Chapter 44

The Uncertainty Principle

Derivation of the uncertainty inequality and how the uncertainty principle has become less central in Quantum Mechanics than it was.

Last Update: 9/4/12

Whilst there has been no revolution, quantum mechanics had been subject to significant evolution over the last 40 years or so. Its home, though, is still Hilbert space. I confess to being rather bemused that the most fundamental of physical theories (we are led to believe) is based on something so banal as a linear vector space. How different from, for example, general relativity with its infinitely rich manifold structure. Notwithstanding this, the last few decades has seen the subject reveal yet more of its strangeness.

More than half the chapters in this book are on quantum mechanics. It is telling that less than half of these rely, even implicitly, on the uncertainty principle or the closely associated non-commutative properties of observables. This stands in contrast to the way in which the subject used to be taught, and perhaps still is, with uncertainty and non-commutativity as the dominant ideas. In truth the central feature of quantum mechanics which give rise to most of its weirdness is superposition rather than non-commutativity, a point noted by Dirac himself, despite being the originator of canonical quantisation based on non-commutation. Nevertheless, the uncertainty principle is still a key concept in QM, and is one of the defining distinctions between QM and classical physics. And the non-zero values of commutators, from which uncertainty follows, also specify or constrain the possible values taken by observables (as illustrated in Chapters 32 and 35).

If we pick an arbitrary pair of Hermetian operators acting on a given Hilbert space, in almost all cases they will not commute. This means they will have different eigenvectors. The proof is simple. Suppose they had the same eigenvectors. Each operator can be expressed in the form $A = U\Lambda_A U^+$ and $B = U\Lambda_B U^+$, where U is the matrix of normalised orthogonal column eigenvectors, and is the same for the two operators by assumption. The matrices Λ_A and Λ_B are diagonal and consist of the eigenvalues, which may differ. Hence we have, $AB = U\Lambda_A U^+ U\Lambda_B U^+ = U\Lambda_A \Lambda_B U^+ = U\Lambda_B \Lambda_A U^+ = U\Lambda_B U^+ U\Lambda_A U^+ = BA$, so that operators with the same eigenvectors necessarily commute, noting that Λ_A and Λ_B commute by virtue of being diagonal.

But quantum mechanics holds that an observation projects the state of the system onto one of the observables eigenstates. A system is in a definite state of an observable is it is already in an eigenstate of that observable, since there is then only one possible outcome to its measurement. It follows that a system cannot be simultaneously in a definite state of two non-commuting observables, since it cannot be in an eigenstate of both. This is the essence of the uncertainty principle.

However, the uncertainty principle as usually stated goes further. It considers a quantification of the uncertainty in each of a pair of complementary observables. For our purposes ‘complementary’ observables can be defined as pairs of observables whose commutator is $i\hbar$, the archetypal example being momentum and position in any Cartesian direction, e.g., $[\hat{P}_x, \hat{x}] = i\hbar$. The uncertainty principle states that the product of the uncertainties in the two quantities has a smallest possible, non-zero value, namely $\Delta x \cdot \Delta P_x \geq \frac{\hbar^2}{4}$. This quantified form of the uncertainty principle follows from the commutation relation. This is proved below.

We assume only that we have two Hermetian operators with commutator $[\hat{P}, \hat{x}] = i\hbar$. Whilst one interpretation of this is that these operators represent Cartesian components of momentum and position, the result to be proved holds for any operators with this commutator.

Consider the modified operators defined as the difference of these operators from some arbitrary, but constant, value,

$$\Delta\hat{x} = \hat{x} - x_0 \qquad \text{and} \qquad \Delta\hat{P} = \hat{P} - P_0 \tag{1}$$

where x_0 and P_0 are just numerical values. Hence, the mean square deviation of the observables $\hat{x}$ and $\hat{P}$ from these constant values are given by,

$$\left\langle \Delta x^2 \right\rangle = \langle \psi | \Delta\hat{x}^2 | \psi \rangle \qquad \text{and} \qquad \left\langle \Delta P^2 \right\rangle = \langle \psi | \Delta\hat{P}^2 | \psi \rangle \tag{2}$$

where the arbitrary state of the system is labelled ψ. This state can be expanded in some arbitrary orthonormal basis as $|\psi\rangle = c_i |\phi_i\rangle$, where summation over repeated subscripts will be assumed hereafter. Note that the basis states $\{|\phi_i\rangle\}$ need not be eigenstates of either $\hat{x}$ or $\hat{P}$. The action of these operators on the basis can be written,

$$\Delta\hat{x}|\phi_i\rangle = a_{ij}|\phi_j\rangle \qquad \text{and} \qquad \Delta\hat{P}|\phi_i\rangle = b_{ij}|\phi_j\rangle \tag{3}$$

