Appendix A

Vector spaces: basic properties and Dirac notation

In quantum computation the integers from 0 to N are associated with $N+1$ orthogonal unit vectors in a vector space of $D = N+1$ dimensions over the complex numbers. The nature of this association is the subject of Chapter 1. Here we review some of the basic properties of such a vector space, while relating conventional vector-space notation to the Dirac notation used in quantum computer science. Usually the dimension D is a power of 2, but this does not matter for our summary of the basic facts and nomenclature.

In conventional notation such a set of $D = N+1$ orthonormal vectors might be denoted by symbols such as $\phi_0, \phi_1, \phi_2, \ldots, \phi_N$. The orthogonality and normalization conditions are expressed in terms of the inner products (ϕ_x, ϕ_y):

$$(\phi_x, \phi_y) = \begin{cases} 0, & x \neq y; \\ 1, & x = y. \end{cases} \tag{A.1}$$

In quantum computation the indices x and y describing the integers associated with the vectors play too important a role to be relegated to tiny fonts in subscripts. Fortunately quantum mechanics employs a notation for vectors, invented by the physicist Paul Dirac, which is well suited for representing such information more prominently. One replaces the symbols ϕ_x and ϕ_y by $|x\rangle$ and $|y\rangle$, and represents the inner product (ϕ_x, ϕ_y) by the symbol $\langle x|y\rangle$. The orthonormality condition (A.1) becomes

$$\langle x|y\rangle = \begin{cases} 0, & x \neq y; \\ 1, & x = y. \end{cases} \tag{A.2}$$

Vectorial character is conveyed by the symbol $|\ \rangle$, with the specific vector being identified by whatever it is that goes between the bent line $\rangle$ and the vertical line $|$. This notational strategy is reminiscent of the notation for vectors in ordinary three-dimensional physical space (which we will use here for such vectors) in which vectorial character is indicated by a horizontal arrow above a symbol denoting the specific vector being referred to: $\overrightarrow{r}$.

Symbols like ϕ and ψ remain useful in the notation of quantum computation for representing generic vectors, but for consistency with the notation for vectors associated with specific integers, and to emphasize their vectorial character, they too are enclosed between a bent line $\rangle$

and a vertical line |, becoming $|\phi\rangle$ and $|\psi\rangle$. Some mathematicians disapprove of this practice. Why write $|\psi\rangle$, introducing the spurious symbols $\rangle$ and |, when ψ by itself does the job perfectly well? This gets it backwards. The real point is that the important information – for example the number 7798 – is easier to read in the form $|7798\rangle$ than when presented in small print in the form ϕ_{7798}. Why introduce in a normal font the often uninformative symbol ϕ, at the price of demoting the most important information to a mere subscript?

The vector space that describes the operation of a quantum computer consists of all linear combinations $|\psi\rangle$ of the $N+1$ orthonormal vectors $|x\rangle$, $x = 0, \ldots, N$, with coefficients α_x taken from the complex numbers:

$$|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle + \cdots + \alpha_N|N\rangle = \sum_{x=0}^{N} \alpha_x|x\rangle, \tag{A.3}$$

where $\alpha_x = u_x + iv_x$, u_x and v_x are real numbers, and $i = \sqrt{-1}$.

The mathematicians' preference for writing ψ instead of $|\psi\rangle$ for generic vectors is explicitly acknowledged in the useful convention that $|\alpha\psi + \beta\phi\rangle$ is nothing more than an alternative way of writing the vector $\alpha|\psi\rangle + \beta|\phi\rangle$:

$$|\alpha\psi + \beta\phi\rangle = \alpha|\psi\rangle + \beta|\phi\rangle. \tag{A.4}$$

In a vector space over the complex numbers the inner product of two general vectors is a complex number satisfying

$$\langle\psi|\phi\rangle = \langle\phi|\psi\rangle^*, \tag{A.5}$$

where * denotes complex conjugation:

$$(u + iv)^* = u - iv, \quad u, v \text{ real}. \tag{A.6}$$

The inner product is linear in the right-hand vector,

$$\langle\phi|\alpha\psi_1 + \beta\psi_2\rangle = \alpha\langle\phi|\psi_1\rangle + \beta\langle\phi|\psi_2\rangle, \tag{A.7}$$

and therefore, from (A.5), "anti-linear" in the left-hand vector,

$$\langle\alpha\phi_1 + \beta\phi_2|\psi\rangle = \alpha^*\langle\phi_1|\psi\rangle + \beta^*\langle\phi_2|\psi\rangle. \tag{A.8}$$

