

Mathematics of Quantum Computation I

Wim van Dam

Departments of Computer Science and Physics, University of California at Santa Barbara, Santa Barbara, CA 93106-5110, USA

These are notes for the Spring 2007 graduate course “Quantum Computation and Quantum Information”.

LINEAR ALGEBRA AND BRAKET NOTATION

Complex Values: Let $\alpha \in \mathbb{C}$, then we can write this complex value as $\alpha = x + yi$ with the real and imaginary components $x, y \in \mathbb{R}$. It is often useful to write the value as $\alpha = re^{i\varphi}$ with the norm $r \in \mathbb{R}_{\geq 0}$ and the phase $\varphi \in [0, 2\pi)$. The “norm squared” of α equals $|\alpha|^2 = x^2 + y^2 = r^2$. The complex conjugate of α is defined by $\bar{\alpha} = \alpha^* = x - yi = re^{-i\varphi}$, which can be used in $|\alpha|^2 = \alpha\alpha^*$. The norm $|\alpha| = \sqrt{x^2 + y^2} = r$ obeys the triangle inequality $|\alpha + \beta| \leq |\alpha| + |\beta|$ for all $\alpha, \beta \in \mathbb{C}$.

Finite Dimensional Hilbert Space: Let \mathcal{A} be a finite set of $N = |\mathcal{A}|$ basis states. A quantum state, which is in superposition over all basis states \mathcal{A} , is represented by a complex valued vector $(\alpha_1, \dots, \alpha_N) \in \mathbb{C}^N$ with ℓ_2 -norm $(\sum_{j=1}^N |\alpha_j|^2)^{1/2} = 1$. In Dirac’s bra-ket notation, a column vector is denoted by a $|ket\rangle$ and a row vector by a $\langle bra|$. If $|\psi\rangle = \sum_{x \in \mathcal{A}} \alpha_x |x\rangle$, then $\langle\psi| := \sum_{x \in \mathcal{A}} \alpha_x^* \langle x|$ (note the complex conjugation of α_x). Given column vector $|\psi\rangle$, the row vector $\langle\psi|$ is also denoted by $|\psi\rangle^\dagger$ and is called the *conjugate transpose* of $|\psi\rangle$. If $\mathcal{A} = \{1, \dots, N\}$, we can write in vector notation:

$$|\psi\rangle^\dagger = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{pmatrix}^\dagger = (\alpha_1^* \ \alpha_2^* \ \dots \ \alpha_N^*) = \langle\psi|. \quad (1)$$

We equip this space with an *inner product* $\langle \cdot | \cdot \rangle : \mathbb{C}^N \times \mathbb{C}^N \rightarrow \mathbb{C}$ such that it becomes a *finite dimensional Hilbert space*. For the vectors

$$|\psi\rangle = \sum_{x \in \mathcal{A}} \alpha_x |x\rangle \text{ and } |\phi\rangle = \sum_{x \in \mathcal{A}} \beta_x |x\rangle \quad (2)$$

the inner product $\langle\psi|\phi\rangle$ is expressed by the *braket*

$$\langle\psi|\phi\rangle = \sum_{x \in \mathcal{A}} \alpha_x^* \beta_x = \langle\phi|\psi\rangle^*. \quad (3)$$

The *norm* of a vector in \mathbb{C}^N is defined by $\| |\psi\rangle \| := \sqrt{\langle\psi|\psi\rangle}$, which for a valid state representation is always one: $\sum_{x \in \mathcal{A}} \alpha_x \alpha_x^* = 1$ (the normalization restriction). For vectors the triangle inequality holds as well: $\| \alpha |\psi\rangle + \beta |\phi\rangle \| \leq \| \alpha |\psi\rangle \| + \| \beta |\phi\rangle \|$.

