Chapter 7

Uncertainty, and the Classical Limit

Previously, we noted that if we precisely measure a spin component of an ensemble of electrons, the other two become effectively randomized, in that their measurements will return $+\frac{\hbar}{2}$ half the times and $-\frac{\hbar}{2}$ half the times. Similarly, if we measure the electrons' positions with precision, the momentum measurement returns will be randomized. These are manifestations of one of the most famous results of quantum mechanics, the Uncertainty Principle, which is the main topic of this chapter. We shall also look at the relation between classical and quantum mechanics with an eye to whether under certain conditions (high energies, for example), the results of standard quantum mechanics are indistinguishable from those of classical mechanics. We start with the Uncertainty Principle, a proper comprehension of which requires an understanding of standard deviation. It is, therefore, to this notion that we now turn.

7.1 Standard Deviation

Consider a cart with seven boxes in it. Two boxes weigh 5*kg* each; three weigh 10*kg* each; one weighs 40*kg*, and one 60*kg*. The average weight (total weight divided by the number of boxes) is $<w>=20kg$. Consider now a second cart with 5 boxes of 20*kg* each, 1 box of 15*kg* and 1 box of 25*kg*. Here too, $<w>=20kg$. Although the averages in the two cases are the same, the distribution of weight among the boxes in the two carts is not: in the second cart the "spread" of weight from $<w>$ is less than in the first, since most boxes weigh exactly 20*kg*. The standard way to measure the spread is as follows. For each *w* we determine how different it is from $<w>$; that is, for each *w* we obtain

$$\Delta w = w - <w>. \qquad (7.1.1)$$

Then, we *square* it and compute its average over all the boxes to obtain the *variance* or *dispersion* σ^2 of the distribution:

$$\sigma^2 = \left\langle (\Delta w)^2 \right\rangle. \qquad (7.1.2)$$

Finally, we take the square root to obtain the *standard deviation* σ, the common measure of spread in distribution: it tells us how much the weights of individual boxes fluctuate about the average. In practice, however, often one obtains the variance by using the following equality here stated without proof:

$$\sigma^2 = < w^2 > - < w >^2. \qquad (7.1.3)$$

That is, the variance is the average of the square minus the square of the average.[1]

EXAMPLE 7.1.1

A cart contains 2 one-kilogram boxes, 3 two-kilogram boxes and 1 three-kilogram box. Let us determine the weight average $< w >$ and the standard deviation. After obtaining

$$< w > = \left(\frac{2+6+3}{6} \right) kg = 1.83 kg, \qquad (7.1.4)$$

we use (7.1.3) to get the variance, and we then take the square root to get the standard deviation. So,

$$< w^2 > = \left(\frac{2 \cdot 1^2 + 3 \cdot 2^2 + 3^2}{6} \right) kg = \frac{23}{6} kg, \text{ and } < w >^2 = \frac{121}{36} kg. \qquad (7.1.5)$$

Hence,

[1] Note that usually, though not always, $< w^2 > \neq < w >^2$. Consider two boxes, one of $2kg$ and one of $4kg$. Then, $< w^2 > = \left(\frac{4+16}{2} \right) kg = 10 kg$, but $< w >^2 = \left(\frac{2+4}{2} \right)^2 kg = 9 kg$.

$$\sigma^2 = \left(\frac{23}{6} - \frac{121}{36}\right)kg = \frac{17}{36}kg = 0.47kg. \qquad (7.1.6)$$

Consequently, the standard deviation is $\sigma = 0.68kg$.

7.2 The Generalized Uncertainty Principle

We can now prove the Generalized Uncertainty Principle (GUP); as we proceed, let us keep in mind that since operator algebra is not commutative, order is important in multiplication.[2]

Let A be an observable with $\hat{A}$ as the corresponding operator. Then, it turns out, although we shall not prove it here, that the variance of A is

$$\sigma_A{}^2 = <(\hat{A} - <A>)^2>, \qquad (7.2.1)$$

where $(\hat{A} - <A>)$ is Hermitian. Remembering that $<Q> = \langle \Psi | \hat{Q}\Psi \rangle$, we get

$$\sigma_A{}^2 = \langle \Psi | (\hat{A} - <A>)^2 \Psi \rangle. \qquad (7.2.2)$$

As the square of an operator is nothing but the double application of the operator, we obtain

$$\sigma_A{}^2 = \langle \Psi | (\hat{A} - <A>)(\hat{A} - <A>)\Psi \rangle. \qquad (7.2.3)$$

Since $(\hat{A} - <A>)\Psi$ is a vector, and $(\hat{A} - <A>)$ is Hermitian, we get

$$\sigma_A{}^2 = \langle (\hat{A} - <A>)\Psi | (\hat{A} - <A>)\Psi \rangle. \qquad (7.2.4)$$

To simplify the notation let

$$\Omega = (\hat{A} - <A>)\Psi. \qquad (7.2.5)$$

Then,

---

[2] The proof is taken from Griffiths, D. J., (1995): 108-10. Elegant as it is, the proof is long and complex, and may therefore be only perused on the first reading.

$$\sigma_A{}^2 = \langle \Omega | \Omega \rangle. \qquad (7.2.6)$$

Similarly, if B is another observable, we can obtain

$$\sigma_B{}^2 = \langle \Pi | \Pi \rangle. \qquad (7.2.7)$$

So,

$$\sigma_A{}^2 \sigma_B{}^2 = \langle \Omega | \Omega \rangle \langle \Pi | \Pi \rangle. \qquad (7.2.8)$$

However, it can be proved that

$$\langle \Omega | \Omega \rangle \langle \Pi | \Pi \rangle \geq \left| \langle \Omega | \Pi \rangle \right|^2.[3] \qquad (7.2.8)$$

