Measuring Quantum Systems

An Introduction to Quantum Theory and the POVM Formalism

Justin Chun

MTH 619: Functional Analysis III

15 May 2025

What are POVMs?

- (**P**ositive **O**perator)–**V**alued **M**easures
- Generalization of *projection-valued measures*, themselves a generalization of ordinary (scalar-valued) measures.
- Each projector in a projection-valued measure corresponds to the measuring of a observable.
- Examples of observables include:
  - Position (commonly denoted X)
  - Momentum (commonly denoted P)
  - Total energy, i.e. the Hamiltonian (commonly denoted H)

Measure Theory Recap

Recall that:

- A σ**–algebra** on a set X is a subset $\Sigma \subseteq 2^X$ stable under complement and countable union.
- The algebra of **Borel sets** of a topological space (X, τ) is the σ-algebra generated by the topology τ.
- A **measure** is a function $\mu : \Sigma \to [0, \infty]$ such that $\mu(\varnothing) = 0$ and such that for $\{E_n\}_{n \geq 1} \subset \Sigma$ countable pairwise disjoint,

$$\mu\left(\bigcup_{n\geq 1} E_n\right) = \sum_{n \geq 1} \mu(E_n).$$

- Slogan:

probability theory is a "normalized measure theory."
quantum theory is a "unitary probability theory."

Dirac Notation

Let $\mathcal{H}$ be a complex separable Hilbert space.

- $|x\rangle$ denotes a **ket**, aka a vector in $\mathcal{H}$.
- $\langle y|$ denotes a **bra**, aka a linear functional in $\mathcal{H}^*$.[1]
- $\langle y|x\rangle$ denotes a **bra-ket**, aka $y(x)$, a scalar.
- Think of $\langle \cdot | \cdot \rangle$ as an inner product, and note that the conjugate linearity is in the first argument, not the second.
- $|x\rangle\langle y|$ denotes a **ket-bra**. Think of this as an outer product.
- In particular, $|x\rangle\langle x| = P_x$ is the orthogonal projector onto the subspace generated by x.
- Generalizing, $P_W = \sum_{i=1}^{k} |i\rangle\langle i|$ for a k-dimensional subspace W with orthonormal basis $\{|1\rangle, \ldots, |k\rangle\}$.

[1] The dual space $\mathcal{H}^*$ is frequently identified with $\mathcal{H}$ via the Riesz Representation Theorem.

Dirac Notation ctd.

- The dagger (Hermitian adjoint) of a bra is a ket, and the dagger of a ket is a bra. The dagger of a scalar is its complex conjugate.

$$\langle x|^\dagger = |x\rangle, \qquad |x\rangle^\dagger = \langle x|, \qquad \alpha^\dagger = \alpha^*.$$

- If Ω and Λ are operators on $\mathcal{H}$,
  then $[\Omega, \Lambda] = \Omega\Lambda - \Lambda\Omega$ denotes the commutator,
  and $[\Omega, \Lambda]_+ = \Omega\Lambda + \Lambda\Omega$ denotes the anticommutator.
- We denote the **expected value** of an operator in a given pure state $|\psi\rangle$ by

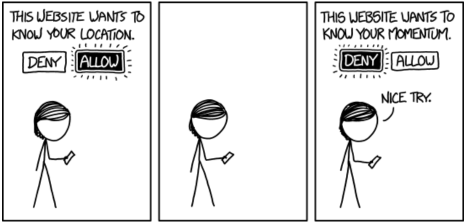
$$\langle\Omega\rangle = \langle\psi|\Omega|\psi\rangle$$

  and

$$\Delta\Omega = \left\langle \psi \left| (\Omega - \langle\Omega\rangle)^2 \right| \psi \right\rangle^{\frac{1}{2}}$$

  denotes the **uncertainty** (i.e., the standard deviation).