The *outer product* $|\cdot\rangle\langle\cdot| : \mathbb{C}^N \times \mathbb{C}^N \rightarrow \mathbb{C}^{N \times N}$ maps two N -dimensional vectors to an $N \times N$ matrix:

$$|\psi\rangle\langle\phi| = \sum_{x, y \in \mathcal{A}} \alpha_x \beta_y^* |x\rangle\langle y|, \quad (4)$$

where $|x\rangle\langle y|$ is the all-zero matrix with one 1 value in the x -th row and the y -th column.

Braket Calculus: The difference between the inner and the outer product shows that ‘multiplying’ bras and kets does not commute: $\langle\psi|\phi\rangle \neq |\phi\rangle\langle\psi|$. However, this multiplication *is* associative and distributive. Hence, for example, $|\psi\rangle(\langle\phi| + \langle\phi'|) = |\psi\rangle\langle\phi| + |\psi\rangle\langle\phi'|$ and $(|\psi\rangle\langle\phi|)(|\psi\rangle\langle\phi|) = |\psi\rangle(\langle\phi|\phi\rangle)\langle\psi| = |\psi\rangle\langle\psi|$ (because $\langle\phi|\phi\rangle = 1$).

Measurement Projection: According to quantum mechanics, the ‘inner product squared’ $|\langle\psi|\phi\rangle|^2 = \langle\psi|\phi\rangle\langle\phi|\psi\rangle$ between two states $|\psi\rangle$ and $|\phi\rangle$ gives the probability that one observes the outcome “ $|\psi\rangle$ ” when one observes the state “ $|\phi\rangle$ ”. It is straightforward to verify that $0 \leq |\langle\psi|\phi\rangle|^2 \leq 1$. If $\langle\psi|\phi\rangle = 0$, the two states are *orthogonal*. If $|\langle\psi|\phi\rangle| = 1$, then the two states are identical *up to a general phase factor* (because we can still have $\langle\psi|\phi\rangle = e^{i\gamma}$). Although mathematically present, such a general phase difference can never be observed in reality, hence it has no physical relevance.

Tensor Product Construction: We can combine the spaces \mathbb{C}^N and \mathbb{C}^M to a joint space $\mathbb{C}^{NM} := \mathbb{C}^N \otimes \mathbb{C}^M$. If \mathcal{A} and \mathcal{B} are the respective basis sets of these two spaces, then the joint basis of $\mathbb{C}^N \otimes \mathbb{C}^M$ is given by the Cartesian product $\mathcal{A} \times \mathcal{B}$. As a result, using the *tensor product* $\otimes : \mathbb{C}^N \times \mathbb{C}^M \rightarrow \mathbb{C}^{NM}$, we can combine the states

$$|\psi\rangle = \sum_{x \in \mathcal{A}} \alpha_x |x\rangle \in \mathbb{C}^N \text{ and } |\phi\rangle = \sum_{y \in \mathcal{B}} \beta_y |y\rangle \in \mathbb{C}^M \quad (5)$$

to the tensor product state

$$|\psi\rangle \otimes |\phi\rangle = |\psi, \phi\rangle = \sum_{x \in \mathcal{A}, y \in \mathcal{B}} \alpha_x \beta_y |x, y\rangle \in \mathbb{C}^{NM}. \quad (6)$$

For the conjugate transpose of a tensor product it holds that $(|\psi\rangle \otimes |\phi\rangle)^\dagger = \langle\psi| \otimes \langle\phi|$.

If we assume $\mathcal{A} = \{1, \dots, N\}$ and $\mathcal{B} = \{1, \dots, M\}$, then this tensor product equation is described in vector notation by

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{pmatrix} \otimes \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_M \end{pmatrix} = \begin{pmatrix} \alpha_1 \beta_1 \\ \alpha_1 \beta_2 \\ \vdots \\ \alpha_1 \beta_M \\ \alpha_2 \beta_1 \\ \vdots \\ \vdots \\ \alpha_N \beta_M \end{pmatrix}. \quad (7)$$

If $|\psi\rangle$ and $|\phi\rangle$ are norm one vectors, then so is $|\psi\rangle \otimes |\phi\rangle$. Note that the tensor product does not commute: $|\psi\rangle \otimes |\phi\rangle \neq |\phi\rangle \otimes |\psi\rangle$, but that it is associative and distributive. For example, $|\psi\rangle \otimes (|\phi\rangle \otimes |\phi'\rangle) = (|\psi\rangle \otimes |\phi\rangle) \otimes |\phi'\rangle$, and $|\psi\rangle \otimes (\alpha |\phi\rangle + \beta |\phi'\rangle) = |\psi\rangle \otimes \alpha |\phi\rangle + |\psi\rangle \otimes \beta |\phi'\rangle = \alpha |\psi, \phi\rangle + \beta |\psi, \phi'\rangle$.