Hence,

$$\sigma_A{}^2 \sigma_B{}^2 = \langle \Omega | \Omega \rangle \langle \Pi | \Pi \rangle \geq \left| \langle \Omega | \Pi \rangle \right|^2. \qquad (7.2.9)$$

Now, for any complex number *u*, it can be shown that

$$|u|^2 = \left[\frac{1}{2i}(u - u^*) \right]^2. \qquad (7.2.10)$$

Hence, by setting

$$u = \langle \Omega | \Pi \rangle, \qquad (7.2.11)$$

and remembering that $\langle \Omega | \Pi \rangle = \langle \Pi | \Omega \rangle^*$, we get

$$\sigma_A{}^2 \sigma_B{}^2 = \left[\frac{1}{2i} \left(\langle \Omega | \Pi \rangle - \langle \Pi | \Omega \rangle \right) \right]^2. \qquad (7.2.12))$$

Let us now work on $\langle \Omega | \Pi \rangle$. We note that

$$\langle \Omega | \Pi \rangle = \left\langle (\hat{A} - <A>)\Psi \,|\, (\hat{B} - <B>)\Psi \right\rangle. \qquad (7.2.13)$$

Since $(\hat{A} - <A>)$ is Hermitian, we obtain

$$\langle \Omega | \Pi \rangle = \left\langle \Psi \,|\, (\hat{A} - <A>)(\hat{B} - <B>)\Psi \right\rangle, \qquad (7.2.14)$$

[3] This is the Schwarz Inequality, whose proof one can find in any linear algebra text.

and by multiplication,

$$\left\langle \Omega | \Pi \right\rangle = \left\langle \Psi | (\hat{A}\hat{B} - \hat{A} < B > - < A > \hat{B} + < A >< B >) \Psi \right\rangle, \quad (7.2.15)$$

that is,

$$\left\langle \Omega | \Pi \right\rangle = \left\langle \Psi | \hat{A}\hat{B}\Psi \right\rangle - \left\langle \Psi | \hat{A} < B > \Psi \right\rangle - \left\langle \Psi | < A > \hat{B}\Psi \right\rangle + \left\langle \Psi | < A >< B > \Psi \right\rangle. \quad (7.2.16)$$

By moving the scalars out of the inner products, we get

$$\left\langle \Omega | \Pi \right\rangle = \left\langle \Psi | \hat{A}\hat{B}\Psi \right\rangle - < B > \left\langle \Psi | \hat{A}\Psi \right\rangle - < A > \left\langle \Psi | \hat{B}\Psi \right\rangle + < A >< B > \left\langle \Psi | \Psi \right\rangle. \quad (7.2.17)$$

Since $< Q > = \left\langle \Psi | \hat{Q}\Psi \right\rangle$, we obtain

$$\left\langle \Omega | \Pi \right\rangle = < \hat{A}\hat{B} > - < B >< A > - < A >< B > + < A >< B >, \quad (7.2.18)$$

or

$$\left\langle \Omega | \Pi \right\rangle = < \hat{A}\hat{B} > - < A >< B >. \quad (7.2.19)$$

By repeating the same procedure on $\left\langle \Pi | \Omega \right\rangle$, one can derive

$$\left\langle \Pi | \Omega \right\rangle = < \hat{B}\hat{A} > - < A >< B >. \quad (7.2.20)$$

Using (7.2.19) and (7.2.20), we have

$$\left\langle \Omega | \Pi \right\rangle - \left\langle \Pi | \Omega \right\rangle = < \hat{A}\hat{B} > - < \hat{B}\hat{A} >. \quad (7.2.21)$$

Remembering that the difference of averages is equal to the average of the differences, we infer

$$\left\langle \Omega | \Pi \right\rangle - \left\langle \Pi | \Omega \right\rangle = < \hat{A}\hat{B} - \hat{B}\hat{A} >. \quad (7.2.22)$$

At this point, we need to remember a crucial property of linear operator or, which is the same, matrix multiplication: it is not always the case that $AB = BA$. In other words, linear operator (matrix) multiplication is not commutative. As we know, the difference between the two orderings is the *commutator*: $AB - BA = [A, B]$. Hence,

$$\rangle \Omega \,|\, \Pi \langle - \rangle \Pi \,|\, \Omega \langle \, =< \left[\hat{A}, \hat{B}\right] > . \qquad (7.2.23)$$

If we plug (7.2.23) into (7.2.12), we get

$$\sigma_A{}^2 \sigma_B{}^2 \geq \left(\frac{1}{2i} < \left[\hat{A}, \hat{B}\right] > \right)^2 . \qquad (7.2.24)$$

Equation (7.2.24) expresses GUP. It regards measurements on ensembles. Since the right part of (7.2.24) may be greater than zero (unless the operators commute, in which case $\left[\hat{A}, \hat{B}\right]$ is *always* zero), GUP says that when $\left[\hat{A}, \hat{B}\right] > 0$, if we obtain consistent (narrowly spread) A-measurement returns on many systems in state Ψ, then we shall get a wide spread in B-measurement returns, and conversely if we obtain consistent B-measurement returns then we shall get a wide spread in A-measurement returns, compatibly with (7.2.24). It does not say that if we perform A- measurements on an ensemble we cannot get any consistent results. It does not say that quick repeated A-measurements on one particle are not consistent, for experience teaches us the contrary. It does not say that if we simultaneously (or in very quick succession) measure A and B on one particle we do not get precise results, for, on the contrary, we do get determinate results that, minimally, must involve measurement errors smaller than the variances in (7.2.24), otherwise GUP would be unverifiable.

