Tutorial on Quantum Mechanics

Prakash Panangaden
Ecole d'Informatique - School of Computer Science
Université McGill University
Montréal, Québec

Thanks to Samson Abramsky, Richard Blute, Bob Coecke, Ellie D'Hondt, Ivan T. Ivanov, Peter Selinger, Rafael Sorkin and especially Gordon Plotkin for interesting discussions.

Terminology of Linear Spaces and Operators

- vector space, inner product, Hilbert space
- tensor product
- linear operator
- adjoint: $\langle \psi, A\phi \rangle = \langle A^\dagger \psi, \phi \rangle$
- hermitian operator: $A = A^\dagger$
- unitary operator: $U^{-1} = U^\dagger$
- projection: $P^2 = P$ and $P = P^\dagger$

Physicists' Notation for Linear Algebra

- element of vector (Hilbert) space $|\text{label}\rangle$
- element of dual space $\langle\text{label}|$
- inner product $\langle a|b\rangle$
- element of $\mathcal{H}_1 \otimes \mathcal{H}_2$, $|ab\rangle$
- element of the space $\mathcal{H} \otimes \mathcal{H}^*$ (a matrix) $|b\rangle\langle a|$.

$$|a\rangle\langle b|(|c\rangle) = \langle b|c\rangle|a\rangle.$$

- projection operator onto subspace spanned by $|a\rangle$, $|a\rangle\langle a|$.

Postulates of Quantum Mechanics

- States form a Hilbert Space $\mathcal{H}$,

- The evolution of an *isolated* system is governed by a *unitary* transformation on $\mathcal{H}$: $U(t, t') = \text{exp}(-i\text{H}(t' - t))$,

- Measurements are described by operators acting on $\mathcal{H}$.

Measurements I

The usual case: The quantity being measured is described by a hermitian operator M.

The possible outcomes are the *eigenvalues* of M.
If M is an observable (hermitian operator) with eigenvalues λ_i and eigenvectors ϕ_i and $\psi = \sum_i c_i \phi_i$ then probabilities and expectation values are given by:

- $Prob(\lambda_i|\psi) = |c_i|^2$

- $E[M|\psi] = \sum_i |c_i|^2 \lambda_i = \sum_i c_i \bar{c}_i \langle \phi_i, M\phi_i \rangle$
  $= \langle \psi, M\psi \rangle.$

These are so-called *projective* measurements and are a special case.

Interpreting Quantum Mechanics (Naive)

What happens during a measurement?

- The result of a measurement is a random event;
- the values obtained belong to the spectrum of the observable;
- if the measurement is *immediately* repeated then the same result is observed.

The process of measurement knocks the system into an eigenvector - *reduction of the state vector*.

What happened to the usual dynamics based on unitary evolution?

The “Theory” of Measurement

- Measurement is an **interaction** between system and apparatus.

- Measurements do not uncover some pre-existing physical property of a system. There is no **objective property** being measured.

- The **record** or **result** of a measurement is an objective property.

- Quantum mechanics is nothing more than a set of rules to compute the outcome of physical tests to which a system may be subjected.

The Stern-Gerlach Experiment

- Send neutral atoms through a varying magnetic field.

- Observe two peaks - the beam is split into an "up" and a "down."

- Rotate the apparatus and still observe the beam split in two.

Suppose that the atom has a magnetic moment $\vec{\mu}$ (a vector) then the observed component of the moment about a unit vector $\mathbf{e}$ is $\vec{\mu} \cdot \mathbf{e}$ which - according to the experiment - must be $\pm\mu$. If we choose three vectors $\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}$ at 120 degrees to each other then we have $\mathbf{e_1} + \mathbf{e_2} + \mathbf{e_3} = \mathbf{0}$. Hence if $\mu_i := \vec{\mu} \cdot \mathbf{e_i}$ we have $\mu_1 + \mu_2 + \mu_3 = 0$ which is impossible if each μ_i has value $\pm\mu$.

Composing Subsystems

A fundamental difference between classical and quantum systems is:

we put systems together by **tensor product** in quantum mechanics rather than by cartesian product.

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$$

This is the key to the power of quantum computing: see various papers by Josza at

`www.cs.bris.ac.uk/Research/QuantumComputing/entanglement.html`

Digression: Spectral Theorem for Undergraduates

Theorem: If M is hermitian then its eigenvalues are all real and it can be written in the form $\sum_i \lambda_i P_i$ where λ_i are the eigenvalues of M and P_i are the projection operators on to the corresponding eigenspaces and the P_i are all mutually orthogonal.