The inner product of a vector with itself is a real number satisfying

$$\langle\phi|\phi\rangle > 0, \quad |\phi\rangle \neq 0. \tag{A.9}$$

It follows from the orthonormality condition (A.2) that the inner product of the vector $|\psi\rangle$ in (A.3) with another vector

$$|\phi\rangle = \beta_0|0\rangle + \beta_1|1\rangle + \cdots + \beta_N|N\rangle = \sum_x \beta_x|x\rangle \tag{A.10}$$

is given in terms of the expansion coefficients α_x and β_x (called *amplitudes* in quantum computation) by

$$\langle\phi|\psi\rangle = \sum_x \beta_x^* \alpha_x. \tag{A.11}$$

The squared magnitude of a vector is its inner product with itself, so (A.11) gives for the squared magnitude

$$\langle\psi|\psi\rangle = \sum_x |\alpha_x|^2, \tag{A.12}$$

where

$$|u + iv|^2 = u^2 + v^2, \quad u, v \text{ real}. \tag{A.13}$$

The form (A.12) gives an explicit confirmation of the rule (A.9).

A linear transformation **A** associates with every vector $|\psi\rangle$ another vector, called $\mathbf{A}|\psi\rangle$, subject to the rule (linearity)

$$\mathbf{A}\big(\alpha|\psi\rangle + \beta|\phi\rangle\big) = \alpha\mathbf{A}|\psi\rangle + \beta\mathbf{A}|\phi\rangle. \tag{A.14}$$

With a nod to the mathematicians, it is notationally useful to define

$$|\mathbf{A}\psi\rangle = \mathbf{A}|\psi\rangle. \tag{A.15}$$

A linear transformation that preserves the magnitudes of all vectors is called *unitary*, because it follows from linearity that all magnitudes will be preserved if and only if unit vectors (vectors of magnitude 1) are taken into unit vectors. It also follows from linearity that if a linear transformation **U** is unitary then it must preserve not only the inner products of arbitrary vectors with themselves, but also the inner products of arbitrary pairs of vectors. This follows straightforwardly for two general vectors $|\phi\rangle$ and $|\psi\rangle$ from the fact that **U** preserves the magnitudes of both of them, as well as the magnitudes of the vectors $|\phi\rangle + |\psi\rangle$ and $|\phi\rangle + i|\psi\rangle$.

One can associate with any given vector $|\phi\rangle$ the *linear functional* that takes every vector $|\psi\rangle$ into the number $\langle\phi|\psi\rangle$. Linearity follows from property (A.7) of the inner product. The set of all such linear functionals is itself a vector space, called the *dual space* of the original space. The functional associated with the vector $\alpha|\phi\rangle + \beta|\psi\rangle$ is the sum of α^* times the functional associated with $|\phi\rangle$ and β^* times the functional associated with $|\psi\rangle$. It is an easy exercise to show that *any* linear functional on the original space is associated with some vector in the dual space. Dirac called vectors in the original space *ket vectors* and vectors in the dual space *bra vectors*. He denoted the bra associated with the ket $|\phi\rangle$ by the symbol $\langle\phi|$, so that the symbol $\langle\phi|\psi\rangle$ can equally well be viewed as the inner product of the two kets $|\phi\rangle$ and $|\psi\rangle$ or as a compact way of expressing the action $\langle\phi|(|\psi\rangle)$ of the associated linear

functional $\langle\phi|$ on the vector $|\psi\rangle$. Note that one has

$$\langle\alpha\phi + \beta\psi| = \alpha^*\langle\phi| + \beta^*\langle\psi|. \tag{A.16}$$

A linear transformation $\mathbf{A}$ on the space of ket vectors induces a linear transformation $\mathbf{A}^\dagger$ (called "A-adjoint") on the dual space of bra vectors, according to the rule

$$\langle\mathbf{A}\psi| = \langle\psi|\mathbf{A}^\dagger. \tag{A.17}$$

The operation adjoint to the trivial linear transformation that multiplies by a given complex number is multiplication by the complex conjugate of that number.