GUP obviously depends on the non-commutativity of operator multiplication. Remarkably, the proof for GUP does not assume that the incompatible observables be measured on the same system. Hence, suppose we have an ensemble of particles all in the same state and we measure only observable *A* on 50% of them and only observable *B* on the *other* 50%, so that on no particle did we measure both *A* and *B*. Then, GUP must be satisfied even in this case, and if *A* and *B* are incompatible $\sigma_A\sigma_B$ must be greater than zero.[4]

EXAMPLE 7.2.1

Let us verify GUP in the following case. Consider an ensemble of particles in state $|\uparrow_z\rangle$, and suppose we measure S_x on 50% of them, and S_y on the other 50%. Then GUP says that

$$\sigma_{S_x}{}^2\sigma_{S_y}{}^2 \geq \left(\frac{1}{2i} < i\hbar S_z >\right)^2 = \left(\frac{\hbar}{2} < S_z >\right)^2. \tag{7.2.25}$$

However, since all the particles are in state $|\uparrow_z\rangle$, $< S_z >= \hbar/2$, and therefore

$$\sigma_{S_x}{}^2\sigma_{S_y}{}^2 \geq \left(\frac{\hbar^2}{4}\right)^2. \tag{7.2.26}$$

Keeping in mind that standard deviation is always positive, we obtain $\sigma_{S_x}\sigma_{S_y} \geq \hbar^2/4$. Now let us measure σ_{S_x} and σ_{S_y} independently. As we know, $\sigma_{S_x}{}^2 =< S_x{}^2 > - < S_x >^2$. However, $< S_x{}^2 >= \hbar^2/4$, and $< S_x >^2 = 0$, so that $\sigma_{S_x} = \hbar/2$. An analogous computation shows that $\sigma_{S_y} = \hbar/2$ as well. Consequently, $\sigma_{S_x}\sigma_{S_y} = \hbar^2/4$, and GUP is satisfied.

7.3 More on GUP and Degeneracy

Standard quantum mechanics has a nice explanation for GUP and some of its surprising consequences. Although any Hermitian operator generates a basis in Hilbert space, two non-commuting Hermitian operators might share some eigenvectors but do not have a common basis. Suppose now that upon measuring observable *A* we obtain *a*;

---

[4] This point is emphasized in Margenau, H., (1963): 469-85.

EE guarantees that the system is now in the eigenstate $|\Psi_A\rangle$. But if $|\Psi_A\rangle$ is not an eigenstate of the operator representing observable *B*, EE prevents *B* from being sharp. Hence, an *A*-measurement return will not always correlate with the same *B*-measurement return. GUP gives us a precise relation expressing this lack of correlation.

GUP should be used cautiously. For example, if $\hat{A}$ and $\hat{B}$ do not commute, it does not follow that either σ_A or σ_B must be greater than zero. For, it might be the case that $\hat{A}$ and $\hat{B}$ do share an eigenvector χ which happens to be the vector state of the systems on which *A* and *B* are measured. Then, both observables are sharp and therefore $\sigma_A{}^2\sigma_B{}^2 = 0$. In general, we should note that there are two ways in which $[\hat{A},\hat{B}]$ may be different from zero. $[\hat{A},\hat{B}]$ may be equal to a non-zero, constant operator, in which case under no circumstances can $[\hat{A},\hat{B}]$ be equal to zero. For example, $[\hat{x},\hat{p}] = i\hbar$, and since $\hbar$ is a constat, under no circumstances can the commutator be zero. However, $[\hat{A},\hat{B}]$ may be equal to a non-constant operator that may be equal to zero in special circumstances, and when they occur the two observables may be sharp. For example, the operators for the components of orbital angular momentum *L* do not commute.[5] However, when the total angular momentum is zero, all the returns for angular momentum components are zero all the times, and therefore are variation free.

If two Hermitian operators $\hat{A}$ and $\hat{B}$ commute it need not be the case that every eigenvector of one is an eigenvector of the other because from the fact that they share a basis it does not follow that they share *all* of their eigenvectors. For example, in cases of degeneracy, $\hat{A}$ may have extra eigenvectors that $\hat{B}$ does not have. This has interesting

[5] In fact, $[\hat{L}_x,\hat{L}_y] = i\hbar\hat{L}_z$; obviously, if $L = 0$, then $L_z = 0$.

consequences. Suppose that a system is in state $|\Psi\rangle$, an eigenstate common to $\hat{A}$ and $\hat{B}$. Now we measure *A*, and as a return we get $\hat{A}$'s eigenvalue λ, which in standard quantum mechanics means that $|\Psi\rangle$ collapsed onto one of $\hat{A}$'s eigenstates associated with λ. Now if $\hat{A}$ has only one eigenstate associated with λ, all is well. For then $|\Psi\rangle$ *is* that eigenstate, and *A*'s measurement changed nothing: the system is still in state $|\Psi\rangle$ and *B* is still sharp. However, in case of degeneracy there is more than one eigenvector of $\hat{A}$ associated with the eigenvalue λ, and $|\Psi\rangle$ could collapse into any one of these three eigenvectors, the theoretical counterpart of the fact that the system has changed state. Of course, if we measure *A* again, we shall still get the same result as before. However, this need not be the same with respect to the other observable *B*. For it may turn out that not all the eigenstates of $\hat{A}$ are eigenstates of $\hat{B}$ as well. So, if the system ends up in an eigenstate of $\hat{A}$ but not of $\hat{B}$, EE requires that *B* will not be sharp, and therefore $\sigma_A{}^2\sigma_B{}^2 > 0$.[6]

