| **6.443/8.371/18.436 Quantum Information Science II** | Lecture 1 |
|---|---|

# Quantum states and operations

*Aram Harrow* — February 7, 2018

## 1.1 Administrivia

For course information, go to http://web.mit.edu/8.371. There you will find a syllabus with information about office hours, grading, a tentative outline, etc.

## 1.2 Basics: States, operations, metrics

You may have heard that quantum mechanics is a generalization of probability theory [1]. There's more to it than that. Let's look at states, both classical and quantum. Furthermore, we can classify these states as deterministic or random.

| **States** | Deterministic | Random |
|---|---|---|
| Classical | $[d] = \{1, 2, ..., d\}$ | $p \in \mathbb{R}^d \geq 0$,<br>$\sum_x p_x = 1$. |
| Quantum | $\|\psi\rangle \in \mathbb{C}^d$,<br>$\langle\psi\|\psi\rangle = 1$ | Density matrix $\rho \in L(\mathbb{C}^d)$<br>$\rho \succeq 0, \operatorname{tr}\rho = 1, \rho = \rho^\dagger$ |

In the upper right we have classical probability theory, and in the lower left we have pure-state quantum mechanics. Both can be viewed as generalizations of classical deterministic state spaces. They are analogous but incomparable. The lower-left corner is a common generalization of both. Since density matrices were discussed in 8.370 we will not review them in great detail here. For more information the course website will link to the 8.370 videos and some lecture notes from 8.06 that discuss density matrices in detail.

To understand where density matrices come from, let's introduce ensembles of quantum states. An ensemble is a collection like $\mathcal{E} = \{(p_1, |\psi_1\rangle), \ldots, (p_n, |\psi_n\rangle)\}$ such that each $|\psi_i\rangle \in \mathbb{C}^d$, $p_i \geq 0$ and $\sum_i p_i = 1$. In other words, $\mathcal{E}$ refers to the event that the state of a quantum system is $|\psi_1\rangle$ with probability $p_1$, or $|\psi_2\rangle$ with probability $p_2$ and so on.

However, there are more degrees of freedom in an ensemble than we need. Due to the linearity of quantum mechanics, observables can only depend on quantum state in a limited way. Indeed, the expected value of an arbitrary observable $M$ according to the ensemble $\mathcal{E}$ is $\sum_i p_i \langle\psi_i| M |\psi_i\rangle = \operatorname{tr}(M\rho_{\mathcal{E}})$, where $\rho_{\mathcal{E}} = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. Hence, $\rho_{\mathcal{E}}$ captures all the essential content we need about $\mathcal{E}$. Even if $n$ is much larger then $d$, $\rho_{\mathcal{E}}$ represents $\mathcal{E}$ using only $d^2$ complex numbers.

Density matrices are in a way generalizations of ensembles. Consider the singular value decomposition of a density matrix $\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$. $\rho \succeq 0$ means $\lambda_i \geq 0$, and $\operatorname{tr}(\rho)$ means $\sum_i \lambda_i = 1$. Hence the eigenvalue vector $\lambda = (\lambda_1, \ldots, \lambda_d)$ of $\rho$ is a probability distribution, and $\{(p_1, |\psi_1\rangle), \ldots, (p_d, |\psi_d\rangle)\}$ is an ensemble with the same density matrix as $\rho$. Consider the extreme cases:

1. $\lambda = (1, 0, 0, ..., 0) \implies \rho$ is a pure state $= |\psi\rangle\langle\psi|$
2. $\lambda = (1/d, 1/d, 1/d, ..., 1/d) \implies \rho = \mathbb{I}/d$. $\rho$ is maximally mixed, or maximally "noisy."

Let $|1\rangle, \ldots, |d\rangle$ be an orthonormal basis for $\mathbb{C}^d$. Density matrices also generalize, deterministic classical states by $j \leftrightarrow |j\rangle\langle j|$, and also classical random states by $(p_1, \ldots, p_d) \leftrightarrow \sum_i p_i |i\rangle\langle i|$. Finally they generalize pure quantum states: the pure state $|\psi\rangle$ becomes the density matrix $|\psi\rangle\langle\psi|$. Note that this representation has the advantage that replacing $|\psi\rangle$ with $e^{i\phi}|\psi\rangle$ does not change the density matrix.

[1] From Scott Aaronson's *Quantum Computing Since Democritus:* "Quantum mechanics is a beautiful generalization of the laws of probability: a generalization based on the 2-norm rather than the 1-norm, and on complex numbers rather than nonnegative real numbers...this generalized probability theory leads naturally to a new model of computation – the quantum computing model..."

Aside: It is assumed that when we say positive semidefinite, we also mean Hermitian. $A = A^\dagger$ is positive semidefinite (denoted $A \succeq 0$) if:

1. eigenvalues(A) $= (\lambda_1, ..., \lambda_d)$ are all $\geq 0$
2. $\langle v|A|v\rangle \geq 0 \; \forall \, |v\rangle$
3. $A = B^\dagger B$ for some $B$

*Exercise:* Prove that all three conditions are equivalent.

