states by application of transformations (like gates in a classical computer). Since states are vectors, and all transformations in quantum are linear, it is time to study linear operators on vectors.

Given two vector spaces, V and W over $\mathbb{C}$, a *linear* operator $M : V \to W$ is defined as an operator satisfying the following properties.

- $M(x + y) = M(x) + M(y)$.
- $M(\alpha x) = \alpha M(x), \ \forall \alpha \in \mathbb{C}$.

These conditions imply that the *zero* of the vector space V is mapped to the *zero* of the vector space W. Also,

$$M(\alpha_1 x_1 + \cdots + \alpha_k x_k) = \alpha_1 M(x_1) + \cdots + \alpha_k M(x_k)$$

Where $x_1, \cdots, x_k$ are elements of V and α_i's are in $\mathbb{C}$. Because of this linearity, it is enough to specify the value of a linear operator on any basis of the vector space V. In other words, a linear operator is uniquely defined by the values it takes on any particular basis of V.

Let us define the addition of two linear operators as $(M + N)(u) = M(u) + N(u)$. Similarly, αM (scalar multiplication) is defined to be the operator $(\alpha M)(u) = \alpha M(u)$. The space of all linear operators from V to W (denoted $L(V, W)$) is a vector space in itself. The space of linear operators from V to V will be denoted by $L(V)$.

Exercise 3. Given the dimension of V and W, what is the dimension of the vector spaces $L(V, W)$?

One of the issue is, how to represent a linear operator. You can give its action on all the elements in vector space, but that will take infinite space.

1.3 Matrices as linear operators

Given two vector spaces $V = \mathbb{C}^n, W = \mathbb{C}^m$ and a matrix M of dimension $m \times n$, the operation $v \in V \to Mv \in W$ is a linear operation. Here, if $v = a_1 v_1 + \cdots + a_n v_n$, we represent it as a vector,

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

So, a matrix acts as a linear operator on the corresponding vector space.

To ask the converse, can any linear operator be specified by a matrix?

Let f be a linear operator from a vector space V (dimension n) to a vector space W (dimension m). Suppose $\{e_1, e_2, \cdots, e_n\}$ is a basis for the vector space V. Denote the images of elements of this basis under f as $\{w_1 = f(e_1), w_2 = f(e_2) \cdots, w_n = f(e_n)\}$.

Exercise 4. What is the lower-bound/ upper-bound on the dimension of the vector space spanned by $\{w_1, w_2, \cdots, w_n\}$?

Define M_f to be the matrix with columns $w_1, w_2, \cdots, w_n$. Notice that M_f is a matrix of dimension $m \times n$. It is a simple exercise to verify that the action of the matrix M_f on a vector $v \in V$ is just $M_f v$. Here we assume that v is expressed in the chosen basis $\{e_1, e_2, \cdots, e_n\}$.

Exercise 5. Convince yourself that Mv is a linear combination of columns of M.

Notice that the matrix M_f and the operator f act exactly the same on the basis elements of V. Since both the operations are linear, they are exactly the same operation. This proves that any linear operation can be specified by a matrix.

The equivalence of matrices and linear operators does not depend upon the chosen basis. We can pick our favorite bases of V and W, and the linear operator can similarly be written in the new basis as a matrix (The columns of this matrix are images of the basis elements of V). In other words, given bases of V and W and a linear operator f, it has a unique matrix representation.

To compute the action of a linear operator, express $v \in V$ in the preferred basis and multiply it with the matrix representation. The output will be in the chosen basis of W. We will use the two terms, linear operator and matrix, interchangeably in future (bases will be clear from the context).

For a matrix A, A^T denotes the transpose of the matrix and A^* denotes the adjoint of the matrix (take complex conjugate and then transpose).

Exercise 6. Why is matrix multiplication defined the way it is? Why can't it be defined in the more natural way of entry-wise multiplication?

Let us look at some simple matrices which will be used later.

- Zero matrix: The matrix with all the entries 0. It acts trivially on every element and takes them to the 0 vector.
- Identity matrix: The matrix with 1's on the diagonal and 0 otherwise. It takes $v \in V$ to v itself.
- All 1's matrix (J): All the entries of this matrix are 1.