EXAMPLE 7.3.1

Consider an ensemble of spin-half particles in state

$$|\Psi\rangle = \frac{1}{\sqrt{9}}\begin{pmatrix} 2-i \\ 2 \end{pmatrix}, \qquad (7.3.1)$$

[6] To simplify things one often talks about *ideal* measurements, namely measurements that minimize the disturbance of the physical properties of the system compatibly with the laws of quantum mechanics. So, for example, an eigenvector will collapse onto itself, in which case $\sigma_A{}^2\sigma_B{}^2 = 0$.

and suppose we try to measure both S_x and S_y. We know that $\left[S_x, S_y\right] = i\hbar S_z$, and that therefore the two observables are not compatible. By a simple calculation, we obtain

$$\sigma_{S_x}{}^2 \sigma_{S_y}{}^2 \geq \frac{\hbar^2}{4} < S_z >^2. \qquad (7.3.2)$$

Now, let us determine $< S_z >$. Remembering that $< S_z > = \left\langle \Psi \middle| S_z \Psi \right\rangle$,

$$< S_z > = \left\langle \begin{pmatrix} \frac{2-i}{\sqrt{9}} \\ \frac{2}{\sqrt{9}} \end{pmatrix} \middle| \begin{pmatrix} \frac{\hbar}{2} & 0 \\ 0 & -\frac{\hbar}{2} \end{pmatrix} \begin{pmatrix} \frac{2-i}{\sqrt{9}} \\ \frac{2}{\sqrt{9}} \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} \frac{2-i}{\sqrt{9}} \\ \frac{2}{\sqrt{9}} \end{pmatrix} \middle| \begin{pmatrix} \frac{\hbar(2-i)}{2\sqrt{9}} \\ -\frac{\hbar}{\sqrt{9}} \end{pmatrix} \right\rangle = \qquad (7.3.3)$$

$$\frac{2+i}{\sqrt{9}} \frac{\hbar(2-i)}{2\sqrt{9}} - \frac{2}{\sqrt{9}} \frac{\hbar}{\sqrt{9}} = \frac{\hbar}{18}.$$

Consequently,

$$\sigma_{S_x}{}^2 \sigma_{S_y}{}^2 \geq \frac{\hbar^2}{4} \frac{\hbar^2}{18^2}, \qquad (7.3.4)$$

that is,

$$\sigma_{S_x} \sigma_{S_y} \geq \frac{\hbar^2}{36}. \qquad (7.3.5)$$

Hence, repeated measurements on the identically prepared particles will not yield consistent results.

7.4 Two Versions of the Position/Momentum Uncertainty Principle

GUP entails the position/momentum uncertainty principle (HUP), which was formally proved by E. H. Kennard in 1927. As we saw, HUP is a purely statistical result, like GUP.[7] However, starting with Heisenberg himself, often the phrase "Uncertainty

---

[7] We use the phrase "Uncertainty Principle" or "Uncertainty Relation" because they are the most common. There is another relation between time and energy that is often

Relation" has been associated with a claim about individual systems, not merely ensembles. Suppose that we measure with great precision a particle's momentum, position, momentum again, position again and so on in very quick succession. If we compare two consecutive momentum returns, we discover that their differences vary unpredictably.[8] In short, precise position measurements are associated with unpredictable changes of momentum, and vice versa. So, it looks as if we cannot precisely experimentally know the simultaneous position and momentum of the particle. In orthodox quantum mechanics this can be explained by the fact that the position and momentum representations of the same quantum state are Fourier transforms of each other: if one peaks, the other spreads out. So, if we measure position and obtain the return value *a*, EE guarantees that the position wave function will spike at *a* and be zero everywhere else; in other words, it will become a Dirac delta function centered on *a*. Consequently, the momentum wave function will become totally flat, a superposition of all momentum eigenfunctions, telling us that the particle's momentum has been effectively randomized. If we then measure momentum very precisely, the exact converse will happen, and the particle's position will be randomized. This is often expressed more precisely in a relation reminiscent of HUP, namely,

$$\Delta_x \Delta_p \geq \frac{\hbar}{2}, \qquad (7.4.1)$$

---

referred to as "Time-Energy Uncertainty." However, this is really a misnomer even in an area in which terminology is often misleading; see appendix 6.

[8] The same is true of spin components or indeed of other incompatible observables.

where Δ_x and Δ_p are the uncertainties, understood as minimal degrees of imprecision, in knowledge of position and momentum. Let us call this ‘version’ of the uncertainty relations applied to individual systems “HUPI”. It is a strange principle, oddly in tension with the very orthodox quantum mechanics that embraces it. Obviously, it makes sense to talk about imprecision in knowledge only if there is something definite (the precise value of momentum, say) the knowledge is about. Hence, HUPI seems to entail that the particle has both definite position and momentum, although we cannot know both. However, EE plus collapse preclude this: if the particle is an eigenstate of position, it has to be in a superposition of eigenstates of momentum, and therefore it has no momentum. The epistemological and the ontological interpretation of HUPI are, at best, at odds.

When thinking about uncertainty relations, we must distinguish HUP from HUPI. HUP is a purely statistical claim not applicable to a single system that appears to entail nothing about the simultaneous measurement of position and momentum in a single system. For example, considered abstractly HUP is compatible with $\Delta_x\Delta_p = 0$ for some, though not all, of the identically prepared particles in an ensemble. Hence, in the absence of quantum mechanical principles tying the two, HUP is compatible with the negation of HUPI.