### 1.2.1 Bloch ball

For the $d = 2$ case the set of density matrices has a simple geometric description. The Pauli matrices $\sigma_i$ and the identity constitute a basis for the space of $2 \times 2$ Hermitian matrices, so we can write

$$\rho = \frac{a_0\mathbb{I} + a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3}{2} \tag{1.1}$$

for arbitrary $a_i \in \mathbb{R}$. Recalling that the trace of a density matrix is 1 and the Pauli matrices are traceless, we conclude that

$$\operatorname{tr}\rho = a_0 = 1 \implies \rho = \frac{\mathbb{I} + \sum_{i=1}^{3} a_i\sigma_i}{2}. \tag{1.2}$$

If you have taken 8.05 and 8.06 you may recall that $\text{eigs}(\sum_{i=1}^{3} a_i\sigma_i) = \pm\sqrt{\sum_i a_i^2}$. If you have not seen this before, pause to justify this to yourself. Combined with the condition that $\rho$ has nonnegative eigenvalues, we have

$$\text{eigs}(\rho) = \frac{1 \pm \sqrt{\sum_i a_i^2}}{2} \geq 0 \implies \|\vec{a}\|^2 \leq 1. \tag{1.3}$$

What this means is that valid $2 \times 2$ density matrices can be represented by a unit ball centered at the origin. This is called the “Bloch ball.” Notably, pure states are on the surface of the ball, called the “Bloch sphere.” That is, $\rho$ is pure iff $\text{eigs}(\rho) = (1, 0)$ iff $\|\vec{a}\| = 1$.

For example, the point (0, 0, 1) corresponds to the state $|0\rangle$, as $\rho = \frac{\mathbb{I}+\sigma_z}{2} = |0\rangle\langle 0|$. Similarly, (0, 0, -1) corresponds to $|1\rangle$. The states $\frac{|0\rangle \pm |1\rangle}{\sqrt{2}}$ on the Bloch sphere lie along the x-axis, and $\frac{|0\rangle \pm i|1\rangle}{\sqrt{2}}$ are along the y-axis (in and out of the page, if you draw them). By contrast, points on the interior of the sphere are mixed states, with $\vec{a} = 0$ corresponding to the maximally mixed state $\rho = \mathbb{I}/2$.

Density matrices generalize both quantum deterministic and classical random states. How do classical probabilities fit into this geometric picture? Suppose $p = \mathbb{P}(\text{bit} = 0)$ and $1 - p = \mathbb{P}(\text{bit} = 1)$. Then the density matrix is

$$\rho = \begin{bmatrix} p & 0 \\ 0 & 1-p \end{bmatrix} = \frac{\mathbb{I} + (2p-1)\,\sigma_z}{2}.$$

Geometrically,such states lie inside the Bloch ball, along the vertical line between the points (0, 0, 1) and (0, 0, -1).

### 1.2.2 Operations

Let’s consider a similar table to that we started with, but this time for operations rather than states.

| States | Deterministic | Random |
|---|---|---|
| Classical | $f : [d] \to [d]$<br>$f$ could destroy information | stochastic $T \in \mathbb{R}^{d\times d}$ |
| Quantum | U(d) | Quantum operations, channels, TPCP maps |

Let's focus on the top right box for the moment. If you are in a state x, you have probabilities to transition to states $1, ..., d$. These transitions are codified by a matrix whose elements are $T_{yx} = \mathbb{P}[y|x]$. The initial distribution $p$ is mapped by this process to $Tp$. Such an operation is called "stochastic." We must have that $T_{yx} \geq 0$ and $\sum_y T_{yx} = 1$, as the total probability must remain 1.

*Exercise:* Find T that is not positive semidefinite.

What are the ways in which T is operationally unlike a unitary matrix?

1. Stochastic matrices can create or destroy information. (Formally they can increase or decrease entropy.)
2. Unlike with unitary matrices, there may not exist a generator G such that $T = e^G$ and $e^{Gt}$ is stochastic for all real $t$.
3. The set of stochastic matrices is convex, i.e. $pT^{(1)} + (1-p)T^{(2)}$ is also stochastic. The set of unitaries is not convex.

As with states, the bottom right box must provide the common generalizations of stochastic operations and unitary operations. Here are some examples of quantum operations $\mathcal{N}$:

1. $\mathcal{N}(\rho) = U\rho U^\dagger$ for $U \in U(d)$, unitary evolution.
2. $\mathcal{N}(\rho_{AB}) = \mathrm{tr}_B(\rho_{AB})$, taking the partial trace.
3. $\mathcal{N}(\rho) = \rho \otimes \sigma$, adjoining an ancilla system.
4. $\mathcal{N}(\rho) = V\rho V^\dagger$, applying an isometry to $\rho$.
5. $\mathcal{N}(\rho) = \sigma$, replacing $\rho$ with some fixed state $\sigma$ independent of the input.

Unitary and isometric operations are *not* the same! An isometry is $V : \mathbb{C}^{d_1} \to \mathbb{C}^{d_2}$ such that $\|V\,|\psi\rangle\,\| = \|\,|\psi\rangle\,\|$. Note that this implies that $d_2 \geq d_1$ and $V^\dagger V = \mathbb{I}_{d_1}$. An example of an isometry is $V\,|\psi\rangle = |\psi\rangle \otimes |0\rangle$ . This is not unitary, as unitary operations preserve dimension.

How can operations outside of the bottom left box, namely deterministic quantum operations, occur in the world? Recall that subsystems of entangled pure states are mixed. That means that a subsystem of a larger state can look random. This is how we get noisy quantum operations. More on this next time...