Proof of the Spectral Theorem

Proof. The facts that the eigenvalues are all real and that the P_i are orthogonal are easy to prove. For the main claim we proceed by induction on the dimension d. The case $d = 1$ is trivial. Let λ be an eigenvalue of M, P the projector onto its eigenspace and Q the orthogonal projector to P. We have $P + Q = I$ and $PQ = 0$. Now clearly $PMP = \lambda P$.

$$\begin{aligned} M &= (P+Q)M(P+Q) \\ &= PMP + QMQ + QMP + PMQ. \end{aligned}$$

Now $QMP = 0$ and $(PMQ)^\dagger = QMP = 0$ so $PMQ = 0$. Also QMQ is hermitian so by inductive hypothesis it has the required form and from

$$M = \lambda P + QMQ$$

we conclude.

Density Matrices (Operators)

An alternative description of states and of the postulates of quantum mechanics due to von Neumann.

- Given $\psi \in \mathcal{H}$ we have a *projection operator* P_ψ or $|\psi\rangle\langle\psi|$. When a system is in the state ψ we say that $|\psi\rangle\langle\psi|$ is the **density matrix** of the system.

- A unitary operator U acts on $|\psi\rangle$ by $U|\psi\rangle$ and hence it acts on $|\psi\rangle\langle\psi|$ by $U|\psi\rangle\langle\psi|U^\dagger$.

- If M is an observable and P_ψ is the density matrix
  - $Pr(\lambda_i|P_\psi) = Tr(P_i P_\psi)$
  - $E[M|P_\psi] = Tr(MP_\psi)$.

Why bother with Density Matrices I?

- Density matrices capture the notion of *statistical mixture*.

- Statistical mixture of states from $\mathcal{H}$

$$\rho = \sum_{j=1}^{m} p_j |\psi_j\rangle\langle\psi_j|$$

where $\sum_j p_j = 1$.

- If all but one p_j are zero then we have a *pure* state otherwise a *mixed* state.

- Thus a density matrix can be seen as a **convex combination** of projection operators corresponding to pure states.

Why bother with Density Matrices II?

Density matrices capture *partial information* about a system. We can get partiality because we are seeing a subsystem of a larger system.

The state space decomposes as $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. If we have a pure state ψ or P_ψ for the entire system but an observer only sees the part described by $\mathcal{H}_1$ we get a density matrix ρ by

$$\rho = Tr_{(2)}(P_\psi).$$

This is almost always a mixed state.

Here $Tr_{(2)}(\cdot)$ means partial trace over $\mathcal{H}_2$.

Recognizing Density Matrices

Proposition - Alternate Definition An operator ρ on $\mathcal{H}$ is a **density matrix** if and only if

- ρ has trace 1 and
- ρ is a positive operator.

Let λ_i be the eigenvalues of ρ. Then $\forall i . \lambda_i \geq 0$ since ρ is a positive operator. Also

$$Tr(\rho) = \sum_i \lambda_i = 1$$

hence

$$Tr(\rho^2) = \sum_i \lambda_i^2 \leq 1$$

with equality if and only if all but one of the λ_i are zero.

Recognizing Pure States

By the spectral theorem $\rho = \sum_i \lambda_i P_i$, where P_i are the projection operators onto the eigenspaces. It follows that if ρ is a density matrix, $Tr(\rho^2) \leq 1$ with equality if and only if ρ is a pure state (i.e. a projection operator).

Density matrices as probability distributions

Note that a positive operator has non negative eigenvalues and trace 1 means that they add up to 1. So by the spectral theorem such an operator is a convex combination of projections operators. So

density matrices are “just” probability distributions on pure states?

No!

Given a density matrix for a mixed state one cannot recover a unique decomposition into pure states.

Mixtures vs Superpositions

A photon can be *polarized* so the state space is a 2-dim Hilbert space. Consider the basis:

- $|\leftrightarrow\rangle$: polarized along the x-axis
- $|\updownarrow\rangle$: polarized along the y-axis

A $|\leftrightarrow\rangle$ photon will not get through a $\updownarrow$ filter. Such a photon will get through a filter oriented at an angle of θ to the vertical with probability proportional to $\sin^2\theta$.

Mixtures vs Superpositions II

A photon can be in a superposed state, i.e. a linear combination of basis states. For example

$$|\nearrow\rangle = \frac{1}{\sqrt{2}}[|\leftrightarrow\rangle + |\updownarrow\rangle].$$

The density matrix for this pure state is

$$\frac{1}{2}[|\leftrightarrow\rangle\langle\leftrightarrow| + |\leftrightarrow\rangle\langle\updownarrow|$$

$$+|\updownarrow\rangle\langle\leftrightarrow| + |\updownarrow\rangle\langle\updownarrow|]$$

Mixtures vs Superpositions III

But we can also make a mixture:

$$\rho = \frac{1}{2}[| \leftrightarrow\rangle\langle\leftrightarrow | + | \updownarrow\rangle\langle\updownarrow |].$$

The density matrix ρ describes a mixed state while $| \nearrow\rangle$ describes a pure state. Easy to see that $Tr(\rho^2) = \frac{1}{2} < 1$.

