Operator moments revisited

Recall that in a ‘measurement of the observable $\mathbf{O}_q$,’ we utilize the spectral decomposition

$$\begin{aligned}\mathbf{O}_q &= \sum_i \lambda_i^q |i^q\rangle\langle i^q| \\ &= \sum_i \lambda_i^q \mathbf{P}_i^q \end{aligned} \tag{1}$$

to specify a complete set of orthogonal projectors and associate the value $q = \lambda_i^q$ with the occurance of the i^{th} outcome.

Suppose we are given many ‘copies’ of a physical system prepared in the state $|\Psi\rangle$, which is not an eigenstate of $\mathbf{O}_q$, and that we perform a measurement of $\mathbf{O}_q$ on each of them. In the sequence of measurements, q is effectively a random variable with probability distribution

$$\Pr(q = \lambda_i^q) = \langle\Psi|\mathbf{P}_i^q|\Psi\rangle. \tag{2}$$

By the way, this kind of relation shows us that the projection operators had better be Hermitian, since we want all probabilities to be real numbers so

$$\langle\Psi|\mathbf{P}_i^q|\Psi\rangle = (\langle\Psi|\mathbf{P}_i^q|\Psi\rangle)^* = \langle\Psi|(\mathbf{P}_i^q)^\dagger|\Psi\rangle.$$

We can easily compute the average (mean) value of q,

$$\begin{aligned}\langle q\rangle &= \sum_i \lambda_i^q \Pr(q = \lambda_i^q) \\ &= \sum_i \lambda_i^q \langle\Psi|\mathbf{P}_i^q|\Psi\rangle \\ &= \langle\Psi|\mathbf{O}_q|\Psi\rangle.\end{aligned} \tag{3}$$

The last expression $\langle\Psi|\mathbf{O}_q|\Psi\rangle$ is called the ‘expectation value of $\mathbf{O}_q$ with respect to $|\Psi\rangle$’ and is commonly written $\langle\mathbf{O}_q\rangle_\Psi$, or even just $\langle\mathbf{O}_q\rangle$.

Just to look at a quick example, consider a quantum model for dice (actually, one six-sided die). We should probably choose a six-dimensional Hilbert space, with one basis ket for each of the die’s faces:

$$\begin{aligned}|\Psi\rangle &\in H_6, \\ H_6 &= \text{span}\{|1\rangle, |2\rangle, |3\rangle, |4\rangle, |5\rangle, |6\rangle\}.\end{aligned}$$

Then the projectors corresponding to a measurement of the property q corresponding to ‘the number showing on the side facing up’ would simply be

$$\mathbf{P}_1^q = |1\rangle\langle 1|, \mathbf{P}_2^q = |2\rangle\langle 2|, etc\ldots$$

and we could indeed define an observable

$$\begin{aligned}\mathbf{O}^q &= \sum_{k=1}^{6} k\mathbf{P}_k^q \\ &= |1\rangle\langle 1| + 2|2\rangle\langle 2| + 3|3\rangle\langle 3| + 4|4\rangle\langle 4| + 5|5\rangle\langle 5| + 6|6\rangle\langle 6|.\end{aligned}$$

(Note that there exist other possible bases in which to measure, *e.g.*,

$$\begin{aligned}
|a_+\rangle &\equiv \frac{1}{\sqrt{2}}(|1\rangle + |2\rangle), \quad |a_-\rangle \equiv \frac{1}{\sqrt{2}}(|1\rangle - |2\rangle),\\
|b_+\rangle &\equiv \frac{1}{\sqrt{2}}(|3\rangle + |4\rangle), \quad |b_-\rangle \equiv \frac{1}{\sqrt{2}}(|3\rangle - |4\rangle),\\
|c_+\rangle &\equiv \frac{1}{\sqrt{2}}(|5\rangle + |6\rangle), \quad |c_-\rangle \equiv \frac{1}{\sqrt{2}}(|5\rangle - |6\rangle).
\end{aligned}$$