However, Jammer has claimed that in standard quantum mechanics there is a direct link between HUP and HUPI, if one makes relatively uncontroversial assumptions. Suppose that HUPI is false. Then, for some members of an ensemble it can be the case that position and momentum are simultaneously measured with greater precision than HUPI allows. Suppose now that we can filter out such systems thus constituting a new ensemble. If we assume that the measurement is immediately repeatable, the returns of position and momentum measurements on the new ensemble will contradict HUP (Jammer, M., (1974): 81). If Jammer is right, HUP entails HUPI.[9]

If it proved anything, Heisenberg's original 1927 paper did not prove HUP but HUPI.[10] Moreover, it attempted to obtain HUPI not from the mathematical formalism of quantum mechanics but from thought experiments, the most famous of which involves a causal story based on a measurement with a hypothetical γ-ray microscope. The basic idea is that a very precise determination of the position of a particle perturbs the particle's momentum uncontrollably. It is this unpredictability that is allegedly reflected in HUPI. Two points should be noted.

First, the γ-ray microscope story is compatible with the idea that the particle has definite position and momentum at all times. Differently put, it is a story about our principled inability to obtain knowledge of simultaneous position and momentum, not about their existence. To conclude that a particle does not have simultaneous position and momentum something else is needed, such as the Fourier transforms account of HUPI we considered above plus EE and the claim that standard quantum mechanics is complete.

Second, as a physical argument Heisenberg's story is far from convincing. In his 1927 paper, probably in order to attack Schrödinger's idea that a quantum particle is a

---

[9] Heisenberg too thought that HUP and HUPI are closely linked. Indeed, in his often quoted *The Physical Principles of the Quantum Theory* (1930), he treated HUPI and HUP as amounting to the same thing.

[10] For a dramatic account of the genesis of HUPI and the role Einstein played in it, see Heisenberg, W., (1971b).

wave packet (an idea that Schrödinger had actually already abandoned when Heisenberg wrote his paper), Heisenberg treated photons and electrons as regular particles, appealing to the Compton recoil to conclude that at the time of collision a photon transfers an uncontrollable amount of momentum to the electron.[11] But aside from the fact that Compton's account uses $E = hf$, where *f* is a wave frequency, thus making wave concepts unavoidable, the Compton effect is a deterministic affair: if we know the energy and momentum of the colliding photon, we can determine the energy and momentum of the electron precisely, as in classical mechanics. The 1930 version of the γ-ray microscope example avoids such problem by appealing to optical diffraction (a wave phenomenon) to the effect that the resolving power of a microscope is given by

$$\frac{\lambda}{2\sin\varepsilon}, \qquad (7.4.2)$$

where λ is the wavelength of the light used and 2ε is the angle subtended by the lens' diameter at the position of the observed object (Heisenberg, W., (1930): 21-2). But it turns out that (7.4.2) does not hold true of new types of imaging techniques like the scanning tunneling microscope which achieve resolutions about $\lambda/50$, much greater than the one allowed by (7.4.2).[12]

Of course, a causal account of HUPI is satisfying because it allows us to see how uncertainty comes about physically. But this does not show that Heisenberg's causal story, or any other causal story, is correct. This is especially true if one tries to generalize the causal account to HUP, which applies even when no casual interaction has taken

---

[11] For an account of the conceptual issues surrounding Heisenberg's paper, see Beller, M., (1999): ch. 4; 6.

[12] For details, see Afriat, A. and Selleri, F., (1999): 54-8.

place, for example when the two incompatible observables are measured on different non-interacting systems in an ensemble.

7.5 Heisenberg, HUPI, and the Past

In 1930, in line with the microscope thought-experiment, Heisenberg commented on HUPI:

> (E)very experiment destroys some of the knowledge of the system which was obtained by previous experiments. This formulation makes it clear that the uncertainty relation does not refer to the past; if the velocity of the electron is at first known and the position then exactly measured, the position for times previous to the measurement may be calculated. Then, for these past times $\Delta p \Delta q$ is smaller than the usual limiting value, but this knowledge of the past is of purely speculative character, since it can never (because of the unknown change in momentum caused by the position measurement) be used as an initial condition in any calculation of the future progress of the electron and thus cannot be subjected to experimental verification. It is a matter of personal belief whether such a calculation concerning the past history of the electron can be ascribed any physical reality or not (Heisenberg, W., (1930): 20).

Here Heisenberg makes three claims. First that HUPI places principled limits to our knowledge; second, that we can obtain exact knowledge about past positions and momentums of a particle; third, that this knowledge is “speculative”. Let us look at them.

The epistemological account of HUPI is certainly in line with Heisenberg's microscope illustration but, as we saw, it fails to agree with the ontological interpretation usually put forth by Copenhagen and orthodox theorists.

Heisenberg's second claim can be illustrated by following an argument put forth by Afriat and Selleri. Consider a one-dimensional electron initially in state

$$\psi(x) = \psi_0 e^{\frac{ipx}{\hbar}}, \tag{7.5.1}$$

which in momentum representation is the eigenvector with p as the momentum's value. Hence, at time t_0 we know p exactly and $\Delta_p = 0$. Of course, the position wave function corresponding to (7.5.1) is flat and consequently it tells us nothing about the electron's location. We could measure the electron's position, but the very act of measurement would alter the value of momentum unpredictably. However, because TDSE is time-reversible, that is, it evolutionary determinism extends not only to the future but also to the past, (15.4.2) guarantees that at time $t < t_0$ the electron's momentum is exactly knowable, and therefore at t we have $\Delta_p = 0$. Consequently, we can obtain the exact velocity of the electron at any time $t < t_0$ as

$$v = \frac{p}{m}.\text{[13]} \tag{7.5.2}$$

Suppose now that at t_0 we measure the electron's position exactly, or at most with finite imprecision Δ_x. Then we can retrodict the electron's position as

$$x(t) = x_0 \pm \Delta_x - v(t_0 - t). \tag{7.5.3}$$

[13] We should remember that momentum is mass times velocity: $p = mv$.