It is convenient to extend the dagger notation to the vectors themselves, defining

$$\big(|\psi\rangle\big)^\dagger = \langle\psi|, \tag{A.18}$$

so that the bra dual to a given ket is viewed as adjoint to that ket. The definition (A.17) of $\mathbf{A}^\dagger$ then becomes

$$\big(|\mathbf{A}\psi\rangle\big)^\dagger = \langle\psi|\mathbf{A}^\dagger, \tag{A.19}$$

or, with (A.15),

$$\big(\mathbf{A}|\psi\rangle\big)^\dagger = \langle\psi|\mathbf{A}^\dagger, \tag{A.20}$$

which provides a simple example of a very general rule that the adjoint of a product of quantities is the product of their adjoints taken in the opposite order. Another instance of the rule which follows from (A.20) is that

$$\langle\phi|\big(\mathbf{AB}\big)^\dagger = \langle\mathbf{AB}\phi| = \langle\mathbf{B}\phi|\mathbf{A}^\dagger = \langle\phi|\mathbf{B}^\dagger\mathbf{A}^\dagger. \tag{A.21}$$

Since this holds for arbitrary $\langle\phi|$ we have

$$\big(\mathbf{AB}\big)^\dagger = \mathbf{B}^\dagger\mathbf{A}^\dagger. \tag{A.22}$$

Although the adjoint $\mathbf{A}^\dagger$ of a linear transformation $\mathbf{A}$ on kets is a linear transformation on bras, one can also define its action on kets. One does so by requiring that the action of $\langle\phi|$ on $\mathbf{A}^\dagger|\psi\rangle$ should be equal to the action of $\langle\phi|\mathbf{A}^\dagger$ on $|\psi\rangle$. This amounts to stipulating that the symbol $\langle\phi|\mathbf{A}^\dagger|\psi\rangle$ should be unambiguous; it does not matter whether it is read as $(\langle\phi|\mathbf{A}^\dagger)|\psi\rangle$ or as $\langle\phi|(\mathbf{A}^\dagger|\psi\rangle)$. Implicit in this definition is the fact that a vector is completely defined by giving its inner product with all vectors. This in turn follows from the fact that a vector $|\psi\rangle$ can be defined by giving all the amplitudes α_x in its expansion (A.3) in the complete orthonormal set $|x\rangle$. But $\alpha_x = \langle x|\psi\rangle$. Similarly, a linear operator $\mathbf{A}$ is completely defined by giving its *matrix elements* $\langle\phi|\mathbf{A}|\psi\rangle$ for arbitrary pairs of vectors, since the subset $\langle x|\mathbf{A}|y\rangle$ is already enough to determine its action on a general vector (A.3).

Note that any matrix element of $\mathbf{A}^\dagger$ is equal to the complex conjugate of the *transposed* (with ϕ and ψ exchanged) matrix element of $\mathbf{A}$:

$$\langle\phi|\mathbf{A}^\dagger|\psi\rangle = \langle\mathbf{A}\phi|\psi\rangle = \langle\psi|\mathbf{A}\phi\rangle^* = \langle\psi|\mathbf{A}|\phi\rangle^*. \tag{A.23}$$

It follows from this that

$$\left(\mathbf{A}^\dagger\right)^\dagger = \mathbf{A}. \tag{A.24}$$

Since a unitary transformation $\mathbf{U}$ preserves inner products, we have

$$\langle\phi|\psi\rangle = \langle\mathbf{U}\phi|\mathbf{U}\psi\rangle = \langle\phi|\mathbf{U}^\dagger\mathbf{U}|\psi\rangle, \tag{A.25}$$

and therefore

$$\mathbf{U}^\dagger\mathbf{U} = \mathbf{1}, \tag{A.26}$$

where $\mathbf{1}$ is the unit (identity) operator that takes every vector into itself. It follows from (A.26) that

$$\mathbf{U}\mathbf{U}^\dagger\mathbf{U} = \mathbf{U}. \tag{A.27}$$

In a finite-dimensional vector space a unitary transformation $\mathbf{U}$ always takes an orthonormal basis into another orthonormal basis, so any $\mathbf{U}$ clearly has a right inverse – the linear transformation that takes the second basis back into the first. Multiplying (A.27) on the right by that inverse tells us that

$$\mathbf{U}\mathbf{U}^\dagger = \mathbf{1}, \tag{A.28}$$

so $\mathbf{U}^\dagger$ and $\mathbf{U}$ are inverses regardless of the order in which they act.