So we can talk about measuring things like

$$\{|a_+\rangle\langle a_+|, |a_-\rangle\langle a_-|, |b_+\rangle\langle b_+|, \ldots\}.$$

)

So, *e.g.*, for the pre-measurement state

$$|\Psi\rangle = c_1|1\rangle + c_4|4\rangle + c_6|6\rangle,$$

we would predict outcome probabilities

$$\begin{aligned}
&\Pr(1) = |c_1|^2,\ \Pr(2) = 0,\ \Pr(3) = 0,\\
&\Pr(4) = |c_4|^2,\ \Pr(5) = 0,\ \Pr(6) = |c_6|^2,
\end{aligned}$$

and

$$\langle q\rangle \equiv \langle \mathbf{O}^q\rangle = |c_1|^2 + 4|c_4|^2 + 6|c_6|^2.$$

In fact we can easily compute the higher statistical moments of q just as easily, using the spectral decomposition:

$$\begin{aligned}
\langle q^n\rangle &= \sum_i (\lambda_i^q)^n \langle\Psi|\mathbf{P}_i^q|\Psi\rangle\\
&= \langle\Psi|\sum_i (\lambda_i^q)^n \mathbf{P}_i^q|\Psi\rangle\\
&= \langle\Psi|\left(\sum_i \lambda_i^q \mathbf{P}_i^q\right)^n|\Psi\rangle\\
&= \langle\mathbf{O}_q^n\rangle.
\end{aligned} \tag{4}$$

Superpositions, uncertainty, and nonorthogonality

Continuing with the above scenario, the variance of q is defined by

$$\begin{align}
(\Delta q)^2 &\equiv \langle q^2\rangle - \langle q\rangle^2 \notag\\
&= \left\langle \mathbf{O}_q^2\right\rangle - \langle \mathbf{O}_q\rangle^2 \tag{5}\\
&\equiv (\Delta\mathbf{O}_q)^2. \tag{\#}
\end{align}$$

The square-root of this quantity, which is the standard deviation of q, is also known as the ‘uncertainty’ $\Delta\mathbf{O}_q$ with respect to the state $|\Psi_A\rangle$:

$$\Delta\mathbf{O}_q \equiv \sqrt{\left\langle \mathbf{O}_q^2\right\rangle_{\Psi_A} - \langle \mathbf{O}_q\rangle_{\Psi_A}^2}\,. \tag{6}$$

(By convention we choose the positive square-root.)

For an arbitrary state $|\Psi_A\rangle$,

$$\Delta \mathbf{O}_q = 0 \tag{7}$$

implies that $|\Psi_A\rangle$ ***must*** be an eigenstate of $\mathbf{O}_q$. To see this, write

$$\begin{aligned}
\left\langle \mathbf{O}_q^2 \right\rangle - \langle \mathbf{O}_q \rangle^2 &= \langle \Psi_A | \sum_i (\lambda_i^q)^2 \mathbf{P}_i^q | \Psi_A \rangle - \left(\langle \Psi_A | \sum_i \lambda_i^q \mathbf{P}_i^q | \Psi_A \rangle \right)^2 \\
&= \sum_i (\lambda_i^q)^2 \langle \Psi_A | \mathbf{P}_i^q | \Psi_A \rangle - \left(\sum_i \lambda_i^q \langle \Psi_A | \mathbf{P}_i^q | \Psi_A \rangle \right)^2 \\
&= \sum_i (\lambda_i^q)^2 \langle \mathbf{P}_i^q \rangle - \sum_{ij} \lambda_i^q \lambda_j^q \langle \mathbf{P}_i^q \rangle \left\langle \mathbf{P}_j^q \right\rangle.
\end{aligned} \tag{8}$$

The latter can only equal zero if $\langle \mathbf{P}_i^q \rangle \left\langle \mathbf{P}_j^q \right\rangle = \delta_{ij} \langle \mathbf{P}_i^q \rangle$, and this can only happen if $\langle \mathbf{P}_i^q \rangle$ only takes on the values 0 and 1. Hence, $|\Psi_A\rangle$ must be an eigenstate of $\mathbf{O}_q$.