Exercise 7. What is the action of matrix J?

1.4 Extra reading: Kernel, image and rank

For a linear operator/matrix (from V to W), the *kernel* is defined to be the set of vectors which map to 0.

$$ker(M) = \{x \in V : Mx = 0\}$$

Here 0 is the zero vector in space W.

Exercise 8. What is the kernel of the matrix J?

The *image* is the set of vectors which can be obtained through the action of the matrix on some element of the vector space V.

$$img(M) = \{x \in W : \exists y \in V,\ x = My\}$$

Exercise 9. Show that $img(M)$ and $ker(M)$ are subspaces.

Exercise 10. What is the image of J?

Notice that $ker(M)$ is a subspace of V, but $img(M)$ is a subspace of W. The dimension of $img(M)$ is known as the *rank* of M ($rank(M)$). The dimension of $ker(M)$ is known as the nullity of M ($nullity(M)$). For a matrix $M \in L(V, W)$, by the famous rank-nullity theorem,

$$rank(M) + nullity(M) = dim(V).$$

Here $dim(V)$ is the dimension of the vector space V.

Proof. Suppose $u_1, \cdots, u_k$ is the basis for $ker(M)$. We can extend it to the basis of V: $u_1, \cdots, u_k, v_{k+1}, \cdots, v_n$. We need to prove that the dimension of $img(M)$ is $n - k$. It can be proved by showing that the set $\{Mv_{k+1}, \cdots, Mv_n\}$ forms a basis of $img(M)$.

Exercise 11. Prove that any vector in the image of M can be expressed as linear combination of $Mv_{k+1}, \cdots, Mv_n$. Also any linear combination of $Mv_{k+1}, \cdots, Mv_n$ can't be zero vector.

□

Given a vector v and a matrix M, it is easy to see that the vector Mv is a linear combination of columns of M. To be more precise, $Mv = \sum_i M_i v_i$ where M_i is the ith column of M and v_i is the ith co-ordinate of v. This implies that any element in the image of M is a linear combination of its columns.

Exercise 12. Prove the rank of a matrix is equal to the dimension of the vector space spanned by its columns (column-space).

The dimension of the column space is sometimes referred as the *column-rank*. We can similarly define the *row-rank*, the dimension of the space spanned by the rows of the matrix. Luckily, row-rank turns out to be equal to column-rank and we will call both of them as the rank of the matrix. This can be proved easily using *Gaussian elimination*.

Operations on matrices Lets look at some of the basic operations on these matrices.

– Trace: The *trace* of a square matrix A is the sum of all the diagonal elements.

$$tr(A) = \sum_i A[i,i]$$

– Entry-wise multiplication: The entry-wise multiplication of two matrices is known as *Hadamard product* and only makes sense when both of them have same number of rows and columns. The Hadamard product of two matrices A, B is

$$(A \circ B)[i,j] = A[i,j]B[i,j].$$

The related operation is when you add up the entries of this Hadamard product.

$$(A \bullet B) = \sum_{i,j} A[i,j]B[i,j]$$

Notice that $A \bullet B$ is a scalar and not a matrix.

Exercise 13. Given a matrix, express $\bullet$ operation in terms of multiplication and trace operation.

– Inverse: Inverse of a matrix M is the matrix M^{-1}, s.t., $MM^{-1} = M^{-1}M = I$. The inverse only exists if the matrix has full rank (columns of M span the whole space).

Exercise 14. What is the inverse of matrix J (all 1's matrix).

Exercise 15. Show that the inverse of a matrix exists iff it has full rank.

2 State of a classical (deterministic) and probabilistic system

For this section, we will again get back to spaces over real numbers, familiar territory to you. A real vector space with dimension n is generally denoted as $\mathbb{R}^n$.

The first task is to represent the state of a classical computer. For the deterministic model, no randomness is used in the algorithm and the final output only depends upon the input. For the randomized model, we have a random string too as part of the input and the output criteria is relaxed.