Substitution into (2) yields,

$$\left\langle \Delta x^2 \right\rangle = |\overline{u}|^2 \qquad \text{and} \qquad \left\langle \Delta P^2 \right\rangle = |\overline{v}|^2 \tag{4}$$

where the complex-valued vectors $\overline{u}$ and $\overline{v}$ are defined by,

$$u_j \equiv c_i a_{ij} \qquad \text{and} \qquad v_j \equiv c_i b_{ij} \tag{5}$$

But Schwartz's inequality tells us that, for any complex valued vectors in any number of dimensions,

$$|\overline{u}|^2 |\overline{v}|^2 \geq \left| \overline{u}^* \cdot \overline{v} \right|^2 \tag{6}$$

This inequality is obvious. It says that the product of the squared-lengths of two vectors must be greater than or equal to the square of their dot product. Since the dot product is the product of the lengths of the vectors times the cosine of the included angle, this is clearly true – the equality holding only when the vectors are parallel and the included angle is zero. The only thing that complicates this is the complex nature of the vectors. However, this has no effect on the LHS, which involves only the moduli of the vectors. On the RHS the effect of the phase difference between the two vectors can only lead to partial cancellations, and hence further reduces the magnitude of the RHS. QED.

Schwartz's inequality suggests we consider the expectation value of the product of $\Delta\hat{x}$ and $\Delta\hat{P}$, i.e. $\langle\psi|\Delta\hat{x}\Delta\hat{P}|\psi\rangle$, which, on substitution of (5) does indeed turn out to equal $|\overline{v}^* \cdot \overline{u}| = |\overline{u}^* \cdot \overline{v}|$. Hence, Schwartz's inequality gives,

$$\left\langle \Delta x^2 \right\rangle \left\langle \Delta P^2 \right\rangle \geq \left|\left\langle \psi \left|\Delta\hat{x}\Delta\hat{P}\right|\psi\right\rangle\right|^2 \tag{7}$$

The operator appearing on the RHS is not Hermetian, because $\hat{x}$ and $\hat{P}$ do not commute. It may be re-written,

$$\Delta\hat{x}\Delta\hat{P} = \frac{1}{2}[\Delta\hat{x},\Delta\hat{P}] + \frac{1}{2}\left(\Delta\hat{x}\Delta\hat{P} + \Delta\hat{P}\Delta\hat{x}\right) = i\frac{\hbar}{2} + \frac{1}{2}\left(\Delta\hat{x}\Delta\hat{P} + \Delta\hat{P}\Delta\hat{x}\right) \tag{8}$$

The last operator on the RHS is now Hermetian, by virtue of its symmetry and the Hermetian nature of $\hat{x}$ and $\hat{P}$. Consequently it has real eigenvalues and hence a real expectation value in any state. In contrast, the first (purely numerical) term on the RHS is imaginary. It follows that, when taking the absolute square of the RHS of (7) there is no cross-product between the two terms, and we have,

$$\left\langle \Delta x^2 \right\rangle \left\langle \Delta P^2 \right\rangle \geq \frac{\hbar^2}{4} + \frac{1}{4}\left|\left\langle \Delta\hat{x}\Delta\hat{P} + \Delta\hat{P}\Delta\hat{x} \right\rangle\right|^2 \tag{9}$$

But the second term on the RHS is positive, so we get,

$$\left\langle \Delta x^2 \right\rangle \left\langle \Delta P^2 \right\rangle \geq \frac{\hbar^2}{4} \tag{10}$$

which is the usual expression of the uncertainty relation. The term we have dropped from the RHS of (9) can be zero, and hence (10) is the strongest general form of the uncertainty principle.

Note that (10) has been derived for arbitrary values of the constants x_0 and P_0. It is usual to interpret these constants as the expectation values of the operators $\hat{x}$ and $\hat{P}$, i.e.,

$$x_0 = \langle\hat{x}\rangle = \langle\psi|\hat{x}|\psi\rangle \qquad \text{and} \qquad P_0 = \left\langle\hat{P}\right\rangle = \langle\psi|\hat{P}|\psi\rangle \tag{11}$$

so that $\langle\Delta\hat{x}\rangle = \langle\psi|\Delta\hat{x}|\psi\rangle = 0$ and $\left\langle\Delta\hat{P}\right\rangle = \langle\psi|\Delta\hat{P}|\psi\rangle = 0$. It is easily seen that adopting (11) leads to the smallest mean square deviations, $\left\langle\Delta x^2\right\rangle$ and $\left\langle\Delta P^2\right\rangle$, and hence provides the strongest form of uncertainty relation.