If the variance of q (with respect to the state $|\Psi_A\rangle$) is nonzero, which implies and is implied by $\Delta \mathbf{O}_q > 0$, then ***the value of*** q ***cannot be predicted with complete certainty in a measurement on*** $|\Psi_A\rangle$. To put it another way, the value of q is not well-defined for system states that are superpositions of different eigenstates (corresponding to distinct eigenvalues) of the observable $\mathbf{O}_q$.

Hence we see that the possibility of 'instrinsic' uncertainty in the value of a given observable follows directly from the structure of the quantum rules for representation and measurement.

Last week I said the same thing about the possibility for distinct preparations of the state of a quantum system to be imperfectly distinguishable by any measurement that could be made on that system. Are these two features of quantum mechanics related?

Simultaneous eigenstates of two operators

Suppose I have two Hermitian operators, **A** and **B**. In some cases, it is possible to find a complete set of basis states $\{|k\rangle\}$ for the Hilbert space that are simultaneously eigenstates of both operators:

$$\begin{aligned}
\mathbf{A}|k\rangle &= a_k|k\rangle, \\
\mathbf{B}|k\rangle &= b_k|k\rangle.
\end{aligned} \tag{9}$$

If the above relations are true, then it must be the case that

$$\begin{aligned}
\mathbf{BA}|k\rangle &= \mathbf{B}a_k|k\rangle = a_k b_k|k\rangle, \\
\mathbf{AB}|k\rangle &= \mathbf{A}b_k|k\rangle = b_k a_k|k\rangle.
\end{aligned} \tag{10}$$

Clearly then,

$$(\mathbf{AB} - \mathbf{BA})|k\rangle = 0. \tag{11}$$

Since the states $\{|k\rangle\}$ form a complete basis, any state $|\Psi\rangle$ in the Hilbert space can be written

$$|\Psi\rangle = \sum_k c_k|k\rangle. \tag{12}$$

Therefore

$$
\begin{aligned}
(\mathbf{AB}-\mathbf{BA})|\Psi\rangle &= \sum_k c_k(\mathbf{AB}-\mathbf{BA})|k\rangle \\
&= 0. \qquad 13
\end{aligned}
$$

It must therefore be true at the operator level that **A** ‘commutes’ with **B**,

$$[\mathbf{A},\mathbf{B}] \equiv \mathbf{AB}-\mathbf{BA} = 0. \qquad 14$$

Hence $[\mathbf{A},\mathbf{B}] = 0$ is a ***necessary*** condition for the existence of a complete set of basis states that are simultaneous eigenstates of **A** and **B**.

As it turns out, this condition is also sufficient **–** see Merzbacher §10.4.

Heisenberg Uncertainty Relations

(Taken from Merzbacher §10.5)

Suppose we have two Hermitian operators **A** and **B**, whose commutator is nonzero:

$$[\mathbf{A},\mathbf{B}] \equiv \mathbf{AB}-\mathbf{BA} = i\mathbf{C}. \qquad 15$$

If not zero, the commutator of **A** and **B** must be some operator, and defining **C** in the above way guarentees that **C** is Hermitian:

$$
\begin{aligned}
(\mathbf{AB}-\mathbf{BA})^\dagger &= -i\mathbf{C}^\dagger \\
&= \mathbf{BA}-\mathbf{AB}, \qquad 16
\end{aligned}
$$

hence $\mathbf{C}^\dagger = \mathbf{C}$.

By a straightforward application of the Schwartz inequality (the inner product operations satisfy the Schwartz inequality by definition)

$$\langle\Psi_A|\Psi_A\rangle\langle\Psi_B|\Psi_B\rangle \geq |\langle\Psi_A|\Psi_B\rangle|^2, \qquad 17$$

we can now prove that the Heisenberg Uncertainty Relation

$$\Delta\mathbf{A}\Delta\mathbf{B} \geq \frac{1}{2}|\langle\mathbf{C}\rangle| \qquad 18$$

holds for every state $|\Psi\rangle$ in the Hilbert space.