A Bit More on Position and Momentum



- In the position basis, we set: $X \mapsto x$, $P \mapsto -i\hbar\frac{\partial}{\partial x}$, $\langle x|x_0\rangle = \delta_{x_0}$.[2]
- In the momentum basis, we set: $X \mapsto i\hbar\frac{\partial}{\partial p}$, $P \mapsto p$, $\langle p|p_0\rangle = \delta_{p_0}$.
- Also, X and P are Fourier conjugates of each other:

$$\langle x|p_0\rangle = \exp\left(-\frac{p_0 x}{i\hbar}\right), \qquad \langle p|x_0\rangle = \exp\left(\frac{x_0 p}{i\hbar}\right)$$

Where $\langle x|p_0\rangle$ is some momentum eigenstate in position space, etc.

[2] Wait, isn't this a separable Hilbert space? Yes, but these basis vectors are actually a basis for a space much larger than $\mathcal{H}$, namely the space Φ^* of distributions on $\mathcal{H}$:

$$\Phi \subset \mathcal{H} \subset \Phi^* \qquad \text{example:} \qquad \mathcal{D} \subset \ell^2 \subset \mathcal{D}^*$$

Example: Heisenberg Uncertainty Principle

Let Ω and Λ be two Hermitian operators with commutator $[\Omega, \Lambda] = i\Gamma$. Then

$$\begin{aligned}(\Delta\Omega)^2(\Delta\Lambda)^2 &= \left\langle \psi \left| (\Omega - \langle\Omega\rangle)^2 \right| \psi \right\rangle \left\langle \psi \left| (\Lambda - \langle\Lambda\rangle)^2 \right| \psi \right\rangle \\ &= \left\langle \psi \left| \hat{\Omega}^2 \right| \psi \right\rangle \left\langle \psi \left| \hat{\Lambda}^2 \right| \psi \right\rangle\end{aligned}$$

where $\hat{\Omega} = \Omega - \langle\Omega\rangle$, $\hat{\Lambda} = \Lambda - \langle\Lambda\rangle$.

Now if $\Omega^\dagger = \Omega$ and $\Lambda^\dagger = \Lambda$, then the same is true for $\hat{\Omega}$ and $\hat{\Lambda}$.

Thus,

$$\begin{aligned}(\Delta\Omega)^2(\Delta\Lambda)^2 &= \left\langle \psi \left| \hat{\Omega}^\dagger \hat{\Omega} \right| \psi \right\rangle \left\langle \psi \left| \hat{\Lambda}^\dagger \hat{\Lambda} \right| \psi \right\rangle = \left\langle \hat{\Omega}\psi | \hat{\Omega}\psi \right\rangle \left\langle \hat{\Lambda}\psi | \hat{\Lambda}\psi \right\rangle \\ &= \left\|\hat{\Omega}\psi\right\|^2 \left\|\hat{\Lambda}\psi\right\|^2 \underset{C-S}{\geq} |\left\langle \hat{\Omega}\psi | \hat{\Lambda}\psi \right\rangle|^2 = |\left\langle \psi | \hat{\Omega}\hat{\Lambda} | \psi \right\rangle|^2\end{aligned}$$

Heisenberg Uncertainty ctd.

Now, $\hat{\Omega}\hat{\Lambda} = \frac{[\Omega,\Lambda]+[\Omega,\Lambda]_+}{2}$.

So, continuing, we have

$$
\begin{aligned}
|\left\langle\psi|\hat{\Omega}\hat{\Lambda}|\psi\right\rangle|^2 &= \left|\left\langle\psi\left|\frac{[\Omega,\Lambda]+[\Omega,\Lambda]_+}{2}\right|\psi\right\rangle\right|^2 \\
&= \left|\left\langle\psi\left|\frac{i\Gamma}{2}+\frac{[\Omega,\Lambda]_+}{2}\right|\psi\right\rangle\right|^2
\end{aligned}
$$

Since Ω and Λ are Hermitian, $[\Omega,\Lambda]_+$ and Γ are too:

$$
\begin{aligned}
(\Delta\Omega)^2(\Delta\Lambda)^2 &\geq \left|\frac{i}{2}\langle\psi|\Gamma|\psi\rangle + \frac{1}{2}\langle\psi|[\Omega,\Lambda]_+|\psi\rangle\right|^2 \\
&= \frac{|\langle\psi|\Gamma|\psi\rangle|^2}{4} + \frac{|\langle\psi|[\Omega,\Lambda]_+|\psi\rangle|^2}{4} \geq \frac{|\langle\psi|\Gamma|\psi\rangle|^2}{4}.
\end{aligned}
$$

Heisenberg Uncertainty ctd.