The vector $|\psi\rangle$ is an *eigenvector* of the linear operator $\mathbf{A}$ if the action of $\mathbf{A}$ on $|\psi\rangle$ is simply to multiply it by a complex number a, called an *eigenvalue* of $\mathbf{A}$:

$$\mathbf{A}|\psi\rangle = a|\psi\rangle. \tag{A.29}$$

Since the number a can be expressed as $a = \langle\psi|\mathbf{A}|\psi\rangle/\langle\psi|\psi\rangle$, it follows from (A.23) that if $\mathbf{A} = \mathbf{A}^\dagger$ (such operators are said to be *self-adjoint* or *Hermitian*) then a is a real number. Eigenvalues of Hermitian operators are necessarily real.

Since $\mathbf{A}$ is Hermitian and a is a real number, it follows from (A.29) (by forming the adjoints of both sides) that

$$\langle\psi|\mathbf{A} = a\langle\psi|, \tag{A.30}$$

so the vector dual to an eigenket of a Hermitian operator is an eigenbra with the same eigenvalue. It follows immediately that if $|\phi\rangle$ is another eigenvector of $\mathbf{A}$ with eigenvalue a', then

$$a\langle\psi|\phi\rangle = \langle\psi|\mathbf{A}|\phi\rangle = a'\langle\psi|\phi\rangle, \tag{A.31}$$

so if $a' \neq a$ then $\langle \psi | \phi \rangle = 0$: eigenvectors of a Hermitian operator with different eigenvalues are orthogonal.

It can be shown that for any Hermitian operator $\mathbf{A}$, one can choose an orthonormal basis for the entire D-dimensional space whose members are eigenvectors of $\mathbf{A}$. The basis is unique if and only if all the D eigenvalues of $\mathbf{A}$ are distinct. In the contrary case (in which $\mathbf{A}$ is said to be degenerate) one can pick arbitrary orthonormal bases within each of the subspaces spanned by eigenvectors of $\mathbf{A}$ with the same eigenvalue. More generally, if $\mathbf{A}, \mathbf{B}, \mathbf{C}, \ldots$ are mutually commuting Hermitian operators then one can choose an orthonormal basis whose members are eigenstates of every one of them.

If $\mathbf{B}$ is any linear operator, then $\mathbf{A}_1 = \mathbf{B} + \mathbf{B}^\dagger$ and $\mathbf{A}_2 = i\left(\mathbf{B}^\dagger - \mathbf{B}\right)$ are both Hermitian, and commute if $\mathbf{B}$ and $\mathbf{B}^\dagger$ commute. Since a joint eigenvector of $\mathbf{A}_1$ and $\mathbf{A}_2$ is also a joint eigenvector of $\mathbf{B} = \mathbf{A}_1 + i\mathbf{A}_2$ and $\mathbf{B}^\dagger = \mathbf{A}_1 - i\mathbf{A}_2$, it follows that if $\mathbf{B}$ commutes with $\mathbf{B}^\dagger$ then one can choose an orthonormal basis of eigenvectors of $\mathbf{B}$. In particular, since a unitary transformation $\mathbf{U}$ satisfies $\mathbf{U}\mathbf{U}^\dagger = \mathbf{U}^\dagger\mathbf{U} = \mathbf{1}$, one can choose an orthonormal basis consisting of eigenvectors of $\mathbf{U}$. Since unitary transformations preserve the magnitudes of vectors, the eigenvalues of $\mathbf{U}$ must be complex numbers of modulus 1. In the quantum theory such complex numbers are often called *phase factors*.