Let us take the easier case first. How would you represent the state of a deterministic algorithm? As mentioned before it is a string of length m, where m is the total number of bits used by the algorithm. In particular, there are 2^m choices and we will be in any one of the choices at a particular step of the algorithm. One *expensive* way to capture this state would be to have a 2^m dimensional vector and keep the single entry 1 (corresponding to the state at that moment).

For example, if $m = 2$, we will have four dimensions corresponding to $00, 01, 10, 11$, the four possible states of the system. If we are in state 01, it will be represented by the vector $|s\rangle\rangle = (0,1,0,0)^T$. Here, we have used $|\rangle\rangle$ to denote a classical state (later $|\rangle$ will be used to denote the quantum state).

Exercise 16. What will be the state vector if we are in state 11?

Since are states are vectors now, the operations will be matrices. In particular, a operation takes one string to another. So the $2^m \times 2^m$ matrix will have the property that every column will have exactly one 1 and rest of the entries will be 0.

Exercise 17. What will be the matrix corresponding to the operation which takes all strings to 00. What will be the matrix for AND/OR gates.

One of the important class of operations for us (later in quantum part) is the set of reversible operations. These are the operations where we can find the state before the operation from the state after the operation. In other words, there is an inverse.

Exercise 18. Convince yourself that the matrix corresponding to a reversible operation is going to be a permutation matrix.

Even though the discussion makes sense, it still begs a question: why are we using such an expensive representation, the state vector and operation matrices are very sparse. It is because we will extend this formulation to randomized computing and then give a similar formulation for quantum computing.

How would we represent the state of a randomized algorithm? If we wanted the state of $A(x, r)$, it is a deterministic algorithm and we already know how to capture its state. Notice that we are interested in the behavior of the algorithm on input x; in randomized case, it means the cumulative behavior over all r's. Our state of the algorithms should capture this cumulative behavior and not just the behavior on each r individually.

Again, let us assume that there are m bits of space on the computer. In the randomized case, after some steps of the algorithm, the data in the computer can be one of the 2^m strings with some probabilities.

To take an example, let us say that space $m = (m_1, m_2)$ has length 2 and our algorithm tosses one fair random coin. This means r has length 1 and it is 0 or 1 with equal probability. The algorithm can be stated as:

– Toss the fair random coin, say its value is r.
– If $r = 0$, then m_1 is replaced with $m_1 \cup m_2$ (apply OR function on m_1 and m_2).
– If $r = 1$, then m_2 is replaced with $\neg(m_1) \cap r$ (apply AND function on $m_1 and r$).

Suppose the starting state is 00, then the starting state of the system should be $|s\rangle\rangle = (1, 0, 0, 0)$. After one step, the state could be 00 with probability 1/2 (if $r = 0$), or 01 with probability 1/2 (if $r = 1$). So the resulting state would be $|s\rangle\rangle = (1/2, 1/2, 0, 0)$.

Suppose the starting state is 01, then the starting state of the system should be $|s\rangle\rangle = (0, 1, 0, 0)$. After one step, the state could be 11 with probability 1/2 (if $r = 0$), or 01 with probability 1/2. (if $r = 1$). So the resulting state would be $|s\rangle\rangle = (0, 1/2, 0, 1/2)$.

Suppose the starting state is 11, then the starting state of the system should be $|s\rangle\rangle = (0, 1, 0, 0)$. After one step, the state could be 11 with probability 1/2 (if $r = 0$), or 10 with probability 1/2. (if $r = 1$). So the resulting state would be $|s\rangle\rangle = (0, 0, 1/2, 1/2)$.

Exercise 19. Show that the resulting state with 10 as input would be $|s\rangle\rangle = (0, 0, 1, 0)$.

To summarize, the entry in front of the coordinate $b \in \{0, 1\}^m$ represents the probability that the space of the computer has that string at that moment. The usual deterministic operations are still allowed, though there are more operations possible. For example, the small step of tossing a coin can be written as (just for one bit),

$$M = \frac{1}{2} \begin{pmatrix} 1 \, 1 \\ 1 \, 1 \end{pmatrix}.$$

In general, since we move from one probability distribution to another, all entries of the matrix should be positive and every column sum should be 1. You will prove that these restrictions are necessary and sufficient in the assignment.

