Chapter III

Quantum Computation

A Mathematical preliminaries

"[I]nformation is physical, and surprising physical theories such as quantum mechanics may predict surprising information processing abilities." (Nielsen & Chuang, 2010, p. 98)

A.1 Complex numbers

If you go to the course webpage, and look under Quantum Computation in the Topics section, you will see a link to "complex number review [FFC-ch4]." Depending on how familiar you are with complex numbers, read or skim it through section 4.4.2.1 (pp. 41–53). This should tell you all you need to know (and a little more).

A.2 Linear algebra review

A.2.a Dirac bracket notation

Much of the math of quantum computation is just elementary linear algebra, but the notation is different (and of course there is a physical interpretation). The Dirac bracket notation will seem peculiar if you are not used to it, but

it is elegant and powerful, as are all good notations. Think of it like a new programming language.

First, the notation $|\psi\rangle$ represents an n-dimensional vector, which we can write in a particular basis as a $n \times 1$ complex column vector, $|\psi\rangle = (v_1, \ldots, v_n)^{\mathrm{T}}$. We pronounce $|\psi\rangle$ "ket psi" or "psi ket." Normally the vectors are finite-dimensional, but they can be infinite-dimensional if the vectors have a finite magnitude (their components are square-summable), $\sum_k |v_k|^2 < \infty$.

The Dirac notation has the advantage that we can use arbitrary names for vectors, for example:

$$|\text{excited}\rangle, |\text{zero}\rangle, |\text{one}\rangle, |\uparrow\rangle, |\nearrow\rangle, |1\rangle, |101\rangle, |5\rangle, |f(\mathbf{x})\rangle, |1 \otimes g(1)\rangle.$$

It may be easier to remember if you notice that it looks kind of like an arrow; compare $|v\rangle$ and $\vec{v}$.

The notation $\langle\phi|$ represents a $1 \times n$ complex *row* vector, $\langle\phi| = (u_1, \ldots, u_n)$ in a particular basis. We pronounce $\langle\psi|$ "bra psi" or "psi bra." If $|\psi\rangle = (v_1, \ldots, v_n)^{\mathrm{T}}$, then $\langle\psi| = (\overline{v_1}, \ldots, \overline{v_n})$, where $\overline{v_k}$ is the complex conjugate of v_k. Recall that

$$\overline{x + iy} = x - iy \text{ and } \overline{re^{i\phi}} = re^{-i\phi}.$$

We define the *adjoint* (*conjugate transpose, Hermitian transpose*) $M^\dagger$ of a matrix M by

$$(M^\dagger)_{jk} = \overline{M_{kj}}.$$

We pronounce it "M dagger," "M adjoint," etc. Note that corresponding bras and kets are adjoints: $\langle\psi| = |\psi\rangle^\dagger$ and $|\psi\rangle = \langle\psi|^\dagger$.

A.2.b Inner product

Suppose $|\phi\rangle = (u_1, \ldots, u_n)^{\mathrm{T}}$ and $|\psi\rangle = (v_1, \ldots, v_n)^{\mathrm{T}}$ in the same basis. Then the *complex inner product* of the vectors is defined $\sum_k \overline{u_k} v_k = \langle\phi|\,|\psi\rangle$. Thus, the inner product of two vectors is the conjugate transpose of the first times the second. This is the convention in physics, which we will follow; mathematicians usually put the complex conjugate on the second argument. It doesn't make any difference so long as you are consistent. Since the inner product multiplies a $1 \times n$ matrix by an $n \times 1$ matrix, the result is a 1×1 matrix, or scalar. This product of a bra and a ket is usually abbreviated $\langle\phi \mid \psi\rangle = \langle\phi|\,|\psi\rangle$, which can be pronounced "ϕ-bra ket-ψ" or "ϕ bra-ket ψ."