Given two vector spaces of dimensions D_1 and D_2, and given any two vectors $|\psi_1\rangle$ and $|\psi_2\rangle$ in the two spaces, one associates with each such pair a *tensor product* $|\psi_1\rangle \otimes |\psi_2\rangle$ (often the tensor-product sign $\otimes$ is omitted) which is bilinear:

$$\begin{aligned} |\psi_1\rangle \otimes \big(\alpha|\psi_2\rangle + \beta|\phi_2\rangle\big) &= \alpha|\psi_1\rangle \otimes |\psi_2\rangle + \beta|\psi_1\rangle \otimes |\phi_2\rangle, \\ \big(\alpha|\psi_1\rangle + \beta|\phi_1\rangle\big) \otimes |\psi_2\rangle &= \alpha|\psi_1\rangle \otimes |\psi_2\rangle + \beta|\phi_1\rangle \otimes |\psi_2\rangle. \end{aligned} \tag{A.32}$$

With the further rule that $|\psi_1\rangle \otimes |\psi_2\rangle = |\phi_1\rangle \otimes |\phi_2\rangle$ only if $|\phi_1\rangle$ and $|\phi_2\rangle$ are scalar multiples of $|\psi_1\rangle$ and $|\psi_2\rangle$, one easily sees that the set of all tensor products of vectors from the two spaces forms a vector space of dimension $D_1 D_2$.

One defines the inner product of $|\psi_1\rangle \otimes |\psi_2\rangle$ with $|\phi_1\rangle \otimes |\phi_2\rangle$ to be the ordinary product $\langle\psi_1|\phi_1\rangle\langle\psi_2|\phi_2\rangle$ of the inner products in the two original spaces. Given orthonormal bases for each of the two spaces, the set of tensor products of all pairs of vectors from the two bases forms an orthonormal basis for the tensor-product space. If $\mathbf{A}_1$ and $\mathbf{A}_2$ are linear operators on the two spaces, one defines the tensor-product operator $\mathbf{A}_1 \otimes \mathbf{A}_2$ to satisfy

$$\big(\mathbf{A}_1 \otimes \mathbf{A}_2\big)\big(|\psi_1\rangle \otimes |\psi_2\rangle\big) = |\mathbf{A}_1\psi_1\rangle \otimes |\mathbf{A}_2\psi_2\rangle = \big(\mathbf{A}_1|\psi_1\rangle\big) \otimes \big(\mathbf{A}_2|\psi_2\rangle\big), \tag{A.33}$$

and easily shows that it can be extended to a linear operator on the entire tensor-product space.

All of this generalizes in the obvious way to n-fold tensor products of n vector spaces.

If $\mathbf{A}$ is a linear operator whose eigenvectors constitute an orthonormal basis – i.e. if $\mathbf{A}$ is Hermitian or, more generally, if $\mathbf{A}$ and $\mathbf{A}^\dagger$ commute – and if f is a function taking complex numbers to complex numbers, then one can define $f(\mathbf{A})$ by specifying that each eigenvector $|\phi\rangle$ of $\mathbf{A}$, in the basis with eigenvalue a, is also an eigenvector of $f(\mathbf{A})$ with eigenvalue $f(a)$. This defines $f(\mathbf{A})$ on a basis, and it can therefore be extended to arbitrary vectors by requiring it to be linear. It follows from this definition that if $f(z)$ is a polynomial or convergent power series in z then $f(\mathbf{A})$ is the corresponding polynomial or convergent power series in $\mathbf{A}$.

In Dirac notation one defines the *outer product* of two vectors $|\phi\rangle$ and $|\psi\rangle$ to be the linear operator, denoted by $|\phi\rangle\langle\psi|$, that takes any vector $|\gamma\rangle$ into $|\phi\rangle$ multiplied by the inner product $\langle\psi|\gamma\rangle$:

$$\big(|\phi\rangle\langle\psi|\big)|\gamma\rangle = |\phi\rangle\big(\langle\psi|\gamma\rangle\big). \tag{A.34}$$

As is always the case with Dirac notation, the point is to define things in such a way that the evaluation of an ambiguous expression such as $|\phi\rangle\langle\psi|\gamma\rangle$ does not depend on how you read it; the notation is designed always to enforce the associative law.