3 State of a quantum system

3.1 Complex Euclidean vector spaces

We will mostly concern ourselves with the vector space $\mathbb{C}^n$ in quantum computing, the vector space of dimension n over the field of complex numbers. The vectors in $\mathbb{C}^n$ are going to be our states of a quantum computer (with more constraints). This means that the scalars used in these vector spaces are complex numbers, i.e., every coordinate will contain a complex number. For example, $\mathbb{C}^2$ is a vector space of dimension 2.

All elements in this vector space can be written as a linear combination of two standard basis vectors,

$$e_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \& \quad e_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

So, every vector $v \in \mathbb{C}^2$ can be written as $\alpha_0 e_0 + \alpha_1 e_1$, where α_0, α_1 are two complex numbers. In Dirac's notation these two standard basis vectors are written as $e_0 = |0\rangle$ and $e_1 = |1\rangle$.

A column vector is the most basic unit of a vector space. Using Dirac's notation, a column vector will be denoted by $|\psi\rangle$. Suppose $\{|v_1\rangle, |v_2\rangle, \cdots, |v_n\rangle\}$ is the basis of the vector space, then any vector $|v\rangle$ can be written as

$$|v\rangle = a_1|v_1\rangle + \cdots + a_n|v_n\rangle,$$

where all a_i's are complex numbers.

For a vector space with dimension n, the standard basis is denoted by $|0\rangle, |1\rangle, \cdots, |n-1\rangle$. Here you can think of $|i\rangle$ as the vector with 1 at the $(i+1)$-th position and 0 otherwise. For example, a 3-dimensional space will have standard basis elements $|0\rangle, |1\rangle$ and $|2\rangle$.

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \& \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \& \quad |2\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

The notation $\langle\psi|$ denotes the row vector whose entries are complex conjugate of the entries of the vector $|\psi\rangle$ (also known as the *adjoint*, $|\psi\rangle = \langle\psi|^*$). If $|\psi\rangle = (x_1\ x_2\ \cdots\ x_n)^T$ and $|\phi\rangle = (y_1\ y_2\ \cdots\ y_n)^T$, the vector space $\mathbb{C}^n$ is equipped with the natural inner product (like dot product),

$$\langle\psi|\phi\rangle = ((x_1, x_2, \cdots, x_n), (y_1, y_2, \cdots, y_n)) = \sum_{i=1}^{n} x_i^* y_i.$$

Here x^* denotes the complex conjugate of a complex number x. In usual vector notation, the inner product between ψ and ϕ will be denoted by $\psi^*\phi$.

This gives the usual notion of distance and lengths. Hence, the complex vector space armed with this inner product is called the *complex Euclidean vector space.*

Remember that a bit in a classical computer is a number in $\{0, 1\}$. Analogously, a *qubit* in a quantum computer is going to be a unit vector in $\mathbb{C}^2$.

Exercise 20. What is a unit vector in this space?

In other words, the state of a qubit can be written as $\alpha|0\rangle + \beta|1\rangle$, where $\alpha^2 + \beta^2 = 1$. You can think of α, β as weights on classical states $|0\rangle, |1\rangle$ respectively. Notice that we did not say probability, more on it later.

Exercise 21. Can you come up with a state of qubit which is not classical? Are your α, β real numbers?

Exercise 22. What is the difference between vector $|0\rangle$ and vector 0?

There is a small difference between vector $|0\rangle$ and vector 0 (the vector with all entries 0). First one is a basis vector with the first entry 1 and rest 0.

In the beginning, you can convert expressions in Dirac notation to the usual vector notation. Slowly, it might become easier to directly manipulate expressions in Dirac notation.

Exercise 23. Suppose a qubit is in state $|\psi\rangle$, where $\psi^T = (1/\sqrt{2}, -i/\sqrt{2})$. Is it a valid qubit state? Write it in Dirac notation.

Exercise 24. Can you think of a linear operator which takes $|0\rangle$ to $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|1\rangle$ to $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$?

In Dirac's notation, action of M on a vector $|v\rangle$ is just $M|v\rangle$. Also, $\langle u|Mv\rangle = \langle M^*u|v\rangle$ is denoted by $\langle u|M|v\rangle$.