First, define

$$
\begin{aligned}
|\Psi_A\rangle &= (\mathbf{A}-\langle\mathbf{A}\rangle\mathbf{1})|\Psi\rangle, \\
|\Psi_B\rangle &= (\mathbf{B}-\langle\mathbf{B}\rangle\mathbf{1})|\Psi\rangle. \qquad 19
\end{aligned}
$$

Plugging this into the Schwartz inequality, we get

$$
\begin{aligned}
\left\langle(\mathbf{A}-\langle\mathbf{A}\rangle)^2\right\rangle_\Psi\left\langle(\mathbf{B}-\langle\mathbf{B}\rangle)^2\right\rangle_\Psi &\geq |\langle\Psi|(\mathbf{A}-\langle\mathbf{A}\rangle\mathbf{1})(\mathbf{B}-\langle\mathbf{B}\rangle\mathbf{1})|\Psi\rangle|^2, \\
(\Delta\mathbf{A})^2(\Delta\mathbf{B})^2 &\geq |\langle\Psi|(\mathbf{A}-\langle\mathbf{A}\rangle\mathbf{1})(\mathbf{B}-\langle\mathbf{B}\rangle\mathbf{1})|\Psi\rangle|^2. \qquad 20
\end{aligned}
$$

Here we have used the fact that the expectation value of a Hermitian operator is always real, and we have noted that, *e.g.*,

$$\begin{aligned}\left\langle(\mathbf{A}-\langle\mathbf{A}\rangle)^2\right\rangle_\Psi &= \left\langle\mathbf{A}^2-2\langle\mathbf{A}\rangle\mathbf{A}+\langle\mathbf{A}\rangle^2\right\rangle_\Psi\\ &= \left\langle\mathbf{A}^2\right\rangle_\Psi-2\langle\mathbf{A}\rangle_\Psi^2+\langle\mathbf{A}\rangle_\Psi^2\\ &= \left\langle\mathbf{A}^2\right\rangle_\Psi-\langle\mathbf{A}\rangle_\Psi^2\\ &= (\Delta\mathbf{A})^2.\end{aligned} \qquad 21$$

Expanding out the RHS of (20),

$$\begin{aligned}(\Delta\mathbf{A})^2(\Delta\mathbf{B})^2 &\geq |\langle\Psi|(\mathbf{AB}-\mathbf{A}\langle\mathbf{B}\rangle-\mathbf{B}\langle\mathbf{A}\rangle+\langle\mathbf{A}\rangle\langle\mathbf{B}\rangle)|\Psi\rangle|^2\\ &= \left|\langle\Psi|\Big(\frac{1}{2}(\mathbf{AB}+i\mathbf{C}+\mathbf{BA})-\mathbf{A}\langle\mathbf{B}\rangle-\mathbf{B}\langle\mathbf{A}\rangle+\langle\mathbf{A}\rangle\langle\mathbf{B}\rangle\Big)|\Psi\rangle\right|^2\\ &= \left|\Big\langle\frac{1}{2}(\mathbf{AB}+\mathbf{BA})-\mathbf{A}\langle\mathbf{B}\rangle-\mathbf{B}\langle\mathbf{A}\rangle+\langle\mathbf{A}\rangle\langle\mathbf{B}\rangle\Big\rangle+\frac{i}{2}\langle\mathbf{C}\rangle\right|^2\\ &\geq \frac{1}{4}|\langle\mathbf{C}\rangle|^2.\end{aligned} \qquad 22$$

The last inequality follows from the fact that the expectation values are real. Taking the positive square-root of both sides, we have the Heisenberg Uncertainty Relation

$$\Delta\mathbf{A}\Delta\mathbf{B} \geq \frac{1}{2}|\langle\mathbf{C}\rangle|. \qquad 23$$

This is certainly an interesting (and sometimes useful) inequality. But does it tell us any interesing physics?