In particular, when $\Gamma = \hbar \mathbf{1}_{\mathcal{H}}$,[3]

$$\boxed{(\Delta\Omega)(\Delta\Lambda) \geq \frac{\hbar}{2}}$$

This is the **Heisenberg Uncertainty Principle**:

the product of two conjugate uncertainties is bounded below by some positive constant, so that it is never zero.

[3] As is the case with position and momentum.

Dynamics of Quantum Theory

The **Schrödinger Equation** is

$$i\hbar|\dot{\psi}\rangle = H|\psi\rangle$$

i.e. $i\hbar|\dot{\psi}\rangle - H|\psi\rangle = 0$. Since $0^\dagger = 0$, we have $\langle\dot{\psi}|i\hbar + \langle\psi|H = 0$ (since the Hamiltonian of any quantum system is Hermitian).

Bra-ing the first equation on the right, ket-ing the second on the left, and adding the results together yields

$$i\hbar(|\psi\rangle\langle\dot{\psi}| + |\dot{\psi}\rangle\langle\psi|) = [H, |\psi\rangle\langle\psi|]$$

or for $\rho = |\psi\rangle\langle\psi|$,

$$\boxed{i\hbar\dot{\rho} = [H, \rho]}$$

This is the **Liouville–von Neumann equation** and it describes how density operators (quantum states) evolve through time.

What is a density operator?

- Associated to every quantum system is a Hilbert space $\mathcal{H}$ of pure states $|\psi\rangle$.[4]
- Let $\mathcal{B}$ be the C^*-algebra of bounded operators on $\mathcal{H}$.[5]
- Within $\mathcal{B}$, the positive semidefinite operators form a cone.
- Intersect this cone with the hyperplane of operators A satisfying $\operatorname{tr} A = 1$: the result is the convex set $\mathcal{S}$ of density operators ρ, which represent *all* quantum states.
- So, a **density operator** is an operator $\rho \in \mathcal{S}$ satisfying

$$\langle\psi|\rho|\psi\rangle \geq 0, \qquad \operatorname{tr} \rho = 1$$

for all pure states $|\psi\rangle \in \mathcal{H}$.

[4] This is a slight inaccuracy, what we really want is a Gelfand triple.

[5] If we allow Hilbert spaces to be over an arbitrary C^*-algebra (instead of just $\mathbf{C}$), then we note that $\mathcal{B}$ forms a Hilbert space over itself (technical term: Hilbert C^*-module). The space $\mathcal{B}$ has inner product $\langle A|B\rangle_{\mathcal{B}} = \operatorname{tr}(A^\dagger B)$.

Purity of States

- Pure states are linear combinations of basis kets in the Hilbert space $\mathcal{H}$. For example,
$$|\psi\rangle = \frac{|0\rangle + |1\rangle + |2\rangle}{\sqrt{3}}$$
is a pure state.
- We can represent pure states as a density operator by bra-ing it on the right:
$$\rho_\psi = |\psi\rangle\langle\psi| = \frac{(|0\rangle + |1\rangle + |2\rangle)(\langle 2| + \langle 1| + \langle 0|)}{3}$$
- But density operators can also represent mixed states, such as
$$\rho_{\mathsf{max}} = \frac{|+\rangle\langle +| + |-\rangle\langle -|}{2}$$
- The purity of a state is the trace of its square (1 is pure):
$$\operatorname{tr}(\rho_\psi^2) = 1, \qquad \operatorname{tr}(\rho_{\mathsf{max}}^2) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

Projection Valued Measures

Let $\mathcal{H}$ be the Hilbert space of pure states and (X, Σ) a measurable space where Σ is the Borel sets of X.[6] Recall that $\mathcal{B}$ is the C^*-algebra of bounded operators on $\mathcal{H}$. Let $\mathcal{O}$ be the subspace of bounded self-adjoint operators.