In Dirac's notation, we denoted inner product between two vectors $|\psi\rangle, |\phi\rangle$ by $\langle\psi|\phi\rangle$. The expression $A = |\psi\rangle\langle\phi|$ is a matrix which takes $|v\rangle$ to $\langle\phi|v\rangle|\psi\rangle$. The analog of this expression in the simple vector notation would be, $A = \psi\phi^*$. If a linear operator M takes an orthonormal basis $\{v_1, v_2, \cdots, v_n\}$ to vectors $\{w_1, w_2, \cdots, w_n\}$, then the matrix representation of M is

$$M = \sum_{i=1}^{n} |w_i\rangle\langle v_i|.$$

Exercise 25. What will be $(|w\rangle\langle v|)|v\rangle$. Prove that M defined in the equation above will have the action as intended.

3.2 First postulate: state of a quantum system

The postulates of quantum mechanics provide us the mathematical formalism over which the physical theory is developed. For people studying quantum computing, it gives the basic laws according to which any quantum system (or a quantum computer) works.

These postulates were agreed upon after a lot of trial and error. We won't be concerned about the physical motivation of these postulates. Most of the material for this lecture is taken from [1]. It is a very good reference for more details.

As discussed before, in this lecture note, we are interested in postulates which allow us to represent data/information in a quantum computer (the state of a quantum computer at an instance). The first postulates specifies, what is meant mathematically by the state of a quantum system.

Postulate 1: A physically isolated system is associated with a Hilbert space, called the *state space* of the system. The system, at a particular time, is completely described by a unit vector in this Hilbert space, called the state of the system.

Intuitively, Hilbert space is a vector space with enough structure so that we can apply the techniques of linear algebra and analysis on it.

Exercise 26. Read more about Hilbert spaces.

For this course, we will only be dealing with vector spaces over complex numbers with inner product defined over them. In almost all these cases, the dimension is going to be finite (say n). In particular, we will assume that our state space is $\mathbb{C}^n$ for some n (n is the dimension of this state space).

The simplest non-trivial state space would be $\mathbb{C}^2$ (dimension being 2), the state space of a qubit. Remember that a qubit is a generalization of bit, the way we store information in a classical computer. It will be spanned by two standard basis vectors, $|0\rangle$ and $|1\rangle$. These two, $|0\rangle$ and $|1\rangle$, represent classical states in a quantum computer. For an example of a non-classical state, take $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ or $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. These states, $|+\rangle$ and $|-\rangle$, are going to be pretty useful in quantum computation.

Exercise 27. Find another basis of $\mathbb{C}^2$.

Any state in this system can be written as,

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad |\alpha|^2 + |\beta|^2 = 1.$$

The coefficients, α and β, are called the amplitude. Specifically, α (β) is the amplitude of the state $|\psi\rangle$ for $|0\rangle$ ($|1\rangle$) respectively. When α and β are non-zero, we say that $|\psi\rangle$ is in *superposition* of states $|0\rangle$ and $|1\rangle$.

The property of *superposition* seems to be one of the major reasons behind the power of quantum computing (other is entanglement, described later). It allows us to compute aggregate properties of an input much faster than the classical computer (we will see this in action very soon, Deutsch's algorithm).

Many people interpret this as, the state $|\psi\rangle$ is in state $|0\rangle$ with probability $|\alpha|^2$ and in state $|1\rangle$ with probability $|\beta|^2$.

Note 1 (Important:). Even though we will formally study measurements later, it is helpful to introduce one simple measurement. Given the state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, the standard measurement will output $|0\rangle$ with probability $|\alpha|^2$ and in state $|1\rangle$ with probability $|\beta|^2$. Many people interpret this as, the state $|\psi\rangle$ is in state $|0\rangle$ with probability $|\alpha|^2$ and in state $|1\rangle$ with probability $|\beta|^2$. This will happen only if we measure the state right away and do not apply any operations (we will learn about measurement formally later). Hence, being in state $|0\rangle$ with probability $|\alpha|^2$ and in state $|1\rangle$ with probability $|\beta|^2$ is only a consequence of $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ and not equivalent to it. See Exercise 47 for more clarification.

