Dirac Notation and Basic Linear Algebra for Quantum Computing

Dave Bacon
Department of Computer Science & Engineering, University of Washington

"Mathematicians tend to despise Dirac notation, because it can prevent them from making important distinctions, but physicists love it, because they are always forgetting such distinctions exist and the notation liberates them from having to remember." - David Mermin

In quantum theory, the basic mathematical structure we deal with is a complex Hilbert space. A complex Hilbert space is a complex vector space with an inner product and which is also complete with respect to the norm defined by the inner product (complete here means that every Cauchy sequence of vectors converges to a vector where convergence is measured by the norm.) In quantum computing we will be dealing almost exclusively with the case where this Hilbert space is the vector space of complex N dimensional vectors, $\mathbb{C}^N$ and the inner product between the vectors $v = [v_0\ v_1\ \cdots\ v_{N-1}]^T$ and $w = [w_0\ w_1\ \cdots\ w_{N-1}]^T$ (here T denotes transpose so we are writing the vectors as "column" vectors) is given by

$$\langle w, v\rangle = \sum_{i=0}^{N-1} w_i^* v_i \tag{1}$$

Here I'd like to introduce you to the Dirac bra-ket notation. Generally this notation can be used form any complex Hilbert space (and really in even more general settings, but we certainly won't need to worry about this) but since we will be dealing with the vector space $\mathbb{C}^N$ and the above inner product, it is useful to introduce this notation and show what it explicitly corresponds to in this complex Hilbert space.

Kets: Vectors in a complex Hilbert space are denoted in the bra-ket notation by kets. Let's call our Hilbert space $\mathcal{H}$. Then we denote a vector in this space as a ket $|v\rangle \in \mathcal{H}$. When the vector space we are dealing with is $\mathbb{C}^N$, then $|v\rangle$ is nothing more than an ordered n-tuple of complex numbers. In particular we will find it useful to think about $|v\rangle$ as a column vector of N complex numbers:

$$|v\rangle \leftrightarrow \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_{N-1} \end{bmatrix} \tag{2}$$

$v_i \in \mathbb{C}$. We can do everything with kets that we can do with vectors: we can add them $|v\rangle + |w\rangle$, multiply them by a scalar $\alpha|v\rangle$, etc. Now there is a special vector in a vector space, the zero vector. For the zero vector we will never write it as $|0\rangle$ (you'll see why soon.) Instead we will always just write it as 0, so $|v\rangle + 0 = |v\rangle$.

Bras: Recall that for a vector space V we can define it a dual vector space, V^*. This is the space of linear functionals on V: i.e. scalar-valued linear transformations on V. What does this mean? Well it means that an element of the dual space takes a vector and turns it into a complex number (functional on V). Further this transform is linear, meaning we can add these transforms and multiply them by a scalar. Elements of the dual vector space for a Hilbert space $\mathcal{H}$ are written as "bra"s: $\langle w| \in \mathcal{H}^*$. When we are dealing with $\mathbb{C}^N$ and the above inner product, then bras are nothing more than row vectors:

$$\langle w| \leftrightarrow [w_0\ w_1\ \cdots\ w_{N-1}] \tag{3}$$

Further for Hilbert spaces, for every ket $|v\rangle$, there is a unique bra $\langle v|$. In $\mathbb{C}^N$, the bra corresponding to a ket is obtained by taking the conjugate transpose (and vice versa):

$$|v\rangle = \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_{N-1} \end{bmatrix} \Leftrightarrow \langle v| = \begin{bmatrix} v_0^* & v_1^* & \cdots & v_{N-1}^* \end{bmatrix} \tag{4}$$

Inner product: Recall that a Hilbert space has an inner product. In bra-ket notation we denote the inner product between the vector $|v\rangle$ and the vector $|w\rangle$ by $\langle v, w\rangle = \langle v|w\rangle$. Notice that when we treat bras as row vectors and kets

as column vectors, then this inner product is just standard matrix multiplication:

$$\langle v|w\rangle = \begin{bmatrix} v_0^* & v_1^* & \cdots & v_{N-1}^* \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{N-1} \end{bmatrix} = \sum_{i=0}^{N-1} v_i^* w_i \tag{5}$$

Notice that $(\langle v|w\rangle)^* = \langle w|v\rangle$.