Unitary Operations: The group of norm preserving, linear operations on \mathbb{C}^N is the group $U(N)$ of unitary, complex valued $N \times N$ matrices $U \in \mathbb{C}^{N \times N}$ that obey the equality $U \cdot U^\dagger = I$. Here U^\dagger is the *Hermitian conjugate* (or the *conjugate transpose*) of U defined by $U_{ij}^\dagger := U_{ji}^*$ for all $1 \leq i, j \leq N$, and I is the N -dimensional identity matrix. As these operations are linear, we have

$$U|\psi\rangle = \sum_{x \in \mathcal{A}} \alpha_x U|x\rangle \quad (8)$$

for all $|\psi\rangle \in \mathbb{C}^N$. Hence, if we know the values of U on the basis states $|x \in \mathcal{A}\rangle$, we know the values of U on all quantum states in \mathbb{C}^N . We can describe $U \in U(N)$ as a summation of outer products by

$$U := \sum_{x,y \in \mathcal{A}} U_{xy} |x\rangle\langle y|, \quad (9)$$

or equivalently $U_{xy} := \langle x|U|y\rangle$, such that by linearity we see that

$$U|\psi\rangle = \sum_{x,y \in \mathcal{A}} U_{xy} |x\rangle\langle y| \sum_{z \in \mathcal{A}} \alpha_z |z\rangle = \sum_{x,z \in \mathcal{A}} \alpha_z U_{xz} |x\rangle. \quad (10)$$

Unitary matrices are inner product preserving (and hence also norm preserving) as is shown by $\langle \phi|\psi\rangle = \langle \phi|I|\psi\rangle = \langle \phi|U^\dagger U|\psi\rangle = \langle \phi'|\psi'\rangle$, where $|\phi'\rangle := U|\phi\rangle$ and $|\psi'\rangle := U|\psi\rangle$. This shows that U is unitary if and only if the row vectors of U form an orthonormal basis of \mathbb{C}^N (similarly for the columns of U).

Just as with vectors, we can define the tensor product between two matrices. Specifically, if $U \in U(N)$ and $W \in U(M)$ are unitary matrices defined by

$$U := \sum_{x,y \in \mathcal{A}} U_{xy} |x\rangle\langle y| \text{ and } W := \sum_{p,q \in \mathcal{B}} W_{pq} |p\rangle\langle q|, \quad (11)$$

then for the tensor product $\otimes : \mathbb{C}^{N \times N} \times \mathbb{C}^{M \times M} \rightarrow \mathbb{C}^{NM \times NM}$ we have

$$U \otimes W = \sum_{x,y \in \mathcal{A}} \sum_{p,q \in \mathcal{B}} U_{xy} W_{pq} |x,p\rangle\langle y,q| \in \mathbb{C}^{NM \times NM}. \quad (12)$$

This matrix acts on the space $\mathbb{C}^{NM} = \mathbb{C}^N \otimes \mathbb{C}^M$ spanned by the set of basis states $\mathcal{A} \times \mathcal{B}$. For the states $|\psi\rangle \in \mathbb{C}^N$ and $|\phi\rangle \in \mathbb{C}^M$ we have $(U \otimes W)(|\psi\rangle \otimes |\phi\rangle) = U|\psi\rangle \otimes W|\phi\rangle \in \mathbb{C}^{NM}$. Again assuming $\mathcal{A} = \{1, \dots, N\}$ and $\mathcal{B} = \{1, \dots, M\}$, the tensor product of two matrices is described in matrix notation by