Exercise 28. Suppose $\frac{1}{3}((1+2i)|0\rangle + (ai)|1\rangle)$ is a quantum state. What is the value of a?

We have talked about the case when $\mathbb{C}^2$ is our Hilbert space, and that case will be very useful. Though, in general, if there are n different classical states, the quantum state would be a unit vector expressed in an orthonormal basis $\{|0\rangle, |1\rangle, \cdots, |n-1\rangle\}$.

Remember that the standard basis is one of the most convenient basis to represent a state, but definitely not *the* only basis to represent a state. We can have any basis $\{|v_0\rangle, |v_1\rangle, \cdots, |v_n - 1\rangle\}$ and a state $|\psi\rangle$ can be written as,

$$|\psi\rangle = \alpha_0|v_0\rangle + \alpha_1|v_1\rangle + \cdots + \alpha_{n-1}|v_{n-1}\rangle, \quad \sum_i |\alpha_i|^2 = 1.$$

We will say that the state $|\psi\rangle$ is in superposition of basis states $\{|v_0\rangle, |v_1\rangle, \cdots, |v_n - 1\rangle\}$ (ideally using only those states whose amplitude is non-zero).

You might already guess from the discussion that the operators on these quantum states will be matrices (linear operators over the vector space). It will turn out that not all linear operators are allowed. Though, that will be discussed in the next lecture (second postulate).

Before that, we have to look at one more postulate. Notice that the state of a quantum computer is described by multiple bits (not just a single bit). How can we describe state of multiple qubits? That takes us to the concept of tensor product.

4 Tensor product

We have described the state of a system as a vector in a Hilbert space. What happens if we have multiple systems. For a classical computer, the answer is pretty simple, you just describe the state of both systems independently.

Interestingly, we can have a state of a composite *quantum* systems such that the individual state of the constituent systems can't be described. This property is known as entanglement and is the reason behind many weird properties of quantum mechanics. To understand this phenomenon, we need to understand the concept of *tensor products.*

Suppose there is a ball which can be colored blue or red. The state of a *quantum* ball is a unit vector in two dimensions,

$$|v\rangle = \alpha|r\rangle + \beta|b\rangle.$$

Where $|r\rangle, |b\rangle$ represent the classical states, the ball being red or blue, the coefficients α, β follow the usual law.

How about if there are two different balls. The classical states possible are $|rr\rangle, |rb\rangle, |br\rangle, |bb\rangle$, i.e., we take the set multiplication of possible states of individual system.

What are the possible states if this system is quantum? By analogy with one quantum ball case, any linear superposition of these classical states (normalized) should be a possible quantum state:

$$|v\rangle = \alpha|rr\rangle + \beta|rb\rangle + \gamma|br\rangle + \delta|bb\rangle,$$

where $|v\rangle$ is a unit vector.

This idea motivates the definition of tensor product. Given two vector spaces V, W equipped with an inner product and spanned by the orthonormal basis $v_1, v_2, \cdots, v_n$ and $w_1, w_2, \cdots, w_m$, the tensor product $V \otimes W$ is the space spanned by the mn vectors $(v_1 \otimes w_1), \cdots, (v_1 \otimes w_n), (v_2 \otimes w_1), \cdots, (v_n \otimes w_m)$.

Exercise 29. What is the dimension of space $V \otimes W$?

Formally, tensor product of two vector spaces is equipped with a bilinear map $\otimes : V \times W \to V \otimes W$ which satisfies the following conditions.

- Scalar multiplication: for $\alpha \in \mathbb{C}, v \in V$ and $w \in W$,

$$\alpha(v \otimes w) = (\alpha v) \otimes w = v \otimes (\alpha w).$$

- Linearity in the first component: for $v_1, v_2 \in V$ and $w \in W$,

$$(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w.$$

- Linearity in the second component: for $v \in V$ and $w_1, w_2 \in W$,

$$v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2.$$

The vector in the tensor product space, $|\psi\rangle \otimes |\phi\rangle$, will be simply written as $|\psi\rangle|\phi\rangle$ in the Dirac notation.