Computational Basis: Recall that a generating set for a vector space is a finite set of vectors such that every vector in this space can be written as a linear combination of these vectors, i.e. if $\{|e_i\rangle\}$ is a generating set for $\mathcal{H}$, then every element of $\mathcal{H}$ can be expressed as $\sum_i v_i|e_i\rangle$. A set of vectors $\{|e_i\rangle\}$ is linearly independent if $\sum_i v_i|e_i\rangle = 0$ only if $v_i = 0$ for all i. A basis is a generating set of vectors which are all linearly independent.

In our model of an information processing machine, our memory cells had configurations chosen from some fixed alphabet. We will always choose this alphabet to be $\{0, 1, \ldots, N-1\}$. The relevant Hilbert space for our memory cell is the vector space $\mathbb{C}^N$ with our standard inner product. Then a very important basis for this space is the computational basis corresponding to being in a configuration with probability unity. The computational basis is labelled $\{|0\rangle, |1\rangle, \ldots, |N-1\rangle\}$. When we translate these over to $\mathbb{C}^N$, these will just be column vectors with a single nonzero element equal to 1:

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots \quad |N-1\rangle = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \tag{6}$$

Since this is a basis, we can express every vector as a sum over these vectors:

$$|v\rangle = \sum_{i=0}^{N-1} v_i|i\rangle \tag{7}$$

Notice that we have put a variable inside of the ket which is summed over the appropriate range of integers. We will do this a lot, so it is best to get used to this: $|i\rangle$ is the ith computational basis state.

Finally we can, using the fact that every ket has a corresponding bra, also construction the computational basis for the bras $\{\langle 0|, \langle 1|, \ldots, \langle N-1|\}$ and expand any bra in terms of this basis

$$\langle w| = \sum_{i=0}^{N-1} w_i\langle i| \tag{8}$$

Further we note that the computational basis is a orthonormal basis. That is each basis element is orthogonal and each has a norm of unity. We express this as

$$\langle i|j\rangle = \delta_{i,j} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ otherwise} \end{cases} \tag{9}$$

Outer products: Another useful way to use bras and kets is to construct, instead of an inner product, outer products. This is an expression like $|v\rangle\langle w|$. When we port this over to $\mathbb{C}^N$ this simply becomes a N dimensional linear operator:

$$|v\rangle\langle w| = \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_{N-1} \end{bmatrix} \begin{bmatrix} w_0^* & w_1^* & \cdots & w_{N-1}^* \end{bmatrix} = \begin{bmatrix} v_0 w_0 * & v_0 w_1^* & \cdots & v_0 w_{N-1}^* \\ v_1 w_0 * & v_1 w_1^* & & \vdots \\ \vdots & & \ddots & \vdots \\ v_{N-1} w_0 * & \cdots & \cdots & v_{N-1} w_{N-1}^* \end{bmatrix} \tag{10}$$

Outer products which involve only computational basis states are especially nice to use because then the linear operator expressed in the computational basis has just a single non-zero element equal to unity. We can thus express any linear transform as a sum over outer product terms, i.e. as

$$M = \sum_{i,j=0}^{N-1} M_{ij}|i\rangle\langle j| \tag{11}$$

where M_{ij} is the matrix element in the ith row and jth column for the matrix corresponding to the linear transform M.