One consequence of equation (23) is that when **A** and **B** do not commute, these operators are ***both*** uncertain for all states in the Hilbert space ***except*** those satisfying $\langle\mathbf{C}\rangle = 0$. This is either a profound or a trivial revelation, depending on whether we are working in finite or infinite dimensions. Note that the derivation of equation (23) is valid in either case.

In finite dimensions, we know that $\Delta\mathbf{A} = \mathbf{0}$ for eigenstates of **A**, and $\Delta\mathbf{B} = \mathbf{0}$ for eigenstates of **B**. Looking at the definition

$$\mathbf{AB}-\mathbf{BA} = i\mathbf{C}, \qquad 24$$

we correspondingly see that $\langle\mathbf{C}\rangle$ must be zero for eigenstates of either **A** or **B** (prove this for yourself as an excercise). For any state in the Hilbert space that is not an eigenstate of either **A** or **B**, it is not surprising that both operators are uncertain – based on the discussion above. And let's face it – most states in the Hilbert space fall into this category! Hence, given an arbitrary pair of observables on a finite-dimensional Hilbert space, both will be uncertain except for a vanishingly small set of states (unless **A** or **B** is proportional to the identity operator).

In infinite dimensions, however, we shall see that that are certain special pairs of observables **A**, **B** whose commutator is proportional to the identity! For these 'conjugate' pairs, the Heisenberg Uncertainty Relation tells us that $\Delta\mathbf{A}\Delta\mathbf{B} > \mathbf{0}$ for ***all*** states in the Hilbert space. Yet we know that **A** and **B** must still have eigenstates, for which either $\Delta\mathbf{A}$ or $\Delta\mathbf{B}$ should be zero – what's going on!?

Ensembles of quantum states; the density operator

Let's say we're working in a nice, simple two-dimensional Hilbert space and that we've chosen orthonormal basis kets $|x\rangle$ and $|y\rangle$.

For the following discussion, we'll need to define two state vectors

$$\begin{aligned} |\Psi_A\rangle &= a_x|x\rangle + a_y|y\rangle, \\ |\Psi_B\rangle &= b_x|x\rangle + b_y|y\rangle. \end{aligned} \qquad 25$$

Suppose I ask you to perform a series of N measurements on this system (here N is just a large integer), corresponding to the projectors

$$\begin{aligned} \mathbf{P}_x &= |x\rangle\langle x|, \\ \mathbf{P}_y &= |y\rangle\langle y|. \end{aligned} \qquad 26$$

The trick is, in this series of measurements I will sometimes prepare the initial state $|\Psi_A\rangle$ (with probability p) and sometimes $|\Psi_B\rangle$ (with probability $1-p$). That is, I will be giving you a ***mixed*** ensemble of quantum states. How shall we predict the overall number of times n_x we expect to obtain the measurement outcome x?

According to simple probability theory,

$$\begin{aligned} n_x &= N[\Pr(\Psi_A)\Pr(x|\Psi_A) + \Pr(\Psi_B)\Pr(x|\Psi_B)] \\ &= N[p\langle\Psi_A|\mathbf{P}_x|\Psi_A\rangle + (1-p)\langle\Psi_B|\mathbf{P}_x|\Psi_B\rangle] \\ &= N\left[p|a_x|^2 + (1-p)|b_x|^2\right]. \end{aligned} \qquad 27$$

Note that since $0 \le p \le 1$, the quantity n_x/N is bounded from below by the smaller of $|a_x|^2$ and $|b_x|^2$. In particular, if both of these quantities are nonzero than n_x must also be greater than zero.