For our purposes, we can define the tensor product of two vectors in a canonical way for the vector spaces $\mathbb{C}^n$ and $\mathbb{C}^m$. The tensor product of two vectors $a = (a_1, \cdots, a_n) \in V$ and $b = (b_1, \cdots, b_m)$ is the vector $a \otimes b \in V \otimes W = \mathbb{C}^{mn}$,

$$a \otimes b = \begin{pmatrix} a_1 b \\ a_2 b \\ \vdots \\ a_n b \end{pmatrix} = \begin{pmatrix} a_1 b_1 \\ a_1 b_2 \\ \vdots \\ a_1 b_m \\ \vdots \\ \vdots \\ a_n b_1 \\ \vdots \\ a_n b_m \end{pmatrix}$$

In Dirac's notation, we further simplify $|v\rangle \otimes |w\rangle$ to $|vw\rangle$ when v and w are symbols. For example, state $|0\rangle \otimes |1\rangle$ is same as $|0\rangle|1\rangle = |01\rangle$.

Exercise 30. What is $\left(\frac{1}{2}|0\rangle + \frac{\sqrt{3}}{2}|1\rangle\right) \otimes |-\rangle$.

Exercise 31. What is the basis of state of two qubits? What about three qubits?

Exercise 32. Show that $(v + w) \otimes (a + b) = v \otimes a + v \otimes b + w \otimes a + w \otimes b$.

We can define the inner product on the tensor product space in the natural way,

$$\langle a \otimes b | c \otimes d \rangle = \langle a | c \rangle \langle b | d \rangle. \tag{1}$$

In other words, we have defined the tensor product between two Hilbert spaces V and W. First, look at them as vector spaces and define the vector space $V \otimes W$. Then, we define the inner product on $V \otimes W$ by the equation above (Eq. 1).

Carrying the analogy further, given two linear operators $A \in L(V)$ and $B \in L(W)$, their tensor product $A \otimes B$ can be defined in the space $L(V \otimes W)$. Precisely, its action is specified by,

$$(A \otimes B)(a \otimes b) = Aa \otimes Bb.$$

We can extend this by linearity to define the action on the complete space $V \otimes W$.

Exercise 33. Write out the matrix representation of $H^{\otimes 2} = H \otimes H$ where $H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ is the Hadamard Matrix. What is $H^{\otimes 2} \left(\frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) \right)$?

Exercise 34. Given the matrix representation of A, B; come up with the matrix representation of $A \otimes B$.

You can easily verify the following properties,

- $(A \otimes B)(C \otimes D) = AC \otimes BD$
- $(A \otimes B)^* = A^* \otimes B^*$

$A \otimes B$ are linear operators in $L(V \otimes W)$. Can there be other linear operators?

The sum of two linear operators is a linear operator. So, any operator of the form $\sum_i c_i (A_i \otimes B_i)$ is also a linear operator.

Are there any more linear operators? It turns out that these are the only linear operators in $L(V \otimes W)$, you can prove this by dimensionality argument.

The description of tensor product is given in a very simplified manner in terms of basis vectors, sufficient for use in our course. Readers are encouraged to check out the formal definitions.

5 Fourth postulate: composite Systems

The fourth postulate deals with composite systems, systems with more than one part. We will cover second and third postulates in the later sections. We will use tensor products for the sake of describing multiple systems.

Postulate 4: Suppose the state space of Alice is H_A and Bob is H_B, then the state space of their combined system is $H_A \otimes H_B$. If Alice prepares her system in state $|a\rangle$ and Bob prepares it in $|b\rangle$, then the combined state is $|a\rangle \otimes |b\rangle$, succinctly written as $|ab\rangle$.

Exercise 35. If Alice's qubit is in state $\frac{1}{2}|0\rangle + \frac{\sqrt{3}}{2}|1\rangle$ and Bob's qubit is in $\frac{1}{\sqrt{2}}|0\rangle - \frac{i}{\sqrt{2}}|1\rangle$, what is the state of the combined system?