The complex inner product satisfies several important properties:

1. It is *positive definite*:

$$\begin{aligned} \langle \psi \mid \psi \rangle &> 0, \quad \text{if } |\psi\rangle \neq \mathbf{0}, \\ \langle \psi \mid \psi \rangle &= 0, \quad \text{if } |\psi\rangle = \mathbf{0}. \end{aligned}$$

2. It has *conjugate symmetry*: $\langle \phi \mid \psi \rangle = \overline{\langle \psi \mid \phi \rangle}$.

3. It is *linear in its second argument* (the one that's not conjugated):

$$\begin{aligned} \langle \phi \mid c\psi \rangle &= c\langle \phi \mid \psi \rangle, \quad \text{for } c \in \mathbb{C}, \\ \langle \phi \mid \psi + \chi \rangle &= \langle \phi \mid \psi \rangle + \langle \phi \mid \chi \rangle. \end{aligned}$$

Note that conjugate symmetry and linearity in the second argument together imply that $\langle c\phi \mid \psi \rangle = \overline{c}\langle \phi \mid \psi \rangle$ (antilinearity in the first argument). The complex inner product is called *sesquilinear*, which means "one-and-a-half linear" (in contrast to the inner product of real vectors, which is linear in both arguments, i.e., *bilinear*).

The *norm* or *magnitude* of a vector is defined $\||\psi\rangle\| = \sqrt{\langle \psi \mid \psi \rangle}$. That is, $\||\psi\rangle\|^2 = |v_1|^2 + \cdots |v_n|^2$. A vector is *normalized* if $\||\psi\rangle\| = 1$. Note that normalized vectors fall on the surface of an n-dimensional hypersphere.

A.2.c Bases and Generalized Fourier Series

Vectors $|\phi\rangle$ and $|\psi\rangle$ are *orthogonal* if $\langle \phi \mid \psi \rangle = 0$. A set of vectors is *orthogonal* if each vector is orthogonal to all the others. An *orthonormal (ON)* set of vectors is an orthogonal set of normalized vectors. A set of vectors $|\phi_1\rangle, |\phi_2\rangle, \ldots$ *spans* a vector space if for every vector $|\psi\rangle$ in the space there are complex coefficients $c_1, c_2, \ldots$ such that $|\psi\rangle = \sum_k c_k |\phi_k\rangle$. A *basis* for a vector space is a linearly independent set of vectors that spans the space. Equivalently, a basis is a *minimal generating set* for the space; that is, all of the vectors in the space can be generated by linear combinations of the basis vectors. An *(orthonormal) basis* for a vector space is an (orthonormal) set of vectors that spans the space. In general, when I write "basis" I mean "orthonormal basis." Any vector in the space has a unique representation as a linear combination of the basis vectors.

A *Hilbert space* is a complete inner-product space. "Complete" means that all Cauchy sequences of vectors (or functions) have a limit in the space. (In a Cauchy sequence, $\|x_m - x_n\| \to 0$ as $m, n \to \infty$.) Hilbert spaces may

be finite- or infinite-dimensional. Suppose $|1\rangle$, $|2\rangle$, ... is an ON basis for a Hilbert space $\mathcal{H}$. Note, that these are just the names of the basis vectors (we could have used $|e_1\rangle$, $|e_2\rangle$, ... or something similar), and have nothing to do with the integers 1, 2, etc. Given such a basis, any $|\psi\rangle$ in $\mathcal{H}$ can be expanded in a *generalized Fourier series*:

$$|\psi\rangle = \sum_k c_k |k\rangle.$$

The *generalized Fourier coefficients* c_k can be determined by taking inner products with the corresponding basis vectors:

$$\langle k \mid \psi\rangle = \langle k| \sum_j c_j |j\rangle = \sum_j c_j \langle k \mid j\rangle = c_k.$$

Therefore, $c_k = \langle k \mid \psi\rangle$. Hence,

$$|\psi\rangle = \sum_k c_k |k\rangle = \sum_k \langle k \mid \psi\rangle \, |k\rangle = \sum_k |k\rangle\langle k \mid \psi\rangle.$$

This is just the vector's representation in a particular basis. (Note that this equation implies the identity matrix $I = \sum_k |k\rangle\langle k|$.)

A *linear operator* $L : \mathcal{H} \to \hat{\mathcal{H}}$ satisfies $L(c|\phi\rangle + d|\psi\rangle) = cL(|\phi\rangle) + dL(|\psi\rangle)$ for all $|\phi\rangle, |\psi\rangle \in \mathcal{H}$ and $c, d \in \mathbb{C}$. For example, differentiation is a linear operator.