A **projection valued measure** is a function

$$\pi : \Sigma \to \mathcal{O}$$

such that:

- $\pi(E)$ is an orthogonal projection for all $E \in \Sigma$.
- $\pi(\varnothing) = 0$ and $\pi(X) = 1_{\mathcal{H}}$, and also $\pi(E_1 \cap E_2) = \pi(E_1)\pi(E_2)$.
- For pairwise disjoint $\{E_n\}_{n \geq 1} \subset \Sigma$,

$$\pi\left(\bigcup_{n \geq 1} E_n\right) x = \sum_{n \geq 1} \pi(E_n) x.$$

[6] X can be $\mathbf{R}$, or perhaps $\mathbf{R}^3$, or even just a finite set (this is common).

Positive Operator Valued Measures

Let $\mathcal{H}$ and $\mathcal{O}$ be as prior, and let $(X, 2^X)$ be such that $X = \{x_j\}_{j=1}^k$ (this is the most frequently encountered case).

A **POVM** is a function

$$\Pi : 2^X \to \mathcal{O}$$

such that

$$\Pi(X) = \mathbf{1}_{\mathcal{H}}$$

and

$$\Pi(x_j) = \pi(x_j)^\dagger \pi(x_j)$$

where π is a projection valued measure.

That is, a POVM is basically a finite set of positive operators $\Pi(x_j)$ that sum to the identity. Note that the elements of a POVM need not commute with each other.

Measuring an Observable

Essentially, X represents the possible values an observable can take on. The probability p_k of reading the value x_k from the measurement $\pi(x_k)$ when the system is in state ρ is then

$$p_k = \mathrm{tr}(\Pi(x_k)\rho) = \langle \rho | \Pi(x_k) \rangle_{\mathcal{B}} \,.$$

But after measuring the state of the system, it changes!

$$\rho \mapsto \frac{\pi(x_k)\rho\pi(x_k)^\dagger}{p_k}$$

The **expected value** of an observable $A \in \mathcal{O}$ can be reframed in terms of density operators:

$$\langle A \rangle_\rho = \mathrm{tr}(A\rho) = \langle \rho | A \rangle_{\mathcal{B}}$$

and simliarly for the uncertainty.

Why POVMs?

In rough analogy, density operators (arbitary quantum states) are to pure states as POVMs are to projection valued measures.

In the real, physical world (a laboratory, a quantum computer, etc.), pure states are impossible to prepare, and projection valued measures are similarly theoretical idealization.

Since every physical state of a quantum system is then somewhat mixed, POVMs are essential to describing how those mixed states are measured.

An Example (from Mike & Ike)

Let $\mathcal{H} = \mathbf{C}^2$ and say $\{|0\rangle, |1\rangle\}$ is a basis for $\mathcal{H}$.

A normalized vector (or **qubit**) in $\mathcal{H}$ is then a complex linear combination $\alpha|0\rangle + \beta|1\rangle$ such that $|\alpha|^2 + |\beta|^2 = 1$.

Suppose Alice gives Bob a qubit, and that Bob knows Alice's qubit is in one of two states:

$$|\psi_1\rangle = |0\rangle, \qquad |\psi_2\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}.$$

It is impossible to know with perfect reliability which state the qubit is in, but Bob can do the next best thing: *perform a measurement that distinguishes state some of the time, but never makes an error (at the cost of sometimes yielding an inconclusive result).*

Example ctd.

Consider the following POVM:

$$E_1 = \frac{\sqrt{2}}{1+\sqrt{2}} |1\rangle\langle 1|$$

$$E_2 = \frac{\sqrt{2}}{1+\sqrt{2}} \frac{(|0\rangle - |1\rangle)(\langle 0| - \langle 1|)}{2}$$

$$E_3 = \mathbf{1}_{\mathcal{H}} - E_1 - E_2$$

If Bob reads E_3 then he knows nothing about the qubit.

