Heisenberg, Matrix Mechanics, and the Uncertainty Principle

S Lakshmibala

Werner Heisenberg was one of the key players in the development of quantum mechanics. Besides enunciating the famous Uncertainty Principle, he was also the principal architect of Matrix Mechanics, one of the two standard formulations of quantum mechanics.

S Lakshmibala is a theoretical physicist at the Department of Physics, IIT Madras. Her current research interests are in dynamical systems and quantum information.

Genesis

Albert Einstein is said to have remarked that intuition is nothing but the collection of prejudices one forms up to the age of eighteen. What he meant was that our so-called physical intuition is little more than a rough feel for the way the physical world around us behaves on everyday scales. At very small or very large scales of length, mass and time, however, this intuition is as likely as not to be misleading or even wrong. This is indeed the main message of the revolutionary advances in the physical sciences in the 20th century. The discovery of quantum mechanics is the centre-piece of that revolution.

By the end of the 19th century, the edifice of physics seemed to be on a firm foundation. The basic equations of classical mechanics and electromagnetism were in place. But, at the same time, the atomic or discrete nature of matter (as well as radiation) was slowly but surely becoming clear. This was accompanied by a gradual realization that classical physics was seriously inadequate in the realm of atoms, and that atomic phenomena were perhaps governed by a 'quantum' theory waiting to be discovered and tested. Originating in the suggestions of Max Planck and Einstein, this theory (now called the 'old quantum theory') was developed by Niels Bohr and

Keywords
Uncertainty principle, matrix mechanics, quantum measurement.

Arnold Sommerfeld, among others, and showed promise initially. But many of its features were quite clearly *ad hoc*, and it was evident that a deeper theory lay hidden beneath.

These investigations climaxed with the advent of quantum mechanics in the 1920s. Under the leadership of a group of remarkably young physicists, the intricacies of the quantum world were unravelled – each step leaving behind an audience stunned by Nature's ways. High drama was involved, as also a clash of personalities primarily from Göttingen, Copenhagen and Berlin in Europe, who were dogmatic in their views on what could be the correct theory. The lessons learnt were many. The workings of the quantum world defied explanations (however elaborate) based on intuition. A much more precise language was required – the unambiguous language of mathematical equations, matrix algebra, differential equations and group theory.

The quantum world required the unambiguous language of mathematical equations, matrix algebra, differential equations and group theory.

It was crucial, however, to identify the fundamental physical quantities whose properties, static and dynamic, constituted the study of quantum physics. And there were really no significant pointers in this regard from classical physics (*Box* 1). The approach to the quantum world was therefore based, to a large extent, on what each researcher believed should be the guiding principle in investigating the unknown.

Werner Heisenberg, one of the founding fathers of quantum mechanics (as opposed to the 'old' quantum theory)

Box 1. Planck's Constant, h

Once the standard framework of quantum mechanics was well established, it became clear that classical physics can, in general, be recovered from quantum mechanics as a limiting case: one must formally let a certain constant h tend to zero in all expressions. This constant, named after Planck, is the trademark of quantum mechanics. It is one of the three fundamental constants of nature, along with c (the speed of light in vacuum) and G (Newton's gravitational constant).

Box 2. Early Reactions

Einstein, for instance, was convinced that "··· it is the theory which decides what we can observe", and not the other way around. Erwin Schrödinger, the creator of wave mechanics, an alternative formulation of quantum mechanics, opined that he "··· felt discouraged, not to say repelled, ··· by the lack of visualizability" in matrix mechanics. (Schrödinger's formalism deals with the nature and dynamics of a fundamental object called the wave function associated with a physical system.) Nevertheless, Schrödinger himself went ahead soon thereafter to establish the equivalence between wave mechanics and matrix mechanics, in 1926.

sought to establish a basis for the theoretical aspects of the quantum mechanics of a system. *This was founded exclusively upon relationships between quantities pertaining to that system which are, in principle, observable.* By 'observables' are meant quantities which are experimentally measurable, either directly or indirectly. This approach led to a successful formulation of quantum mechanics based on the theory of matrices (matrix mechanics), though it initially met with vehement objections (*Box* 2).

As a first step, Heisenberg had to identify the relevant observables. In atomic physics, observational data related to atomic transitions arising from interactions of the atoms with light quanta. Heisenberg therefore argued that these transition-related quantities were the basic relevant objects. A detailed investigation on these lines led him to propose the first coherent mathematical structure for the quantum theory of atoms, in 1925. Together with Max Born and Pascual Jordan, who recognized that these quantities obeyed rules prescribed by matrix algebra, Heisenberg developed the essentials of matrix mechanics later that year.

Together with Max Born and Pascual Jordan, Heisenberg developed the essentials of matrix mechanics in 1925.