A linear operator $L : \mathcal{H} \to \hat{\mathcal{H}}$ can be represented by a (possibly infinite-dimensional) matrix relative to bases for $\mathcal{H}$ and $\hat{\mathcal{H}}$. To see this, suppose $|1\rangle$, $|2\rangle$, ... is a basis for $\mathcal{H}$ and $|\hat{1}\rangle$, $|\hat{2}\rangle$, ... is a basis for $\hat{\mathcal{H}}$. Consider $|\phi\rangle = L|\psi\rangle$ and represent the vectors in these bases by their Fourier coefficients: $b_j = \langle \hat{j} \mid \phi\rangle$ and $c_k = \langle k \mid \psi\rangle$. Hence $|\phi\rangle$ is represented by the vector $\mathbf{b} = (b_1, b_2, \ldots)^{\mathrm{T}}$ and $|\psi\rangle$ by the vector $\mathbf{c} = (c_1, c_2, \ldots)^{\mathrm{T}}$. Apply the linearity of L:

$$\begin{aligned} b_j &= \langle \hat{j} \mid \phi\rangle \\ &= \langle \hat{j} \mid L \mid \psi\rangle \\ &= \langle \hat{j}| L \left(\sum_k c_k |k\rangle \right) \\ &= \langle \hat{j}| \left(\sum_k c_k L|k\rangle \right) \\ &= \sum_k \langle \hat{j} \mid L \mid k\rangle c_k. \end{aligned}$$

Therefore, define the matrix $M_{jk} \stackrel{\text{def}}{=} \langle \hat{\jmath} \mid L \mid k \rangle$ and we see $\mathbf{b} = \mathbf{Mc}$. For this reason, an expression of the form $\langle \hat{\jmath} \mid L \mid k \rangle$ is sometimes called a *matrix element* of the operator L. Note that the matrix depends on the bases we choose.

A.2.d Outer product or dyad

We can form the product of a ket and a bra, which is called a *dyad* or *outer product*. Consider first the finite dimensional case. If $|\phi\rangle$ is an $m \times 1$ column vector, and $|\psi\rangle$ is an $n \times 1$ column vector (so that $\langle\psi|$ is a $1 \times n$ row vector), then the outer product $|\phi\rangle\langle\psi|$ is an $m \times n$ matrix. In most cases of interest $m = n$. Since matrix multiplication is associative, $(|\phi\rangle\langle\psi|)\,|\chi\rangle = |\phi\rangle\,\langle\psi \mid \chi\rangle$. More generally, we can form outer products of infinite-dimensional vectors in Hilbert spaces. If $|\phi\rangle \in \mathcal{H}'$ and $|\psi\rangle \in \mathcal{H}$, then $|\phi\rangle\langle\psi|$ is the linear operator $L : \mathcal{H} \to \mathcal{H}'$ defined, for any $|\chi\rangle \in \mathcal{H}$:

$$L|\chi\rangle = (|\phi\rangle\langle\psi|)|\chi\rangle = |\phi\rangle\,\langle\psi \mid \chi\rangle.$$

That is, $|\phi\rangle\langle\psi|$ is the operator that returns $|\phi\rangle$ scaled by the inner product of $|\psi\rangle$ and its argument. To the extent that the inner product $\langle\psi \mid \chi\rangle$ measures the similarity of $|\psi\rangle$ and $|\chi\rangle$, the result $|\phi\rangle$ is weighted by this similarity. The product of a ket and a bra $|\phi\rangle\langle\psi|$ can be pronounced "ϕ-ket bra-ψ" or "ϕ ket-bra ψ," and abbreviated $|\phi\rangle\!\langle\psi|$. This product is also called a *dyad*.

The special case of $|\phi\rangle\!\langle\phi|$ in which $|\phi\rangle$ is normalized is called a *projector* onto $|\phi\rangle$. This is because $|\phi\rangle\!\langle\phi|\,|\psi\rangle = |\phi\rangle\,\langle\phi \mid \psi\rangle$, that is, $|\phi\rangle$ scaled by the projection of $|\psi\rangle$ on $|\phi\rangle$. More generally, if $|\eta_1\rangle, \ldots, |\eta_m\rangle$ are orthonormal, then $\sum_{k=1}^{m} |\eta_k\rangle\!\langle\eta_k|$ projects into the m-dimensional subspace spanned by these vectors.