$$\begin{aligned} U \otimes W &= \begin{pmatrix} U_{11}W & U_{12}W & \cdots & U_{1N}W \\ U_{21}W & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ U_{N1}W & \cdots & \cdots & U_{NN}W \end{pmatrix} \\ &= \begin{pmatrix} U_{11}W_{11} & U_{11}W_{12} & \cdots & U_{1N}W_{1M} \\ U_{11}W_{21} & \ddots & & U_{1N}W_{2M} \\ \vdots & & \ddots & \vdots \\ U_{N1}W_{M1} & \cdots & \cdots & U_{NN}W_{MM} \end{pmatrix} \in \mathbb{C}^{NM \times NM}. \end{aligned} \quad (13)$$

As was the case with vectors, the tensor product of matrices is not commutative, but it is distributive and associative. Also, if $U, U' \in U(N)$ and $W, W' \in U(M)$, then $(U \otimes W)(U' \otimes W') = UU' \otimes WW'$; if U, W are unitary, then so is $U \otimes W$ and $(U \otimes W)^\dagger = U^\dagger \otimes W^\dagger$.

Eigenvector / Eigenvalue Decomposition: We can decompose a unitary matrix $U \in U(N)$ into its eigenvectors $|\psi_1\rangle, \dots, |\psi_N\rangle$ and its corresponding eigenvalues $\lambda_1, \dots, \lambda_N \in \mathbb{C}$. With these values we can express the operator as

$$U = \sum_{i=1}^N \lambda_i |\psi_i\rangle\langle \psi_i|. \quad (14)$$

The unitarity of U corresponds with the requirement that all eigenvalues λ_i have norm one, and that the eigenvectors form an orthonormal basis of \mathbb{C}^N . The identity matrix I has for all eigenvalues $\lambda_i = 1$. The conjugate transpose of this U is given by

$$U^\dagger = \sum_{i=1}^N \lambda_i^* |\psi_i\rangle\langle \psi_i|. \quad (15)$$


When U is as above and $W \in U(M)$ has eigenvector decomposition $W = \sum_{j=1}^M \mu_j |\phi_j\rangle\langle \phi_j|$ then for the tensor product we have


$$U \otimes W = \sum_{i=1}^N \sum_{j=1}^M \lambda_i \mu_j |\psi_i, \phi_j\rangle\langle \psi_i, \phi_j|. \quad (16)$$


QUANTUM COMPUTING WITH CIRCUITS


Quantum Bits: The typical setting for a quantum circuit is quantum mechanical system that is described by an n -fold tensor product of two dimensional Hilbert spaces: $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \dots \otimes \mathbb{C}^2 = \mathbb{C}^{2^n}$ (where each \mathbb{C}^2 corresponds to a single qubit). The elementary quantum gates that we can apply to an initial state $|0, \dots, 0\rangle$ are unitary operators that act only a small number of qubits. For example, if we apply the NOT gate $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ to the second qubit, then the overall unitary operator is described by $I \otimes X \otimes I \otimes \dots \otimes I \in U(2^n)$, where the I are the identity operators on the qubits $1, 3, \dots, n$. For operators that act on two non-adjacent qubits, the notation becomes a bit tricky. Consider for example a CNOT gate that acts on the first and the last qubit. To avoid these problems one can introduce the notation where the identity operators are omitted, and a subscript is used to indicate on which qubit the gates act. Hence the previous NOT circuit has the much shorter description $X_2 \in U(2^n)$, and the CNOT example becomes $\text{CNOT}_{1,n} \in U(2^n)$. Regardless, it is often advisable to draw a quantum circuit diagram to explain the operation.