Note that $|\psi\rangle\langle\psi|$ is the projection operator onto the one-dimensional subspace spanned by the unit vector $|\psi\rangle$; i.e. any vector $|\gamma\rangle$ can be written as the sum of a vector $|\gamma\rangle_\parallel$ in the one-dimensional subspace and a vector $|\gamma\rangle_\perp$ perpendicular to the one-dimensional subspace, and

$$\big(|\psi\rangle\langle\psi|\big)|\gamma\rangle = |\gamma\rangle_\parallel. \tag{A.35}$$

Similarly, if one has a set of orthonormal vectors $|\psi_i\rangle$ then $\sum_i |\psi_i\rangle\langle\psi_i|$ projects onto the subspace spanned by all the $|\psi_i\rangle$. If the orthonormal set is a complete orthonormal set – for example $|x\rangle$, $x = 0, \ldots, N$ – then the set spans the entire vector space and the projection operator is the unit operator $\mathbf{1}$:

$$\sum_{x=0}^{N} |x\rangle\langle x| = \mathbf{1}. \tag{A.36}$$

This trivial identity can be surprisingly helpful. Any vector $|\psi\rangle$, for example, satisfies

$$|\psi\rangle = \mathbf{1}|\psi\rangle = \sum_x |x\rangle\langle x|\psi\rangle, \tag{A.37}$$

which tells us that the amplitudes α_x appearing in the expansion (A.3) of $|\psi\rangle$ are just the inner products $\langle x|\psi\rangle$. Similarly, any linear operator

A satisfies

$$|\mathbf{A}\rangle = \mathbf{1A1} = \left(\sum_x |x\rangle\langle x|\right)\mathbf{A}\left(\sum_y |y\rangle\langle y|\right) = \sum_{xy} |x\rangle\langle y|\big(\langle x|\mathbf{A}|y\rangle\big), \tag{A.38}$$

which reveals the matrix elements $\langle x|\mathbf{A}|y\rangle$ to be the expansion coefficients of the operator **A** in the "operator basis" $|x\rangle\langle y|$. And note that

$$\langle x|\mathbf{AB}|y\rangle = \langle x|\mathbf{A1B}|y\rangle = \sum_z \langle x|\mathbf{A}|z\rangle\langle z|\mathbf{B}|y\rangle, \tag{A.39}$$

which gives the familiar matrix-multiplication rule for constructing the matrix of a product out of the matrix elements of the individual operators.

If you prefer to think of vectors in terms of their components in a specific basis, then you might note that the (ket) vector $|\psi\rangle$, with the expansion (A.3) with amplitudes α_x in the orthonormal basis $|x\rangle$, can be represented by a column vector:

$$|\psi\rangle \longrightarrow \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_N \end{pmatrix}. \tag{A.40}$$

The associated bra vector is then the row vector:

$$\langle\psi| \longrightarrow \begin{pmatrix} \alpha_0^* & \alpha_1^* & \dots & \alpha_N^* \end{pmatrix}. \tag{A.41}$$

If

$$|\phi\rangle \longrightarrow \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_N \end{pmatrix}, \tag{A.42}$$

then the inner product $\langle\phi|\psi\rangle$ is given by the ordinary matrix product of the row and column vectors:

$$\langle\phi|\psi\rangle = \begin{pmatrix} \beta_0^* & \beta_1^* & \dots & \beta_N^* \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_N \end{pmatrix}. \tag{A.43}$$

The outer product $|\psi\rangle\langle\phi|$ is also a matrix product:

$$|\psi\rangle\langle\phi| = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_N \end{pmatrix} \begin{pmatrix} \beta_0^* & \beta_1^* & \dots & \beta_N^* \end{pmatrix}. \tag{A.44}$$

Note that in Dirac notation (A.43) is nothing more than the statement that

$$\langle \phi|\psi\rangle = \langle \phi|\mathbf{1}|\psi\rangle = \sum_x \langle \phi|x\rangle\langle x|\psi\rangle = \sum_x \langle x|\phi\rangle^*\langle x|\psi\rangle, \quad \text{(A.45)}$$

while (A.44) asserts that

$$\langle x|\Big(|\psi\rangle\langle\phi|\Big)|y\rangle = \langle x|\psi\rangle\langle\phi|y\rangle = \langle x|\psi\rangle\langle y|\phi\rangle^*. \quad \text{(A.46)}$$