## Thursday 9/9: Operations on Qubits

Here is a statement that uses a lot of notational fuss to express the simplest of ideas:

**Proposition**: For any $m \times n$ matrix $A$, $p \times q$ matrix $B$, $n$-vector $\mathbf{x}$ and $q$-vector $\mathbf{y}$,

$$(A \otimes B) \cdot (\mathbf{x} \otimes \mathbf{y}) \;=\; (A\mathbf{x}) \otimes (B\mathbf{y})\,.$$

**Proof.** The dimensions are consistent: both sides give a column vector of $mp$ entries. Showing equality is where our effort to interpret vectors $\mathbf{x}$ as functions $\mathbf{x}(u)$ of their indices in binary notation may help. Under this view, $\mathbf{z} \;=\; \mathbf{x} \otimes \mathbf{y}$ gives the function $\mathbf{z}(uv) \;=\; \mathbf{x}(u)\mathbf{y}(v)$, where $uv$ means concatenation of binary strings, while the right-hand side is an ordinary numeric product. And a matrix $A$ gives the two-argument function $A(u,w) \;=\; a_{u,w}$. The vector $\mathbf{x}' = A\mathbf{x}$ becomes the function mapping a row-index $u$ to $\mathbf{x}'(u) \;=\; \sum_w A(u,w)\mathbf{x}(w)$. Thus, putting $\mathbf{z}' = (A\mathbf{x}) \otimes (B\mathbf{y})$, the right-hand side is the function

$$\mathbf{z}'(uv) \;=\; \mathbf{x}'(u)\mathbf{y}'(v) \;=\; \left(\sum_w A(u,w)\mathbf{x}(w)\right)\left(\sum_t B(v,t)\mathbf{y}(t)\right)$$

Now by usual rules of re-ordering summations, the right-hand side of this can be rearranged as

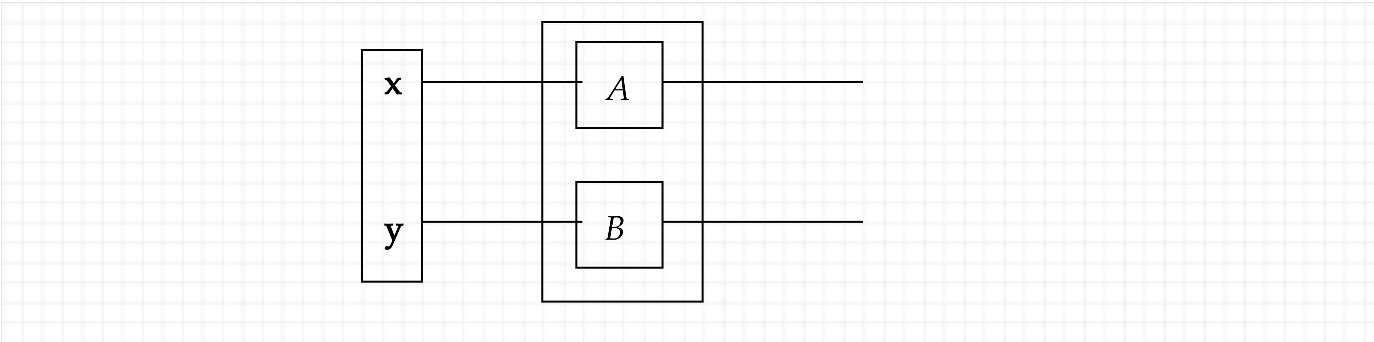
$$\sum_w \sum_t A(u,w)B(v,t)\mathbf{x}(w)\mathbf{y}(t)$$

With $\mathbf{z} \;=\; \mathbf{x} \otimes \mathbf{y}$, we can already recognize that the $\mathbf{x}(w)\mathbf{y}(t)$ part is the same as $\mathbf{z}(wt)$. And $A(u,w)B(v,t)$ is the same as $(A \otimes B)(uv, wt)$. So the whole thing becomes

$$\sum_{w,t} (A \otimes B)(uv, wt) \cdot (\mathbf{x} \otimes \mathbf{y})(wt)\,,$$

which is exactly the meaning of $(A \otimes B) \cdot (\mathbf{x} \otimes \mathbf{y})$. So the two sides are equal. ☒

The simple idea is that $(A \otimes B) \cdot (\mathbf{x} \otimes \mathbf{y})$ does the $A$ operation on $x$ side-by-side with $B$ doing its operation on $y$, but with no connection at all between them. We will soon have diagrams like this---




---note that we picture the inputs coming in from the left but when writing them as matrix arguments they will swing around to the right. As a tandem, this is formally the tensor product $\mathbf{x} \otimes \mathbf{y}$ coming in to $(A \otimes B)$. But really---and **locally**---it is just $A\mathbf{x}$ happening in one place and $B\mathbf{y}$ happening independently in another place. The upshot is this:

> **When we have entanglement, not independence, between the x part and the y part, then the notation will stay the same but the interpretation will change a whole lot.**

[Notation note: The boldfacing on vectors $\mathbf{x}$ and $\mathbf{y}$ is to distinguish them when strings $x$ and $y$ are nearby, and also to convey that they may represent specific physical quantities. The bolding of matrices has the latter idea---in particular, quantum operators like $\mathbf{H}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}$ are bolded. The textbook uses a smoother bolding that I don't see how to get in MathCha.]