Matrix Mechanics

Generalizing the lesson learnt from the transition-related quantities mentioned above, the matrix formulation was built on the premise that *all physical observables must*

be represented by matrices. The set of eigenvalues of the matrix representing an observable is the set of all possible values that could arise as outcomes of experiments conducted on a system to measure the observable. Since the outcome of an experiment to measure a real observable must be a real number, Hermitian matrices would represent such observables (as their eigenvalues are real). If the result of a measurement is a certain eigenvalue, the corresponding eigenvector represents the state of the system immediately after the measurement. The act of measurement is taken to 'collapse' the state of the system to that eigenvector (or eigenstate). Examples of such eigenstates are those of position, momentum, energy, etc. It may be possible sometimes to make simultaneous measurements of two or more observables. In that case the system will collapse to a common eigenstate of these observables right after the measurement.

If the result of a measurement is a certain eigenvalue, the act of measurement is taken to 'collapse' the state of the system to that eigenvector

It was soon realized that the matrices representing the observables are really operators in an appropriate linear vector space, and that their eigenstates are column vectors in that space. This space could even be infinite-dimensional (*Box* 3).

Box 3. Linear Vector Space

The space in which the state of a quantum system 'lives' is not to be confused with the three-dimensional physical space in which we live. To get the idea right, consider the simple example of a point particle that is free to move on a line. An observable in this case could be the position of the particle, represented by the matrix X. Since the particle could be anywhere on the line, the possible outcome of a measurement of its position could be any one of an infinite set of eigenvalues of X, denoted by x. Thus X must be an infinite-dimensional matrix, and hence so is the corresponding linear vector space. Thus even one-dimensional motion could have an infinite-dimensional linear vector space associated with it!

In fact, this example is not as trivial as it appears to be. The set of eigenvalues x is a *continuous* infinity of values, in contrast to a discrete infinite set of values such as $1, 2, 3, 4, \cdots$. This poses certain technical problems: for instance, we must extend the concept of matrices to objects with a *continuously infinite* number of rows and columns: namely, *operators*. Operators, functions spaces, and so on, thus make their appearance in quantum mechanics quite naturally.

From matrix theory we know that eigenvectors corresponding to distinct eigenvalues of a Hermitian matrix are orthogonal to each other.

Further, from matrix theory we know that eigenvectors corresponding to distinct eigenvalues of a Hermitian matrix are orthogonal to each other. In this sense, they are analogous to the x y z axes of the Cartesian coordinate system – except that now there could even be an infinite number of distinct eigenvalues, and hence as many mutually orthogonal eigenvectors 'pointing' along different independent directions in the linear vector space. Again, just as we have unit vectors $\hat{e}_x$ $\hat{e}_y$ $\hat{e}_z$ along the Cartesian axes, we can normalize each eigenvector to have unit magnitude.

From $\hat{e}_x$ and $\hat{e}_y$ we can form the linear combination $\vec{r} = a\,\hat{e}_x + b\,\hat{e}_y$, where a and b are any two real numbers. In much the same way, linear superpositions of the normalized independent eigenvectors can be formed, with coefficients that are complex numbers. It is important to note that, in general, such superpositions are *not* eigenstates of the matrix concerned.

We now face an interesting situation. Recall that, by performing an appropriate measurement on the system, we know the state of the system just after the measurement. Was this the state of the system before the measurement? Not necessarily! For, prior to the measurement, the system could have been in a linear superposition of different eigenstates, with unknown (and unguessable) coefficients. It is like saying that in a coin toss experiment whose outcome is a 'head' the coin could have been in a state which was a combination of head and tail before it was tossed! Of course, this would never be the case for actual coins, governed as they are by the laws of classical physics.

The state of the system just after the measurement is not necessarily the state of the system before the measurement.

But then, what was the precise state of the quantum system before the measurement? The answer is: we *cannot* know. The Copenhagen interpretation is concerned only with *outcomes* of experiments. Deep philosophical questions, peculiar to quantum mechanics, now arise (*Box* 4).

Box 4. Is Quantum Mechanics Complete?

If we can never know the pre-measurement state of a system, is not the theory inadequate, or at least incomplete? For, after all, the system surely has an existence of its own, independent of the act of measurement! (This question is also applicable to wave mechanics, for it too cannot predict the pre-measurement state.) Numerous proposals, including a variety of so-called *hidden variable theories*, have been made to overcome this inadequacy, but none of these is fully satisfactory. The last word has probably not been said yet in this regard.

We will not digress into these here.

We know that, to obtain the average value of any observable, a large number of trials have to be conducted, i.e., repeated measurements have to be made. But, for a quantum system, a single measurement of any observable A yields one of the eigenvalues of A as the outcome, and collapses the state of the system to the corresponding eigenstate. Subsequent measurements made immediately thereafter would continue to yield the same eigenvalue. Does that mean that a single measurement is enough? No! We have tampered with the pre-measurement state of the system and collapsed it during the first measurement. Therefore the post-measurement state is not the same as the initial state, and so it would not make sense to repeat measurements on a single system. The correct thing to do would be to prepare a collection of a very large number of identical copies of the system (an *ensemble*), and conduct a single trial on each copy. The arithmetic mean of all the results thus obtained is the average value we seek, denoted by $\langle A \rangle$. This is in sharp contrast to what we do in usual experiments on a classical system.