Of course (7.5.3) contains the error Δ_x, the legacy of the original position measurement; however, Δ_x is finite. So, at *t* we know the electron's momentum exactly ($\Delta_p = 0$) and the electron's position with a finite degree of error. Hence, in the past $\Delta_x \Delta_p = 0$.

Our ability precisely to determine simultaneous position and momentum in the past seems to provide a further forceful argument against any ontological interpretation of HUPI. For if we know the values of position and momentum at *t*, the electron did have definite position and momentum then, and if it had them then it would seem to follow that the electron would also have them at later times, including now (Afriat, A., and Selleri, F., (1999): 55-7). However, Heisenberg argued that since the values of position and momentum at *t* could not be experimentally tested, one could not cogently argue that the electron really had both position and momentum. In effect, he adopted the radical empiricist view that symbols in a language have meaning only to the extent that they can be made part of sentences whose truth can be empirically tested, directly or indirectly. So, 'Mount Everest is 5000 meters high' is a perfectly meaningful sentence since we have both direct and indirect measurements showing that it is false. However, Heisenberg claimed, the statement that there exists a parallel world with no connection to ours is meaningless because it cannot possibly have any experimental consequence. Hence, any reference to the position or velocity of a particle with an accuracy exceeding that allowed by HUPI is "just as meaningless as the use of words whose sense is not defined" because it cannot be made part of any verifiable statement (Heisenberg, W., (1930): 15). Note that Heisenberg did not require immediate empirical testability; in other words he did not adopt the view 'if it cannot be measured by direct experiment it does not exist.' However, he did require at least the possibility of mediate experimental verification such as the capacity to derive conclusions which can in principle be directly tested. Even so, Heisenberg's reliance on a verificationist theory of meaning makes his argument highly questionable.

One can conclude that the ontological interpretation of HUPI is far from inevitable. True, standard quantum mechanics is bound to accept it, but only because of two strong assumptions, the state completeness principle and EE, both of which can, and have been, questioned, as we shall see later.

7.6 The Classical Limit

The natural domain of quantum mechanics, the domain for which it was designed, is that of the atomic and, eventually, the subatomic world, where the energies involved are immensely smaller than in the macroscopic world of our experience. So, what happens if we introduce the high energy figures typical of the macroscopic world into standard quantum mechanical formulas? The question is a natural one, and an answer was first provided by Bohr's Correspondence Rule, which, roughly put, states that in the classical limit the predictions of quantum mechanics should closely approximate those of classical physics. It was a powerful methodological principle employed systematically by Bohr from around 1918 in his own early versions of quantum mechanics and even later by him and others in modern quantum mechanics. Indeed, its heuristic role in the development of the new theory can be hardly overestimated.

To get a sense of what happens when energies become large, consider the classical limit of the quantum harmonic oscillator. Let us concentrate on the fractional difference F in energy between two adjacent levels:

$$F = \frac{E_{n+1} - E_n}{E_n}. \qquad (7.6.1)$$

As the difference between any two energy levels is always $\hbar\omega$,

$$F = \frac{\hbar\omega}{\left(\frac{1}{2}+n\right)\hbar\omega} = \frac{1}{\frac{1}{2}+n}. \tag{7.6.2}$$

Hence, as *n* grows larger and larger, *F* approaches zero. In other words, as energy increases, the difference between two contiguous energy levels E_{n+1}, E_n becomes infinitesimal when compared to E_n. So, for very large quantum numbers, the energy gaps due to quantization are so small in proportion to the total energy that energy will appear continuously distributed.

There are some serious caveats. Even a high-energy quantum harmonic oscillator is not a classical oscillator: it can still tunnel, its density function still has nodes, and its energy remains quantized. In addition, not all quantum harmonic oscillators will become empirically indistinguishable from a classical one once the energy levels becomes sufficiently large since the superpositions of harmonic oscillator eigenstates can grossly violate classical laws even for very large energies.[14]

7.7 The Ehrenfest Theorem

In 1927, Paul Ehrenfest proved an important theorem showing a connection between quantum and classical mechanics. Here we merely state it, providing the proof in appendix 4. In its most general formulation the Ehrenfest Theorem is

$$\frac{\partial}{\partial t} < O >= \frac{i}{\hbar}\left\langle\left[\hat{H},\hat{O}\right]\right\rangle + \left\langle\frac{\partial\hat{O}}{\partial t}\right\rangle, \tag{7.7.1}$$

[14] Here we follow Home, D., (1997): 146-47.

where O is an observable, $\hat{O}$ its quantum mechanical operator and H the Hamiltonian.[15] The theorem is more easily understood if we assume that the second term on the right of (7.7.1) is zero, that is, if we assume that the operator $\hat{O}$ does not explicitly depend on time (the time variable does not appear in the expression for $\hat{O}$).[16] Then (7.7.1) says that the rate of change of the expectation value of O is proportional to the commutator of the Hamiltonian with the operator $\hat{O}$.

To see how the Ehrenfest Theorem works in detail, let us apply it to the conservation of energy. The general formulation of the Ehrenfest Theorem tells us that

---

[15] $\frac{\partial}{\partial t}$ is the partial derivative with respect to time. To see what it does, consider a straight line x with hot and cold spots on it. Suppose now that the temperature on x varies both with respect to position and with respect to time: different positions on x have different temperatures, and the same position has different temperatures at different times. Mathematically, this is expressed by saying that temperature T is a function of both position x and time t: $T = f(x,t)$. Then, $\frac{\partial}{\partial t} f(x,t)$ tells us how quickly temperature changes (its rate of change) at a given position (that is, time changes but position does not). By contrast, $\frac{\partial}{\partial x} f(x,t)$, the partial derivative with respect to space, tells us the rate of change of temperature as position varies but time does not. For an introduction to derivatives, see appendix one.