### Reversal, Adjoint, and Duality.

The reversal $x^R$ of a string $x$ just means writing it "backwards": $01001^R = 10010$, FACED$^R$ = DECAF, and so on. A string $x$ is a palindrome if $x^R = x$, for instance $1001$. The empty string $\epsilon$ counts as a palindrome since $\epsilon^R = \epsilon$. The rule for reversal and concatenation is that for any strings $x$ and $y$,

$$(xy)^R = y^R x^R.$$

For example,

$$(\mathrm{PUCK}-\mathrm{FACED})^R = (\mathrm{FACED})^R(\mathrm{PUCK}-)^R = \mathrm{DECAF}-\mathrm{KCUP}\,.$$

Actually, if the minus sign is a $-1$ factor which could go anywhere, this would be equivalent to say "DECAF K-CUP" meaning a certain pod for a Keurig coffee-maker.

This gives intuition for how matrix transpose, matrix adjoint, and matrix inverse all work like reversal with regard to matrix product. The rules for any (invertible) matrices $A$ and $B$ are:

1. $(AB)^T = B^T A^T$

2. $(AB)^* = B^* A^*$
3. $(AB)^{-1} = B^{-1} A^{-1}$.

Rule 2 follows from rule 1 because the only difference with $^*$ is doing complex conjugates of individual entries. Rule 3 follows since $(AB)\left(B^{-1}A^{-1}\right) = ABB^{-1}A^{-1} = AA^{-1} = \mathbf{I}$. So why does rule 1 hold? Here our functional view might help: The transpose $A^T$ is the function with the two index arguments reversed: $A^T(j,i) = A(i,j)$. So:

$$(AB)^T(i,j) = (AB)(j,i) = \sum_k A(j,k)B(k,i) = \sum_k B(k,i)A(j,k) = \sum_k B^T(i,k)A^T(k,j) = B^T A^T(i,j)$$

for all arguments (i.e., indices) $i$ and $j$, so $(AB)^T = B^T A^T$. (Note that the switch $A(j,k)B(k,i) = B(k,i)A(j,k)$ in the middle step was just ordinary multiplication of numbers.)

The ideas of transpose and adjoint work also for vectors. The transpose of a column vector is a row-vector. Likewise, the adjoint $\mathbf{x}^*$ of a column vector $\mathbf{x}$ is a row vector. When we multiply a row vector and a column vector---in that order---we get a single number, i.e., a **scalar**. In particular,

$$\mathbf{x}^*\mathbf{x} = \sum_i \mathbf{x}^*(i)\mathbf{x}(i) = \sum_i \overline{\mathbf{x}[i]}\mathbf{x}[i] = \sum_i |\mathbf{x}[i]|^2 = ||\mathbf{x}||^2,$$

which is just the square of the Euclidean length of the vector $\mathbf{x}$. Now if you buy in to the reversal rule for adjoints, we can give a short and snappy proof of Lemma 3.1 in the text.

**Lemma 3.1**: If $\mathbf{U}$ is a unitary matrix and $\mathbf{a}$ is a vector then $||\mathbf{Ua}|| = ||\mathbf{a}||$.

**Proof**: $||\mathbf{Ua}|| = \sqrt{||\mathbf{Ua}||^2} = \sqrt{(\mathbf{Ua})^*(\mathbf{Ua})} = \sqrt{\left(\mathbf{a}^*\mathbf{U}^*\right)(\mathbf{Ua})} = \sqrt{\mathbf{a}^*\left(\mathbf{U}^*\mathbf{U}\right)\mathbf{a}} = \sqrt{\mathbf{a}^*\mathbf{a}} = ||\mathbf{a}||$.☒

The proof became a one-liner. Thus a unitary matrix always preseves the lengths of vectors, and in particular, it always maps a unit vector to a unit vector. This is what makes it "legal" from the quantum probability point of view. The fact works the other way: if a matrix $U$ always preserves the lengths of vectors, then it must be unitary.