Example ctd.

However!

If Bob reads E_1, then he knows Alice didn't send the qubit in state $|\psi_1\rangle$, so it must be in state $|\psi_2\rangle$:

$$\langle\psi_1|E_1|\psi_1\rangle = \frac{\sqrt{2}}{1+\sqrt{2}} \langle 0|1\rangle \langle 1|0\rangle = 0$$

Example ctd.

Similarly, if Bob reads E_2, then he knows Alice didn't send the qubit in state $|\psi_2\rangle$, so it must be in state $|\psi_1\rangle$:

$$\langle\psi_2|E_2|\psi_2\rangle = \frac{\sqrt{2}}{1+\sqrt{2}}\left\langle \frac{\langle 0| + \langle 1|}{\sqrt{2}} \middle| \frac{(|0\rangle - |1\rangle)(\langle 0| - \langle 1|)}{2} \middle| \frac{|0\rangle + |1\rangle}{\sqrt{2}} \right\rangle$$
$$= \frac{\sqrt{2}}{4(1+\sqrt{2})}(\langle 0| + \langle 1|)(|0\rangle\langle 0| - |0\rangle\langle 1| - |1\rangle\langle 0| + |1\rangle\langle 1|)(|0\rangle + |1\rangle)$$

Letting $A = |0\rangle\langle 0| - |0\rangle\langle 1| - |1\rangle\langle 0| + |1\rangle\langle 1|$ we have

$$\langle\psi_2|E_2|\psi_2\rangle = \frac{\sqrt{2}}{4+4\sqrt{2}}\left(\langle 0|A|0\rangle + \langle 0|A|1\rangle + \langle 1|A|0\rangle + \langle 1|A|1\rangle\right)$$

where

$$\langle 0|A|0\rangle = 1, \qquad \langle 0|A|1\rangle = -1, \qquad \langle 1|A|0\rangle = -1, \qquad \langle 1|A|1\rangle = 1$$

That is, $\langle\psi_2|E_2|\psi_2\rangle = 0$.

Probability Data of a POVM

Is this POVM any good? Yes!

- $\langle E_1 | P_{\psi_0} \rangle_{\mathcal{B}} = \mathrm{tr}(\frac{\sqrt{2}}{1+\sqrt{2}} |1\rangle \langle 1|0\rangle \langle 0|) = \mathrm{tr}(\mathbf{0}) = 0.0$ (as shown)
- $\langle E_1 | P_{\psi_1} \rangle_{\mathcal{B}} = \mathrm{tr}(\frac{\sqrt{2}}{1+\sqrt{2}} |1\rangle\langle 1| \frac{(|0\rangle+|1\rangle)(\langle 0|+\langle 1|)}{2}) = \frac{\sqrt{2}}{2+2\sqrt{2}} = 0.292$
- $\langle E_2 | P_{\psi_0} \rangle_{\mathcal{B}} = \mathrm{tr}(\frac{\sqrt{2}}{1+\sqrt{2}} \frac{(|0\rangle-|1\rangle)(\langle 0|-\langle 1|)}{2} |0\rangle\langle 0|) = \frac{\sqrt{2}}{2+2\sqrt{2}} = 0.292$
- $\langle E_2 | P_{\psi_1} \rangle_{\mathcal{B}} = 0.0$ (as shown)

So we get a defnititive result (either ψ_0 or ψ_1, without a doubt) roughly 58 percent of the time.

Thank you. Questions?

References:

- Besides the relevant Wikipedia pages, more information on POVMs can be found in the book *Quantum Computation and Quantum Information* by Michael Nielsen and Issac Chuang. This is commonly known as "Mike and Ike," and it is the source from which I drew my example.
- For an more detailed look at what we did, I recommend consulting the first three chapters of Landau-Lifshitz Vol 3: *Quantum Mechanics (nonrelativistic theory)*.
- I also made use of `arXiv:1902.00967v2` in the process of solidifying my understanding of density operators.