[16] For an operator explicitly dependent on time, consider the potential energy of a harmonic oscillator in which the spring's elastic powers changes with time. Then the potential energy is an explicit function of time, and so is the total energy.

when $\hat{O}$ and $\hat{H}$ commute $<O>$ is constant (its expectation value is conserved) as long as $\hat{O}$ does not explicitly depend on time. However, any operator trivially commutes with itself. Therefore, energy is conserved, that is

$$\frac{d}{dt}<H>=0 \tag{7.7.2}$$

for any state as long as $\hat{H}$ does not explicitly depend on time.

As another example, consider the momentum operator $\hat{p}=\frac{\hbar}{i}\frac{\partial}{\partial x}$, so that

$$\frac{\partial}{\partial t}<p>=\frac{i}{\hbar}\left\langle\left[\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+V\right),\frac{\hbar}{i}\frac{\partial}{\partial x}\right]\right\rangle+\left\langle\frac{\partial}{\partial t}\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)\right\rangle. \tag{7.7.3}$$

Now, the last term is zero because $\hbar\partial/i\partial x$ is time independent (there is no time variable t in the expression). Moreover, since the commutator in (7.7.3) is equal to $-\frac{\hbar\partial V}{i\partial x}$, we obtain

$$\frac{\partial}{\partial t}<p>=-\frac{i}{\hbar}\frac{\hbar}{i}\left\langle\frac{\partial V}{\partial x}\right\rangle=-\left\langle\frac{\partial V}{\partial x}\right\rangle, \tag{7.7.4}$$

which, in terms of its form, looks like Newton's Equation of Motion applied to expectation values.[17] This is suggestive, for if quantum-mechanical expectation values of

---

[17] As we know from (1.1.8), in classical mechanics, $\frac{d}{dt}p=F$; in addition, when the system under study is conservative (a system in which energy is conserved), force is related to potential energy by the relation $F=-\frac{\partial V}{\partial x}$: force is equal to the negative of the rate of change of potential energy with respect to space. For example, imagine a particle

physical quantities can be identified with the classical values of those quantities, then Newton's Equation of Motion could be "derived" from quantum mechanics via the Ehrenfest Theorem. This identification is somewhat plausible, since usually macroscopic particles are greatly localized and therefore their position wave functions are very narrow in comparison with the measuring apparatus; this is why, experimental error aside, the standard deviation of measurement returns for an ensemble of macroscopic systems is zero, and therefore the return value of any observable O is $\langle O \rangle$.[18] Still, we must remember that, for example, any quantum harmonic oscillator, no matter the energy level, can still tunnel and cannot be at the nodes.

At this point, perhaps one should hold that our understanding of the world requires both quantum and classical mechanics because there is a real gap between the microscopic and the macroscopic worlds.[19] The problem, however, is that physically

---

suspended from the ceiling by a wire. If we cut the wire, the particle begins to fall under the effect of the force of gravity $F = -mg$, where g is the acceleration of gravity (the negative sign simply indicates that g, and therefore F are directed towards the floor). The potential energy of the particle is $V = mgx$, and $F = -\frac{\partial}{\partial x} mgx$.

[18] Things look even smoother for Bohm's theory, as we shall see.

[19] Notice that, as we shall see, this is not a restatement of Bohr's views. Bohr did not hold that there is an ontological gap between the quantum world and the macroscopic world. The gap is epistemological: we are macroscopic creatures bound to try to understand the quantum world in terms of concepts resulting from our macroscopic experience.

there seems to be no such sharp division. For example, the first experiments with Stern-Gerlach devices were performed on atoms rather than on, say electrons. But many atoms are big enough that their positions and velocities, once spin (a quantum phenomenon, to be sure) is taken into account, can be determined in quasi-classical terms. So, although an electron is certainly not a classical object, in some respects some atoms are. Conversely, experiments conducted on superconducting quantum interference devices (SQUIDS) have shown that, under special circumstances, macroscopic objects do tunnel, and therefore behave according to quantum mechanical laws (Greenstein, G., and Zajonc, A., (1997): 171-77). More recently, a pair of cesium gas clouds containing 10^{12} atoms each has been entangled, showing that in some respects they could behave like two macroscopic atoms (Physics Update, in *Physics Today*, (2001): 9). In sum, the claim that ‘being macroscopic’ and ‘obeying classical physics’ are extensionally equivalent (in the sense that they identify the very same events) or even merely empirically equivalent (in the sense that they identify the very same events as far as we can experimentally see) is, strictly speaking, false. Moreover, it is not a consequence of standard quantum mechanics, although it is certainly true that the macroscopic systems we encounter in everyday life outside physics laboratories seem to obey classical laws.

Exercises

Exercise 7.1

1. Can the square of the average be greater than the average of the squares? [Hint: look at (7.1.3)].

2. A cart contains 4 one-kilogram boxes, 1 three-kilogram box, and 1 four-kilogram box. Determine the weight average and the standard deviation.

Exercise 7.2

1. Prove that $\hat{A} - < A >$ is Hermitian.
2. Prove HUP, the position-momentum Uncertainty Principle. [Hint: we know that $[\hat{x}, \hat{p}] = i\hbar$. Plug this into GUP. Remember that standard deviation is always positive.]