Standard Quantum Gates: The following gates are standard single qubit gates:


Identity: $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ 


NOT: $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ 


$Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ 

$Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ 

Hadamard: $H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ 


$S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$ 


$\frac{\pi}{8}$ -phase gate: $T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}$ 


θ -Phase Rotation: $R_z(\theta) = \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix}$ 

Note that a global phase difference between two matrices does not really change their physical effect, and that the R_z gates can also be defined by $R_z(\theta) : \alpha|0\rangle + \beta|1\rangle \mapsto \alpha|0\rangle + e^{i\theta}\beta|1\rangle$. Note also that therefore the $\frac{\pi}{8}$ -gate equals the gate $R_z(\frac{\pi}{4})$.


Typically, two qubit gates are of the kind where a *control bit* determines whether or not a single qubit operation is applied to the *target bit* or not.

CNOT = $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ 

$C-R_z(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\theta} \end{pmatrix}$ 

Controlled- $U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & U_{11} & U_{12} \\ 0 & 0 & U_{21} & U_{22} \end{pmatrix}$ 

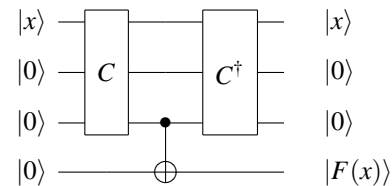
The three qubit, CCNOT gate is crucial for the implementation of classical, reversible computation.

CCNOT = $\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ 

Circuit Calculus: Each quantum gate has a matrix representation, and we assume that the ordering of the dimensions is determined by the alphabetical ordering on the bit strings $\{0, 1\}^n$. (See the Exercises of the course.) The joint behavior of gates that work in parallel is calculated with the help of tensor products, while sequential gates are expressed by matrix products. In both cases one should pay attention to the order of multiplication.

Universal Reversible Computation: With a CCNOT gate we can implement an AND-operation by CCNOT : $|a, b, 0\rangle \mapsto |a, b, ab\rangle$ for all $a, b \in \{0, 1\}$. Hence, using CCNOT and NOT gates, we can implement any Boolean function $F : \{0, 1\}^n \rightarrow \{0, 1\}$ as efficient as is possible classically. However, such an implementation will not erase the input value x and typically also produces ‘garbage’ bits that are a side-effect of the computation. (Take for example the computation of $F(x, y, z) = xyz$ by the sequence of transformations $|x, y, z, 0, 0\rangle \mapsto |x, y, z, xy, 0\rangle \mapsto |x, y, z, xy, xyz\rangle = |x, y, z, xy, F(x, y, z)\rangle$, which has as garbage the intermediate bit value xy .) Because of the reversibility requirement it is impossible to get rid of the input bits, but it is possible to erase the garbage bits as follows.

Note that each quantum circuit can be reversed. Hence, if there is a circuit C that implements the transformation $|x, 0, 0\rangle \mapsto |x, g_x, F(x)\rangle$ (with g_x the garbage bits specifically for the input $x \in \{0, 1\}^n$), then the inverse circuit $C^{-1} = C^\dagger$ implements the mapping $|x, g_x, F(x)\rangle \mapsto |x, 0, 0\rangle$. Now, by applying a CNOT between C and C^\dagger we get the following circuit



which implements the desired transformation $|x, 0, 0, 0\rangle \mapsto |x, 0, 0, F(x)\rangle$ for all x .

Note that this construction applies to all possible quantum circuits C , which gives us the following important result.

Clean Quantum Computation Theorem: If there is a quantum circuit C that implements the unitary mapping $|x, 0, 0\rangle \mapsto |x, g_x, F(x)\rangle$ for a Boolean function $F : \{0, 1\}^n \rightarrow \{0, 1\}^m$, then there exists a quantum circuit (only twice the size of C) that implements the clean computation $|x, 0, 0\rangle \mapsto |x, 0, F(x)\rangle$. Notice that it is crucial for this construction that we have clean working qubits bits lying around that we can use during the computation.

Further Reading: For more information see Chapters 2–4 in “An Introduction to Quantum Computation”, Phillip Kaye, Raymond Laflamme and Michele Mosca, Oxford University Press (2007)

Acknowledgment: The circuits in these exercises were drawn using the Q-circuit L^AT_EXpackage of Bryan Eastin and Steven T. Flammia.