Any linear operator can be represented as a weighted sum of outer products. To see this, suppose $L : \mathcal{H} \to \hat{\mathcal{H}}$, $|\hat{\jmath}\rangle$ is a basis for $\hat{\mathcal{H}}$, and $|k\rangle$ is a basis for $\mathcal{H}$. Consider $|\phi\rangle = L|\psi\rangle$. We know from Sec. A.2.c that

$$\langle \hat{\jmath} \mid \phi \rangle = \sum_k M_{jk} c_k, \text{ where } M_{jk} = \langle \hat{\jmath} \mid L \mid k \rangle, \text{ and } c_k = \langle k \mid \psi \rangle.$$

Hence,

$$|\phi\rangle \;=\; \sum_j |\hat{\jmath}\rangle\,\langle \hat{\jmath} \mid \phi \rangle$$

$$\begin{aligned} &= \sum_j |\hat{j}\rangle \left(\sum_k M_{jk} \langle k \mid \psi \rangle \right) \\ &= \left(\sum_j |\hat{j}\rangle \sum_k M_{jk} \langle k| \right) |\psi\rangle \\ &= \left(\sum_{jk} M_{jk} |\hat{j}\rangle\langle k| \right) |\psi\rangle. \end{aligned}$$

Hence, we have a sum-of-outer-products representation of the operator in terms of its matrix elements:

$$L = \sum_{jk} M_{jk} |\hat{j}\rangle\langle k|, \text{ where } M_{jk} = \langle \hat{j} \mid L \mid k \rangle.$$

A.2.e Tensor products

In this section we define the *tensor product*, which is of central importance in quantum computation. To see where we are headed, suppose $|\phi\rangle$ and $|\psi\rangle$ are the quantum states of two objects (e.g., the states of two *qubits*, or quantum bits). Then the state of the composite system (e.g., the pair of qubits) will be represented by the tensor product $|\phi\rangle \otimes |\psi\rangle$, which we can think of as a sort of concatenation or structure formed of $|\phi\rangle$ and $|\psi\rangle$. If $\mathcal{H}$ and $\mathcal{H}'$ are the Hilbert spaces from which these states are drawn, then the *tensor product space* $\mathcal{H} \otimes \mathcal{H}'$ is the space of all possible states of the two-qubit system (i.e., all such pairs or structures). We can apply the tensor product to operators as well: $L \otimes M$ is the operator that applies, in parallel, L to the first qubit of the pair and M to the second qubit:

$$(L \otimes M)(|\phi\rangle \otimes |\psi\rangle) = (L|\phi\rangle) \otimes (M|\psi\rangle).$$

For vectors, operators, and spaces, we pronounce $L \otimes M$ as "L tensor M." As we will see, the tensor product is essential to much of the power of quantum computation. Next we develop these ideas more formally.

Suppose that $|\eta_j\rangle$ is an ON basis for $\mathcal{H}$ and $|\eta'_k\rangle$ is an ON basis for $\mathcal{H}'$. For every pair of basis vectors, define the *tensor product* $|\eta_j\rangle \otimes |\eta'_k\rangle$ as a sort of couple or pair of the two basis vectors; that is, there is a one-to-one correspondence between the $|\eta_j\rangle \otimes |\eta'_k\rangle$ and the pairs in $\{|\eta_0\rangle, |\eta_1\rangle, \ldots\} \times \{|\eta'_0\rangle, |\eta'_1\rangle, \ldots\}$. Define the *tensor product space* $\mathcal{H} \otimes \mathcal{H}'$ as the space spanned

by all linear combinations of the basis vectors $|\eta_j\rangle \otimes |\eta'_k\rangle$. Therefore each element of $\mathcal{H} \otimes \mathcal{H}'$ is represented by a unique sum $\sum_{jk} c_{jk} |\eta_j\rangle \otimes |\eta'_k\rangle$. In order to make $\mathcal{H} \otimes \mathcal{H}'$ a Hilbert space, we need an inner product:

$$\langle \phi_1 \otimes \phi_2 \mid \psi_1 \otimes \psi_2 \rangle = \langle \phi_1 \mid \psi_1 \rangle \, \langle \phi_2 \mid \psi_2 \rangle.$$

That is, we multiply the inner products of the corresponding elements of the tensor product pairs.