Conjugate Transpose: To convert between our column kets and row bras we used the conjugate transpose operation. This operation for the Hilbert space is the adjoint operation and is written by superscripting † to the expression of which we wish to take the adjoint. Thus $(|v\rangle)^\dagger = \langle v|$. Similarly, we can apply this to a linear operator, i.e.

$$\left(\sum_{i,j=0}^{N-1} M_{ij}|i\rangle\langle j|\right)^\dagger = \sum_{i,j=0}^{N-1} M_{ji}^*|i\rangle\langle j| \tag{12}$$

Ordering: As you have seen the order in which we express bras and kets changes how we are using these operators, i.e. $\langle v|w\rangle \neq |w\rangle\langle v|$. However it is always possible to move scalars (complex numbers) through our expressions. Thus for example $|v\rangle\langle v|w\rangle\langle w| = (\langle v|w\rangle)|v\rangle\langle w|$ since $\langle v|w\rangle$ is just a complex number. Further we can ask what happens to an expression in which we apply the conjugate transpose to the entire expression. In this case the entire order of the elements is reversed and the conjugate transpose is applied separately to each element. Thus, for example, $(\alpha|v\rangle\langle w|M)^\dagger = (M)^\dagger(\langle w|)^\dagger(|v\rangle)^\dagger(\alpha)^\dagger = M^\dagger|w\rangle\langle v|\alpha^* = \alpha^* M^\dagger|w\rangle\langle v|$ where in the second to last step we have used that conjugate transposing a scalar is just complex conjugation and in the last step we have pulled the complex number α^* through the entire expression.

We can now see why we often express linear transforms in the outer product form. To see why this is nice, consider the linear transform $M = \sum_{i,j=0}^{N-1} M_{ij}|i\rangle\langle j|$ acting on $|v\rangle = \sum_{k=0}^{N-1} v_k|k\rangle$. This is

$$M|v\rangle = \sum_{i,j=0}^{N-1} M_{ij}|i\rangle\langle j| \sum_{k=0}^{N-1} v_k|k\rangle = \sum_{i,j,k=0}^{N-1} M_{ij}v_k|i\rangle\langle j|k\rangle = \sum_{i,j,k=0}^{N-1} M_{ij}v_k|i\rangle\delta_{j,k} = \sum_{i=0}^{N-1}\left(\sum_{j=0}^{N-1} M_{ij}v_j\right)|i\rangle \tag{13}$$

The bra-ket notation is handy because it allows us to perform manipulations even more complicated than this is a fairly simple manner. Of course, after a while you get to be old hat at this and it is just a second language. And then you learn that there are even more interesting notations which you can learn (see for example the diagramatic methods of Roger Penrose and others.)

Completeness: For completeness, it is often useful to express the identiy transform in outer product form. This is just the simple formula

$$I = \sum_{i=0}^{N-1} |i\rangle\langle i| \tag{14}$$

Types of Linear Transforms: We won't review all of linear algebra here, but it is useful to recall the different types of linear transforms. In particular we will focus on linear transforms from a vector space to itself (i.e. the corresponding matrix representation is a square matrix.) First recall that a transform M is "hermitian" if $M^\dagger = M$. A transform is "normal" if $M^\dagger M = MM^\dagger$. Notice that hermitian transforms are always normal. A transform whose inverse is its adjoint, $U^\dagger U = I$, is a "unitary" transform. Unitary transforms are normal. A transform whose square is equal to itself, $M^2 = M$ is a "projector".

Eigensystems: Recall that if $M|v\rangle = \lambda_v|v\rangle$, then $|v\rangle$, nonzero, is an eigenvector of M with eigenvalue λ_v (which is just a complex number.) Sometimes it is possible to find $dim(M)$ eigenvectors which form a basis for $\mathcal{H}$. In this case, we say that M is diagonalizable. If $\{|\phi_i\rangle\}$ is such a basis with $|\phi_i\rangle$ being an eigenvector with eigenvalue λ_i, then we can express M as $M = \sum_{i=0}^{N-1} \lambda_i|\phi_i\rangle\langle\phi_i|$. A fundamental theorem of linear algebra says that an operator is diagonalizable if and only if it is normal. This is a good time to remind you that not all linear transforms are diagonalizable! Further recall that hermitian transforms always have real eigenvalues. In fact, when we say the word hermitian, the next words out of our mouth are almost always "diagonalizable" and sometimes "real eigenvalues."