Exercise 7.3

1. Consider an ensemble of spin-half particles in state $|\Psi\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. What happens if we measure both S_z and S_y?
2. Consider an ensemble of spin-half particles in state $|\Psi\rangle = \frac{1}{\sqrt{15}} \begin{pmatrix} 3+i \\ 2-i \end{pmatrix}$. What happens if we measure both S_z and S_x?
3. True or false: if $\hat{A}$ and $\hat{B}$ have just *some* common eigenvector then they commute.
4. True or false: if $\hat{A}$ and $\hat{B}$ commute, then GUP tells us only that $\sigma_A{}^2 \sigma_B{}^2 \geq 0$, namely, that it *may* happen that $\sigma_A{}^2 \sigma_B{}^2 = 0$, not that it must happen.
5. True or false: if two operators commute, then they share all of their eigenvectors.

Answers to the Exercises

Exercise 7.1

1. Since σ^2 is a square and a real number it cannot be negative. Hence, the answer is "no."

2. $<w>=\frac{11}{6}$; $<w^2>=\frac{29}{6}$. Hence, $\sigma=\sqrt{\frac{29}{6}-\frac{121}{36}}=1.21.$

Exercise 7.2

1. The result follows immediately from the fact that $\hat{A}$ is Hermitian and $<A>$ a real number.

2. By plugging $[\hat{x},\hat{p}]=i\hbar$ into the general formula for the Uncertainty Principle, we get $\sigma_x{}^2\sigma_p{}^2\geq\left(\frac{1}{2i}i\hbar\right)^2=\left(\frac{\hbar}{2}\right)^2$, that is, $\sigma_x\sigma_p\geq\pm\left(\frac{\hbar}{2}\right)$. Since a standard deviation is always positive, we can discard the negative root and obtain the original Heisenberg Principle, $\sigma_x\sigma_p\geq\frac{\hbar}{2}$.

3. The mistake is that A and B together do not constitute an ensemble of identically prepared systems, and GUP applies only to such ensembles.

Exercise 7.3

1. The Uncertainty Principle says that $\sigma_{S_z}{}^2\sigma_{S_y}{}^2\geq\left[\frac{1}{2i}<\left[\hat{S}_z,\hat{S}_y\right]>\right]^2$, and $\left[\hat{S}_z,\hat{S}_y\right]=-i\hbar S_x$.

Hence, $\sigma_{S_z}{}^2\sigma_{S_y}{}^2\geq\frac{\hbar^2}{4}<S_x>^2$. Now, $\hat{S}_x=\frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$, and therefore

$$< S_x >= \left\langle \begin{pmatrix}1\\0\end{pmatrix} \middle| \begin{pmatrix}0 & \frac{\hbar}{2}\\ \frac{\hbar}{2} & 0\end{pmatrix} \begin{pmatrix}1\\0\end{pmatrix} \right\rangle = \left\langle \begin{pmatrix}1\\0\end{pmatrix} \middle| \begin{pmatrix}0\\ \frac{\hbar}{2}\end{pmatrix} \right\rangle = 1 + \frac{\hbar}{2}.$$ Hence, $< S_x >^2 = \left(1+\frac{\hbar}{2}\right)^2$. So,

$$\sigma_{S_z}\sigma_{S_y} \ge \frac{\hbar}{2}\left(1+\frac{\hbar}{2}\right).$$

2. The Uncertainty Principle says that $\sigma_{S_z}{}^2\sigma_{S_x}{}^2 \ge \left[\frac{1}{2i} < \left[\hat{S}_z, \hat{S}_x\right] > \right]^2$, and $\left[\hat{S}_z, \hat{S}_x\right] = i\hbar S_y$.

Hence, $\sigma_{S_z}{}^2\sigma_{Sx}{}^2 \ge \frac{\hbar^2}{4} < S_y >^2$. Now, $\hat{S}_y = \frac{\hbar}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}$, and therefore

$$< S_y >= \left\langle \frac{1}{\sqrt{15}}\begin{pmatrix}3+i\\2-i\end{pmatrix} \middle| \begin{pmatrix}0 & -i\frac{\hbar}{2}\\ i\frac{\hbar}{2} & 0\end{pmatrix} \frac{1}{\sqrt{15}}\begin{pmatrix}3+i\\2-i\end{pmatrix} \right\rangle = \left\langle \frac{1}{\sqrt{15}}\begin{pmatrix}3+i\\2-i\end{pmatrix} \middle| \begin{pmatrix}\frac{-i\hbar(2-i)}{2\sqrt{15}}\\ \frac{i\hbar(3+i)}{2\sqrt{15}}\end{pmatrix} \right\rangle =$$

$$= \frac{-6i\hbar - 3\hbar - 2\hbar + i\hbar + 6i\hbar - 2\hbar - 3\hbar - i\hbar}{30} = -\frac{\hbar}{3}.$$ So, $\sigma_{S_z}{}^2\sigma_{Sx}{}^2 \ge \frac{\hbar^2}{4}\frac{\hbar^2}{9}$, and therefore

$$\sigma_{S_z}\sigma_{Sx} \ge \frac{\hbar^2}{6}.$$

3. False. Two Hermitian operators $\hat{A}$ and $\hat{B}$ commute if and only if they have a complete set of common eigenvectors. For example, when discussing angular momentum we noticed that $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$ do not commute. However, when the angular momentum L is zero any measurement of $\hat{L}_x, \hat{L}_y$, and $\hat{L}_z$ will certainly return zero. Hence, the vector state in which $L = 0$ is a shared eigenvector of $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$.

4. GUP tells us only that $\sigma_A{}^2\sigma_B{}^2 \ge 0$, so that it may, although it need not, happen that $\sigma_A{}^2\sigma_B{}^2 = 0$.

5. False.