What is the observable difference?

With a polarizer aligned along the diagonal $\nearrow$:

- All $|\nearrow\rangle$ photons get through
- Only half the ρ photons get through.

How to make a mixed state I

Consider a two state system with the states $|0\rangle$ and $|1\rangle$. We write $|01\rangle$ for $|0\rangle \otimes |1\rangle$. Prepare the EPR state

$$\frac{1}{\sqrt{2}}[|00\rangle - |11\rangle]$$

or in terms of density matrices

$$\frac{1}{2}[|00\rangle\langle 00| - |11\rangle\langle 00| - |00\rangle\langle 11| + |11\rangle\langle 11|].$$

How to make a mixed state II

Now separate the two particles and allow experimenter access to the first particle only. We compute the partial trace over the second subsystem.

$Tr_2(\frac{1}{2}[|00\rangle\langle 00| - |11\rangle\langle 00| - |00\rangle\langle 11| + |11\rangle\langle 11|])$
which is

$$\rho = \frac{1}{2}[|0\rangle\langle 0| + |1\rangle\langle 1|].$$

Clearly ρ is not a pure state since $Tr(\rho^2) = \frac{1}{2}$

Non-locality and Entanglement

A key idea due to Einstein, Podolsky and Rosen (1935) which was intended as an attack on quantum mechanics. Turned out to be revolutionary and led to the notion of non-locality and entanglement.

A different manifestation of non locality was discovered by Aharanov and Bohm in 1959 in which electrons react to the presence of a magnetic field that is "far away." [Don't ask me now!]

EPR - Bohm's version

Two-state quantum particle: $|\uparrow\rangle$ for spin up and $|\downarrow\rangle$ for spin down.

Two-particle basis states written: $|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$.

Consider the state: $\frac{1}{\sqrt{2}}[|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle]$. This state can be *prepared in a laboratory*. Measuring the spin of one particle "makes" the other one have the opposite spin.

This is action at a distance or *non-locality*.

EPR - Consequences

Information is *nonlocal*, a quantum mechanical state is nonlocal. We can substitute entanglement for communication.

What EPR **cannot** be used for is superluminal communication.

Bell's Analysis

Assuming that (local realism)

- there are objective values waiting to be measured (realism)

- interactions are local.

There is an upper bound on the correlations between distant independent measurements; *no quantum mechanical assumptions* are needed to derive this inequality. The Bell inequality contradicts the predictions of quantum mechanics. Moreover, these inequalities are violated *experimentally*. Thus local realism does not hold in nature. Usually one keeps locality of interaction and jettisons realism. [Bohm]

Summary of Quantum States

- Pure quantum states can be superposed; linear structure

- States can be non local (entangled)

- States can be mixed

- Mixed states are described by density operators.

Evolution of States

Evolution of pure states is governed by *unitary* operators.

$$U^\dagger = U^{-1}$$

which implies

$$\forall \psi, \phi \langle U\psi, U\phi\rangle = \langle U^\dagger U\psi, \phi\rangle = \langle \psi, \phi\rangle.$$

Typically $U(t, t_0) = \exp -iH(t - t_0)$ but we will not worry about the detailed structure of evolution operators.

On a density matrix ρ we have

$$\rho \mapsto \rho' = U\rho U^\dagger.$$

Evolution of States

If M (hermitian) describes a projective measurement with outcomes (eigenvalues) λ_i so that $M = \sum_i \lambda_i P_i$ then

$$\rho \mapsto \sum_i \frac{1}{p_i} P_i \rho P_i$$

where p_i is the probability of observing λ_i.

If it is known that the outcome is λ_i then

$$\rho \mapsto \frac{P_i \rho P_i}{p_i}.$$

Evolution of States III

More generally, measurements are described by positive operator-valued measures - the usual projective measurements are a special case. Outcomes labelled by $\mu \in \{i, \ldots, N\}$, to every outcome we have an operator F_μ. The transformation of the density matrix is

$$\rho \mapsto \rho' = \frac{1}{p_\mu} F_\mu \rho F_\mu^\dagger.$$

Let $E_\mu := F_\mu^\dagger F_\mu$; these are positive operators. For a measurement they satisfy $\sum_\mu E_\mu = I$ and the probability of observing outcome μ is $Tr(E_\mu \rho) = p_\mu$ as we had claimed.

Evolution of States IV: Intervention Operators

More general interaction: part of the quantum system gets discarded during the measurement. The transformation of the density matrix is given by:

$$\rho'_\mu = \frac{1}{p_\mu} \sum_m A_{\mu m} \rho A^\dagger_{\mu m}$$

where μ labels the degrees of freedom observed, m labels the degrees of freedom discarded and each $A_{\mu m}$ now maps between two Hilbert spaces of (perhaps) different dimensionality.