What about the eigenvalues of unitary matrices? Suppose $|v\rangle$ is an eigenvector of a unitary matrix U with eigenvalue λ_v. Then since $U^\dagger U = I$ we find that $\langle v|U^\dagger U|v\rangle = \langle v|v\rangle$ so $\langle v|\lambda_v^*\lambda_v|v\rangle = \langle v|v\rangle$. Thus since $|v\rangle$ is not the zero vector, $\lambda_v^*\lambda_v = 1$ or $|\lambda_v|^2 = 1$. This implies that $|\lambda_v|$ is a complex root of unity: $\lambda_v = \exp(2\pi i\nu_v)$ for some real $0 \leq \nu_v < 2\pi$.

Positive operators: A transform is positive if $\langle v|M|v\rangle$ is real and is greater than or equal to zero for all vectors $|v\rangle$. If $\langle v|M|v\rangle$ is always greater than zero for $|v\rangle \neq 0$, then we say that the matrix is positive definite. Positive operators are necessarily Hermitian.

Polar decomposition: It is easy to show that $M^\dagger M$ is always positive. Thus we can diagonalize $M^\dagger M$: $M^\dagger M = \sum_i \lambda_i|\phi_i\rangle\langle\phi_i|$. Since $\lambda_i \geq 0$ we can then define the square root of $M^\dagger M$ by taking the square roots of these eigenvalues: $\sqrt{M^\dagger M} = \sum_i \sqrt{\lambda_i}|\phi_i\rangle = langle\phi_i|$. The polar decomposition theorem states that every matrix M can be expressed as the product UP where U is a unitary matrix and P is positive and in fact is equal $\sqrt{M^\dagger M}$.

Singular valued decomposition: Using the polar decomposition theorem and the fact that positive matrices are diagonalizable, the singular valued decomposition states that every M can be decomposed as UDV where U and V are unitary and D is a diagonal matrix. Thus while it is not always possible to diagonalize every matrix, it is always possible to produce a singular valued decomposition of the above form.

Tensor products: Tensor products usually give people fits when they are first learning about quantum computation. Suppose we have two Hilbert space $\mathcal{H}_1$ and $\mathcal{H}_2$. Then we can form a new Hilbert space, which is the tensor product of these two Hilbert spaces, $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. Suppose that a basis for $\mathcal{H}_1$ is $\{|i\rangle\}$ and a basis for $\mathcal{H}_2$ is $\{|j\rangle\}$. Then a basis for $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ is $\{|i\rangle \otimes |j\rangle$. In particular a vector in $\mathcal{H}$ can be expanded as $|v\rangle = \sum_{i,j} v_{i,j}|i\rangle \otimes |j\rangle$. At this point it is useful to note that some vectors in $\mathcal{H}$ can be expressed as $|v\rangle \otimes |w\rangle$, but that not all vectors can be expressed in this manner. For example if $\mathcal{H}_1 = \mathcal{H}_2$ with the relevant vector space spanned by $|0\rangle$ and $|1\rangle$, then the vector $|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle$ cannot be expressed as $|v\rangle \otimes |w\rangle$.

Since $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ is a new Hilbert space, we can consider all of the things we normally consider on this new Hilbert space. For example we can talk about linear transform on this Hilbert space. Just as sometimes we can think of vectors in $\mathcal{H}$ as one vector in $\mathcal{H}_1$ tensored with one vector in $\mathcal{H}_2$, some of the linear transforms on $\mathcal{H}$ can be expressed as $A \otimes B$. Note that $A \otimes B|v\rangle \otimes |w\rangle = (A|v\rangle) \otimes (B|w\rangle)$.

Commutator and Anticommutator: The commutator of two matrices A and B is denoted $[A, B]$ and is equal to $AB - BA$. The anticommutator of two matrices A and B is denoted $\{A, B\}$ and is equal to $AB + BA$.

Trace: The trace of a matrix M is defined to be the sum of its diagonal elements $Tr(M) = \sum_i M_{ii}$. The trace satisfies $Tr(A + B) = Tr(A) + Tr(B)$ and $Tr(AB) = Tr(BA)$. It is also useful to note that $Tr(A|v\rangle\langle v|) = \langle v|A|v\rangle$.