Similarly, if operator A is applied on Alice's system and operator B is applied on Bob's system, then operator $A \otimes B$ is applied to the combined system. This follows from the property,

$$(A \otimes B)(|a\rangle \otimes |b\rangle) = A|a\rangle \otimes B|b\rangle.$$

Generally, it is quite clear which part of the system belongs to which party. In case of confusion, we will use subscripts to resolve it. So if A is an operator on first system and B is an operator on second system, the combined operator is $A_1 \otimes B_2$.

The most useful example will be of k qubits.

Exercise 36. What will be the dimension of the state space of k qubits?

The state space for k qubits is $\mathbb{C}^{2^k}$. It is a vector space of dimension 2^k. A natural way to represent this vector space is through basis $|0\rangle, |1\rangle, |2\rangle, \cdots, |2^k-1\rangle$. Though, a better way to represent the basis is by binary strings of length k. This way, we keep the structure of k qubits and not a general vector space with 2^k dimension.

To take an example, the state space of 2 qubits is spanned by $|00\rangle, |01\rangle, |10\rangle, |11\rangle$. The state space of k qubits is spanned by $|B_n\rangle$, where B_n is the binary representation of number n and n ranges from 0 to 2^k-1. These basis states are product states, e.g., state $|001\rangle = |0\rangle \otimes |0\rangle \otimes |1\rangle$.

Notice the rise in the dimension of the state space for k qubits. If we described the state space of k bits, it will have dimension k (over the field of two elements). For k qubits, it rises to dimension 2^k over complex numbers. There is an exponential growth, which makes it hard to simulate quantum computer on a classical computer.

Exercise 37. Are the two states $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ and $|+\rangle \otimes |+\rangle$ same?

The tensor product structure of the composite system gives rise to a very interesting property called *entanglement*. As explained before, there are states in the composite system which cannot be decomposed into the states of their constituent systems. Such states are called *entangled states*.

The most famous example of an entangled state is called the *Bell state*,

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).$$

Exercise 38. Show that the Bell state can't be written as $|\psi\rangle \otimes |\phi\rangle$.

Exercise 39. Read about Pauli matrices X,Y and Z. Let $\psi = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$, calculate the value of $\langle\psi|X_1 \otimes Z_2|\psi\rangle$.

It is clear that every state in the composite system $H_1 \otimes H_2$ can be written as $\sum_{i=1}^{l} |\psi_i\rangle \otimes |\phi_i\rangle$ (Why?).

Exercise 40. Prove a bound of $dim(H_1) \times \dim(H_2)$ on l for any state in the composite system.

Can you give a better bound? Read about *Schmidt decomposition* for a better bound. We have defined when a state is entangled and when is it not. But how can we quantify entanglement? In other words, how entangled is an state? These are very interesting questions and lot of research is currently being done to answer them.

6 Assignment

Exercise 41. Prove that $rank(AB) \leq rank(A)$.

Exercise 42. Prove that $rank(A) = rank(A^*A)$ without using singular or spectral decomposition.

Hint: $rank(A) \geq rank(A^*A)$ is easy. For the other direction, reduce A to its reduced row echelon form.

Exercise 43. Show that $\langle v|A|w\rangle = \sum_{ij} A_{ij} v_i^* w_j$.

Exercise 44. Show that $tr(A|v\rangle\langle v|) = \langle v|A|v\rangle$.

Exercise 45. Suppose M, N are two square matrices, show that $MN = I \Rightarrow NM = I$. Notice that it is not true if matrix is not square, find a counterexample.

Exercise 46. Write the matrix representation of the operator which takes $\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)$ to $|0\rangle$ and $\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle)$ to $|1\rangle$.

Exercise 47. Both states $\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)$ and $\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle)$ give $|0\rangle$ and $|1\rangle$ with probability $1/2$. Using a simple operator show that these states are different. Hint: Can you think of two states which are definitely different?

Exercise 48. Prove that the matrix takes probability distributions to probability distributions iff all the entries are positive and column sums should be 1.

References

1. M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information.* Cambridge, 2010.
2. Gilbert Strang. *Introduction to Linear Algebra.* Wellesley-Cambridge Press, 2009.