How can we come up with the general form of intervention operators on physical and mathematical grounds?

This will occupy the rest of the lecture.

Positive Operators

$A : \mathcal{H} \longrightarrow \mathcal{H}$ is **positive** if $\forall x \in \mathcal{H}.\ \langle x, Ax\rangle \geq 0$. Implicit in this definition is the assumption that $\langle x, Ax\rangle$ is always real; hence, all eigenvalues of A are real and A is hermitian.

Positivity Abstractly

Any vector space V can be equipped with a notion of positivity. A subset C of V is called a **cone** if

- $x \in C$ implies that for any *positive* α, $\alpha x \in C$,
- $x, y \in C$ implies that $x + y \in C$ and
- x and $-x$ both in C means that $x = 0$

We can define $x \geq 0$ to mean $x \in C$ and $x \geq y$ to mean $x - y \in C$.

An **ordered** vector space is just a vector space equipped with a cone.

Proposition: The collection of positive operators in the vector space of linear operators forms a cone.

Positive Maps

Abstractly, $L : (V, \leq_V) \longrightarrow (W, \leq_W)$ is a **positive map** if

$$\forall v \in V.\ v \geq_V 0 \Rightarrow L(v) \geq_W 0.$$

Do not confuse "positive maps" and "positive operators."

If we are transforming states (density matrices) then the legitimate transformations obviously take density matrices to density matrices. They have to be *positive maps* considered as maps between the appropriate ordered vector spaces. The appropriate ordered vector spaces are the vector spaces of linear operators on $\mathcal{H}$ the Hilbert space of pure states.

The problem with positive maps

Unfortunately the tensor product of two positive maps is not positive in general. We really want this! If I can perform transformation T_1 on density matrix ρ_1 and transformation T_2 on density matrix ρ_2 then I should be able to regard $\rho_1 \otimes \rho_2$ as a composite system and carry out $T_1 \otimes T_2$ on this system. We certainly want this if, say, T_2 is the identity. But even when T_2 is the identity this may fail!!

Example of a "Bad" Positive Map

Take $\mathcal{H}$ to be the two-dimensional Hilbert space with basis vectors $|0\rangle$ and $|1\rangle$, take T_1 to be the map that transposes a density matrix over $\mathcal{H}$ and T_2 the identity. The transpose map is positive because it does not change any of the eigenvalues.

Consider the entangled state

$$\frac{1}{\sqrt{2}}[|00\rangle + |11\rangle]$$

with density matrix

$$\frac{1}{2}[|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|].$$

In matrix form this is:

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

If we apply $T_1 \otimes I$ to the above density matrix we flip the 0 and the 1 of the first basis pair to get

$$\frac{1}{2}[|00\rangle\langle 00| + |10\rangle\langle 01| + |01\rangle\langle 10| + |11\rangle\langle 11|].$$

In matrix form this is

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

which has a negative eigenvalue so it is not a positive operator.

Completely Positive Maps

- A**completely positive map** K is a positive map such that for every identity map $I_n : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ the tensor product $K \otimes I_n$ is positive.

- The tensor of completely positive maps is always a completely positive map.

A Convenient(?) Category for Quantum Computation

Finding the “right” category in which to develop quantum computation is a pressing problem. It has to be closed under the right things. The category of ordered vector spaces with completely positive maps as the morphsism is a nice monoidal category but it is too big. It has the great virtue of being a traced monoidal category.

However if we think about physical effects the morphisms must preserve trace or - if we do not care about normalization - at least be trace decreasing. However, this latter category is not traced.

The Kraus Representation Theorem

The important result is the Kraus representation theorem (due to M.-D. Choi)

The general form for a completely positive map $\mathcal{E} : \mathcal{B}(\mathcal{H}_1) \longrightarrow \mathcal{B}(\mathcal{H}_2)$ is

$$\mathcal{E}(\rho) = \sum_m A_m \rho A_m^\dagger$$

where the $A_m : \mathcal{H}_1 \longrightarrow \mathcal{H}_2$.

Here $\mathcal{B}(\mathcal{H})$ is the Banach space of bounded linear operators on $\mathcal{H}$.

In other words: *Completely positive maps* and *intervention operators* are the same thing.

A more abstract version of this theorem was proved by Stinespring in 1955.

The trace requirement

If, in addition, we require that the trace of $\mathcal{E}(\rho) \leq 1$ then the A_m will satisfy

$$\sum_m A_m^\dagger A_m \leq I.$$

Final Summary

- States are density matrices; i.e. positive operators with trace 1.

- Physical transformations are trace preserving (decreasing) completely positive maps.