Usually, we are dealing with finite-dimensional spaces, in which case the tensor products can be defined less abstractly in terms of matrices. Suppose in a given basis $|\phi\rangle = (u_1, \ldots, u_m)^{\mathrm{T}}$ and $|\psi\rangle = (v_1, \ldots, v_n)^{\mathrm{T}}$, then their tensor product in that basis can be defined by the *Kronecker product*:

$$\begin{aligned} |\phi\rangle \otimes |\psi\rangle &= \begin{pmatrix} u_1|\psi\rangle \\ \vdots \\ u_m|\psi\rangle \end{pmatrix} \\ &= \left(u_1|\psi\rangle^{\mathrm{T}}, \ldots, u_m|\psi\rangle^{\mathrm{T}}\right)^{\mathrm{T}} \\ &= (u_1 v_1, \ldots, u_1 v_n, \ldots, u_m v_1 \ldots, u_m v_n)^{\mathrm{T}}. \end{aligned}$$

Note that this is an $mn \times 1$ column vector and that

$$(|\phi\rangle \otimes |\psi\rangle)_{(j-1)n+k} = u_j v_k.$$

This combinatorial explosion of dimension is what gives quantum computation its power.

The following abbreviations are frequent: $|\phi\psi\rangle = |\phi, \psi\rangle = |\phi\rangle|\psi\rangle = |\phi\rangle \otimes |\psi\rangle$. Note that $|\phi\rangle|\psi\rangle$ can only be a tensor product because it would not be a legal matrix product. These are some useful properties of the tensor product (which is bilinear):

$$\begin{aligned} (c|\phi\rangle) \otimes |\psi\rangle &= c(|\phi\rangle \otimes |\psi\rangle) = |\phi\rangle \otimes (c|\psi\rangle), \\ (|\phi\rangle + |\psi\rangle) \otimes |\chi\rangle &= (|\phi\rangle|\chi\rangle) + (|\psi\rangle|\chi\rangle), \\ |\phi\rangle \otimes (|\psi\rangle + |\chi\rangle) &= (|\phi\rangle \otimes |\psi\rangle) + (|\phi\rangle \otimes |\chi\rangle). \end{aligned}$$

The tensor product of linear operators is defined

$$(L \otimes M) \, (|\phi\rangle \otimes |\psi\rangle) = L|\phi\rangle \otimes M|\psi\rangle.$$

Using the fact that $|\psi\rangle = \sum_{jk} c_{jk}|\eta_j\rangle \otimes |\eta'_k\rangle$ you can compute $(L \otimes M)|\psi\rangle$ for an arbitrary $|\psi\rangle \in \mathcal{H} \otimes \mathcal{H}'$ (exercise). If $\mathbf{M}$ is a $k \times m$ matrix and $\mathbf{N}$ is a $l \times n$ matrix, then their Kronecker product is a $kl \times mn$ matrix:

$$\mathbf{M} \otimes \mathbf{N} = \begin{pmatrix} M_{11}\mathbf{N} & M_{12}\mathbf{N} & \cdots & M_{1m}\mathbf{N} \\ M_{21}\mathbf{N} & M_{22}\mathbf{N} & \cdots & M_{2m}\mathbf{N} \\ \vdots & \vdots & \ddots & \vdots \\ M_{k1}\mathbf{N} & M_{k2}\mathbf{N} & \cdots & M_{km}\mathbf{N} \end{pmatrix}.$$

Matrices, of course, are linear operators and satisfy $(\mathbf{M} \otimes \mathbf{N})\ (|\phi\rangle \otimes |\psi\rangle) = \mathbf{M}|\phi\rangle \otimes \mathbf{N}|\psi\rangle$.

For a vector, operator, or space M, we define the *tensor power* $M^{\otimes n}$ to be M tensored with itself n times:

$$M^{\otimes n} = \overbrace{M \otimes M \otimes \cdots \otimes M}^{n}.$$

A.2.f Properies of operators and matrices

Several properties of operators and matrices are important in quantum computation. An operator $L : \mathcal{H} \to \mathcal{H}$ is *normal* if $L^\dagger L = LL^\dagger$. The same applies to square matrices. That is, normal operators commute with their adjoints.