The adjoint $\mathbf{x}^*$ of a vector $\mathbf{x}$ has another interpretation. It stands ready to pounce on any column vector $\mathbf{y}$ of the same length as $\mathbf{x}$ and wrangle it down to the scalar

$$\mathbf{x}^*\mathbf{y} = \sum_i \overline{\mathbf{x}[i]}\mathbf{y}[i] = \langle \mathbf{x}, \mathbf{y} \rangle,$$

which is the inner product of $\mathbf{x}$ and $\mathbf{y}$. As such, $\mathbf{x}^*$ defines the **linear functional** $f_{\mathbf{x}} : \mathbb{H}^n \to \mathbb{H}$ by

$$f_{\mathbf{x}}(\mathbf{y}) = \langle \mathbf{x}, \mathbf{y} \rangle.$$

Whereas a column vector is to be interpreted as "data", the row-vector form is "code". The resulting inner product finally suggested---to the physicist Paul Adrien Maurice Dirac in particular---to write the adjoint of $\mathbf{x}$ as $\langle x|$ instead, to go with writing $|y\rangle$ in place of $\mathbf{y}$. Some nerdy things to note:

- There is no $*$ or complex-conjugation $\overline{x}$ in $\langle x|$. The complex inner product $\langle x|y\rangle$ (if we write it that way) already does the conjugation.
- Put another way, the adjoint $|x\rangle^*$ of $|x\rangle$ is exactly what $\langle x|$ is---no further $^*$ required.
- If the vector $\mathbf{x}$ has no complex entries then $\langle \mathbf{x}, \mathbf{y} \rangle$ is the same as the ordinary real dot product $\mathbf{x} \cdot \mathbf{y} = \sum_i \mathbf{x}[i]\mathbf{y}[i]$ anyway.
- Hey, did you forget to write the bold for vectors? Why $\langle x|$ and $|y\rangle$ not $\langle \mathbf{x}|$ and $|\mathbf{y}\rangle$? The answer is that the angle brackets already identify the contents as physically meaningful vectors. Not only to they distinguish $\langle x|$ and $|y\rangle$ from strings $x$ and $y$, we want to write $\langle x|$ and $|y\rangle$ precisely when $x$ and $y$ **are** strings. Such as when writing $|10010\rangle$, for instance.
- There is nothing wrong with writing $\langle \mathbf{x}|$ and $|\mathbf{y}\rangle$, in our opinion---it just might be redundant. Where this matters is in Chapter 14 where we follow the common usage of the Greek letters $\phi, \psi$ etc. to represent quantum states. Then writing $|\phi\rangle$, $|\psi\rangle$, etc., makes them look "more quantum" but usually does not have any further significance.
- If $\mathbf{z} = a\mathbf{x}$ where $\mathbf{x}$ and $\mathbf{z}$ are numeric vectors and $a$ is a (possibly complex) scalar, then we have the rule $\mathbf{z}^* = \overline{a}\mathbf{x}^*$. We have to remember to conjugate any factor we pull out of the adjoint. About a minute into this Khan Faculty video they write the rules $|a\psi\rangle = a|\psi\rangle$ and $\langle a\psi| = a^*\langle\psi|$, but you have to be careful that $\psi$ stands for a numeric vector here. It makes no sense to say e.g. that $3|1\rangle = |3\rangle$ when the $|1\rangle$ is the binary-bit attribute, nor that $3|7\rangle = |21\rangle$ if the "7" is the rank of a playing card. (Note that it is more convenient to write $a^*$ rather than $\overline{a}$ for the complex conjugate of a scalar, as if it were a "$1 \times 1$" dimensioned entity. We will do so on occasion.)

The $\langle \cdot |$ form is called a **bra** to go with $| \cdot \rangle$ being a **ket** just so that the combination $\langle \cdot | \cdot \rangle$ becomes a **bra**c**ket**. The genius of the notation is liberating the inner product into a product with interchangeable parts. The bras and kets can be combined, with these resulting rules:

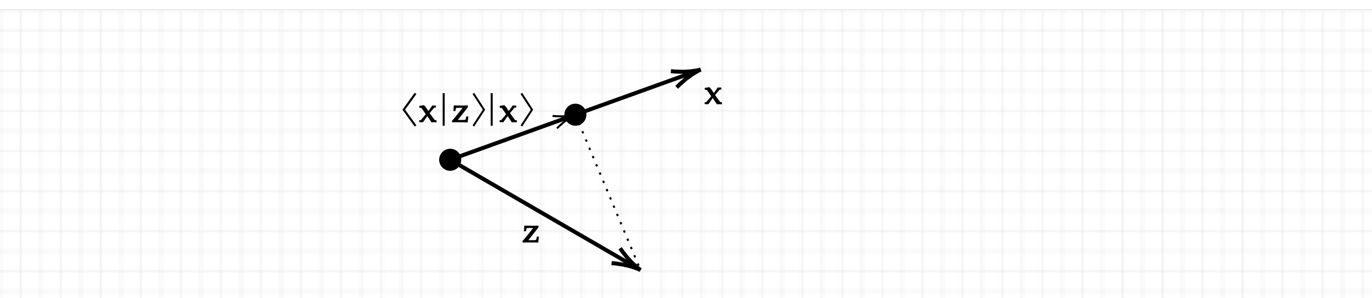
1. $\langle x| \cdot |y\rangle = \langle x|y\rangle$. The product dot first goes invisible, then the two vertical bars combine to be one.
2. $\langle y|x\rangle = \langle y| \cdot |x\rangle = |y\rangle^* \cdot \langle x|^* = (\langle x| \cdot |y\rangle)^* = \langle x|y\rangle^*$ by the reversal rule. So the flipped-around inner product $\langle y|x\rangle$ is just the complex conjugate of the scalar $\langle x|y\rangle$.
3. Two consecutive kets as in $|x\rangle|y\rangle$ is a gray area. It is tempting to equate it to $|x\rangle \otimes |y\rangle$ so that we could have cases like $|1\rangle|0\rangle|0\rangle|1\rangle|0\rangle = |10010\rangle$. But the product of two column vectors is not really defined. If you have something like $\langle w|$ before your $|x\rangle|y\rangle$, then you want it to become $\langle w|x\rangle \cdot |y\rangle$, where the $\cdot$ is ordinary multiplication.

4. Two consecutive bras like $\langle x|\langle y|$ are even grayer. Would they be the adjoint of $|y\rangle|x\rangle$ or of $|x\rangle|y\rangle$? Note what happens for tensor products of matrices: For all indices $u, v, w, t$,

$$\begin{aligned}(A \otimes B)^*(uv, wt) &= \overline{(A \otimes B)(wt, uv)} = \overline{A(w,u)B(t,v)} = \overline{A(w,u)} \cdot \overline{B(t,v)} \\ &= A^*(u,w)B^*(v,t) = \left(A^* \otimes B^*\right)(uv, wt).\end{aligned}$$

So $(A \otimes B)^* = A^* \otimes B^*$. Did you expect the $A$ and $B$ to reverse? Maybe not if you realize that they operate in independent systems.

5. $|x\rangle\langle y|$ --- ? The product of a $p \times 1$ column vector $\mathbf{x}$ and a $1 \times q$ row vector $\mathbf{y}$ is well defined algebraically. It gives a $p \times q$ matrix $A$ of entries $A[i,j] = \mathbf{x}[i]\mathbf{y}[j]$. If $\mathbf{y}$ is given as a numeric vector inside a bra then we have to remember to conjugate its entries, so that $A[i,j] = \mathbf{x}[i]\overline{\mathbf{y}[j]}$. The resulting matrix $A$ has **rank** one---so it is as far from being invertible as possible without being the zero matrix. It is called the **outer product** and has the following important relation to inner product when given any column vector $|z\rangle$: It pounces on $|z\rangle$, wrangles it into the scalar $a = \langle y|z\rangle$, and multiplies $|x\rangle$ by that.
6. In particular, the outer product $|x\rangle\langle x|$ of a vector $|x\rangle$ with itself becomes an operator that makes any vector $|z\rangle$ multiply $|x\rangle$ by the extent to which $|z\rangle$ itself aligns with $|x\rangle$. This gives the **projection** of $|z\rangle$ **onto** $|x\rangle$. One rule of projections is that repeating it doesn't change the result, at least not when $|x\rangle$ is a unit vector: $|x\rangle\langle x|$ applied to $\left(|x\rangle\langle x|\right)|z\rangle$ gives $|x\rangle\langle x|\left(|x\rangle\langle x|\right)|z\rangle = |x\rangle\langle x|x\rangle\langle x|z\rangle = |x\rangle\langle x|z\rangle = \langle x|z\rangle|x\rangle$ since $\langle x|z\rangle$ is a scalar.




The issues with the possible rules 3 and 4 still make us suspicious of Dirac notation and require being careful with $\langle x|z\rangle|x\rangle$ here. Can we read it as the single-tier bra $\langle x|$ multiplying the double-tier quantity $|z\rangle|x\rangle$ read as $|z\rangle \otimes |x\rangle$? Then the dimensions don't even align for multiplying on the left by the row vector $\langle x|$. The issue is that the "invisible dot" between the $|z\rangle$ and the $|x\rangle$ is a scalar product in $\langle x|z\rangle|x\rangle$, but gets morphed into a tensor product in $|z\rangle \otimes |x\rangle$. In online forums one can find it explained that the tensor way of interpreting $|z\rangle|x\rangle$ doesn't stay within the *algebra* of the "single-tier" vectors.

But regardless, the identity $\left(|x\rangle\langle x|\right) \cdot |z\rangle = |x\rangle$ multiplied by $\langle x|z\rangle$ is real. Indeed, there is a strong argument for saying that all reality goes through it: it is the basis of defining the **density matrix** of a quantum state as will come later in chapter 14.