A very different expression for n_x would be obtained if, instead of a mixed ensemble, I were to present you with a 'coherent superposition' of the states $|\Psi_A\rangle$ and $|\Psi_B\rangle$, corresponding to the 'pure' state $|\Psi(p_A,p_B)\rangle$

$$\begin{aligned} |\Psi(p_A,p_B)\rangle &= p_A|\Psi_A\rangle + p_B|\Psi_B\rangle \\ &= (p_A a_x + p_B b_x)|x\rangle + (p_A a_y + p_B b_y)|y\rangle. \end{aligned} \qquad 28$$

(Note that we should be careful in choosing p_A and p_B such that $|\Psi(p_A,p_B)\rangle$ is normalized.) In this case, we would predict

$$\begin{aligned} n_x &= N\langle\Psi(p_A,p_B)|\mathbf{P}_x|\Psi(p_A,p_B)\rangle \\ &= N|p_A a_x + p_B b_x|^2. \end{aligned} \qquad 29$$

Note that in certain cases, *e.g.* $p_A a_x = -p_B b_x \neq 0$, it is possible to choose p_A, p_B such that $n_x = 0 < |a_x|^2, |b_x|^2$ through the phenomenon of *destructive interference*. This is the truly important distinction between coherent superpositions (of the type that produce a single pure state) and incoherent admixtures (of the type that produce a mixed ensemble of quantum states).

A simplifying notation can be introduced for performing computations with mixed quantum ensembles. If the states composing the ensemble are labelled $|\Psi_i\rangle$ and have probabilities p_i, then the ***density operator*** for this ensemble is defined as

$$\boldsymbol{\rho} = \sum_i p_i|\Psi_i\rangle\langle\Psi_i|. \qquad 30$$

Note that $\boldsymbol{\rho}$ is indeed an operator on the Hilbert space, and has the form of a linear combination of projection operators. However, the ensemble representation used to define a density operator is not necessarily also a spectral decomposition, as the various $|\Psi_i\rangle$ that constitute an ensemble do not need to be mutually orthogonal.

The density operator is automatically Hermitian, and furthermore has the property that

$$\mathrm{Tr}\boldsymbol{\rho} = 1. \qquad 31$$

Here Tr denotes the ‘trace’ operation

$$\mathrm{Tr}\boldsymbol{\rho} = \sum_k \langle k|\boldsymbol{\rho}|k\rangle, \qquad 32$$

where $\{|k\rangle\}$ is ***any*** orthonormal basis for the Hilbert space – the numerical result is independent of choice of basis.

In particular, since $\boldsymbol{\rho}$ is Hermitian we can chose to take the trace in its own eigenbasis, which makes it clear that

$$\mathrm{Tr}\boldsymbol{\rho} = \sum_i \lambda_i^\rho, \qquad 33$$

where λ_i^ρ are the eigenvalues of $\boldsymbol{\rho}$. Note that if $\boldsymbol{\rho}$ happens to be available in matrix form, we can further make use of the fact that the sum of the eigenvalues of a matrix is equal to the sum of its diagonal elements.

Density operators can represent either pure states,

$$\boldsymbol{\rho} = |\Psi\rangle\langle\Psi|, \qquad 34$$

or mixed states

$$\boldsymbol{\rho} = \sum_i p_i |\Psi_i\rangle\langle\Psi_i| \qquad 35$$

where there is more than one $p_i > 0$.

Note that in the former (pure state) case $\boldsymbol{\rho}$ is a true projection operator, so

$$\textbf{pure}: \quad \boldsymbol{\rho}^2 = \boldsymbol{\rho}. \qquad 36$$

In particular, $\mathrm{Tr}\boldsymbol{\rho}^2 = 1$ for a pure state.

For a mixed state, however, we can use the spectral decomposition to show that $\mathrm{Tr}\boldsymbol{\rho}^2 < 1$. We start by writing

$$\begin{aligned}\boldsymbol{\rho}^2 &= \left(\sum_i \lambda_i^\rho \mathbf{P}_i^\rho\right)^2 \\ &= \sum_i (\lambda_i^\rho)^2 \mathbf{P}_i^\rho, \qquad 37\end{aligned}$$

and note that since $\mathrm{Tr}\boldsymbol{\rho} = \sum_i \lambda_i^\rho = 1$, each of the λ_i^ρ must be strictly less than one for a mixed state. Hence the eigenvalues of $\boldsymbol{\rho}^2$, which are equal to the $(\lambda_i^\rho)^2$, must add up to *less than* one.