An operator $L : \mathcal{H} \to \mathcal{H}$ is *Hermitian* or *self-adjoint* if $L^\dagger = L$. The same applies to square matrices. (Hermitian matrices are the complex analogues of symmetric matrices.) Note that Hermitian operators are normal. It is easy to see that L is Hermitian if and only if $\langle\phi \mid L \mid \psi\rangle = \langle\psi \mid L \mid \phi\rangle$ for all $|\phi\rangle, |\psi\rangle$ (since $\langle\psi \mid L \mid \phi\rangle = \langle\phi \mid L^\dagger \mid \psi\rangle = \langle\phi \mid L \mid \psi\rangle$). A normal matrix is Hermitian if and only if it has real eigenvalues (exercise). This is important in quantum mechanics, since measurement results are usually assumed to be real.

An operator U is *unitary* if $U^\dagger U = UU^\dagger = I$. That is, a unitary operator is invertible and its inverse is its adjoint: if U is unitary, then $U^{-1} = U^\dagger$. Obviously every unitary operator is also normal. (A normal matrix is unitary if and only if its *spectrum* — the multiset of its eigenvalues — is contained in the unit circle in the complex plane.)

Unitary operators are like rotations of a complex vector space (analogous to orthogonal operators, which are rotations of a real vector space). Just as orthogonal transformations preserve the angles between vectors, unitary operators preserve their inner products (which are analogous to angles).

Consider the inner product between $U|\phi\rangle$ and $U|\psi\rangle$:

$$(U|\phi\rangle)^\dagger U|\psi\rangle = \langle\phi \mid U^\dagger U \mid \psi\rangle = \langle\phi \mid \psi\rangle.$$

Hence, the inner product of $U|\phi\rangle$ and $U|\psi\rangle$ is the same as the inner product of $|\phi\rangle$ and $|\psi\rangle$. Therefore, unitary operators are *isometric*, that is, they preserve norms:

$$\|U|\psi\rangle\|^2 = \langle\psi \mid U^\dagger U \mid \psi\rangle = \langle\psi \mid \psi\rangle = \||\psi\rangle\|^2.$$

Unitary operators are important in quantum computation because the evolution of quantum systems is unitary.

A.2.g Spectral decomposition and operator functions (supplementary)

For any normal operator on a finite-dimensional Hilbert space, there is an ON basis that diagonalizes the operator, and conversely, any diagonalizable operator is normal. This is called a *spectral decomposition* of the operator. The ON basis comprises its set of eigenvectors, which we can write $|0\rangle, |1\rangle, \ldots, |n\rangle$ and the corresponding eigenvalues λ_k are the diagonal elements (cf. Sec. A.2.d, p. 72): $L = \sum_{k=1}^{n} \lambda_k |k\rangle\langle k|$. Therefore, a matrix is normal if and only if it can be diagonalized by a unitary transform (see A.2.f, above). That is, it is normal if and only if there is a unitary U such that $L = U\Lambda U^\dagger$, where $\Lambda = \mathrm{diag}(\lambda_1, \ldots \lambda_n)$. If $|0\rangle, |1\rangle, \ldots, |n\rangle$ is the basis, then $U = (|0\rangle, |1\rangle, \ldots, |n\rangle)$ and

$$U^\dagger = \begin{pmatrix} \langle 0| \\ \langle 1| \\ \vdots \\ \langle n| \end{pmatrix}.$$

(More generally, this property holds for compact normal operators.)

It is often convenient to extend various complex functions (e.g., $\ln, \exp, \sqrt{}$) to normal matrices and operators. If $f : \mathbb{C} \to \mathbb{C}$ and $L : \mathcal{H} \to \mathcal{H}$, then we define:

$$f(L) \stackrel{\text{def}}{=} \sum_{k=1}^{n} f(\lambda_k) |k\rangle\langle k|,$$

where $L = \sum_{k=1}^{n} \lambda_k |k\rangle\langle k|$ is the spectral decomposition of L. Therefore, for a normal linear operator or matrix L we can write $\sqrt{L}$, $\ln L$, e^L, etc. This is not an arbitrary extension, but follows from the power series expansion for f.