So, can we now claim that ensemble averages should give us precise answers for *all* the observables of a system, measured simultaneously? (We will allow for instrumental precision and errors of observation.) The answer is an emphatic 'No!' It turns out that certain pairs of ob-

Certain pairs of observables simply cannot be simultaneously measured to an arbitrarily high level of precision. In general, two matrices A and B, unlike mere numbers, do not commute with each other.

servables simply *cannot* be simultaneously measured to an arbitrarily high level of precision. The source of the problem is neither observational error nor the least count of the instruments used, but quantum mechanics itself! It stems from the fact that, in general, two matrices A and B, unlike mere numbers, do not *commute* with each other: the *commutator* $AB - BA \equiv [A, B] \neq 0$. This fact has far-reaching implications. Chief among them is the uncertainty relation due to Heisenberg, subsequently called the *Uncertainty Principle* (apparently for the first time by A Eddington in his Gifford Lectures in 1928), because of its profound consequences.

The Uncertainty Principle

The Uncertainty Principle was discovered by Heisenberg when he attempted to understand the so-called 'Transformation Theory' proposed by P A M Dirac and Jordan. This theory gives a cogent picture of quantum mechanics using linear vector spaces. It clarifies the roles of both wave mechanics and matrix mechanics, and is essentially the modern formulation of quantum mechanics.

Heisenberg realised that genuine, intrinsic, uncertainties or imprecisions cropped up in the simultaneous measurement of the position and linear momentum of a moving particle, and also in the simultaneous measurement of energy and time.

Heisenberg investigated the question of measurement in this framework. He realised that genuine, intrinsic, uncertainties or imprecisions cropped up in the simultaneous measurement of the position and linear momentum of a moving particle, and also in the simultaneous measurement of energy and time. He attempted to explain this novel feature through a *gedanken* or thought experiment, which uses a hypothetical gamma-ray microscope to observe electrons. His original argument, however, is not part of our current understanding of the actual Uncertainty Principle, for it treats interactions between quantum objects somewhat unrealistically, analogous to mechanical collisions of classical particles. Nevertheless, it was Heisenberg's seminal paper in 1927, written after serious discussions with Bohr and Wolfgang Pauli, that

Box 5. Infinite-dimensional Matrices

For any two $(n \times n)$ matrices R and S, Trace $(\mathsf{R}\,\mathsf{S})$ = Trace $(\mathsf{S}\,\mathsf{R})$. This property (called the 'cyclic invariance of the trace') need not hold good for infinite-dimensional matrices like X and P_x. Taking the trace on both sides of equation(1), Trace $(\mathsf{X}\,\mathsf{P}_x)$ − Trace $(\mathsf{P}_x\,\mathsf{X})$ = $i\hbar$ Trace $(\mathbb{I})$. But the RHS does not vanish in fact, it is infinite in magnitude!

disclosed to the world the existence of the quantum uncertainty relations.

The first mathematically exact derivation of the Uncertainty Principle for position and momentum seems to have been given in 1927 by E H Kennard. It hinges crucially on the fact that

$$[\mathsf{X} \quad \mathsf{P}_x] = i\hbar\mathbb{I}, \tag{1}$$

where $\hbar \equiv h/(2\pi)$ and $i \equiv \sqrt{-1}$. The x-components of the position and momentum of a particle are eigenvalues of the operators X and P_x respectively. $\mathbb{I}$ is the identity operator (*Box* 5). The commutator algebra in (1) is associated with a rich underlying group-theoretic structure. This has wide applications, most notably in quantum optics.

Kennard's derivation pertains to the variances in the simultaneous measurements of the observables X and P_x. These variances are defined as

$$(\Delta\mathsf{X})^2 \equiv \left\langle (\mathsf{X} - \langle\mathsf{X}\rangle)^2 \right\rangle \quad \text{and} \quad (\Delta\mathsf{P}_x)^2 \equiv \left\langle (\mathsf{P}_x - \langle\mathsf{P}_x\rangle)^2 \right\rangle \tag{2}$$

Suppose these measurements are made exactly once on each copy of the system in an ensemble. Kennard showed that it follows from (1) that the product $\Delta\mathsf{X}\ \Delta\mathsf{P}_x$ satisfies

$$\Delta\mathsf{X}\ \Delta\mathsf{P}_x \geq \frac{\hbar}{2} \tag{3}$$

The commutator algebra in equation (1) is associated with a rich underlying group-theoretic structure. This has wide applications, most notably in quantum optics.

Box 6. Variance and Higher Moments

The variance is a measure of the departure from the average value of the outcomes in a series of measurements. Its positive square root (ΔX or ΔP_x in our case) gives the spread of the outcomes about the average. Clearly, if we conduct N trials in which we measure an observable A and the outcome is a in *each* of them, the average value is just a, and $\Delta A = 0$. This happens if the system is in an eigenstate of A. Now $(\Delta A)^2$ can also be written as $\langle A^2 \rangle - \langle A \rangle^2$. Higher-order counterparts of the variance involve the moments $\langle A^k \rangle$ where $k > 2$. Generalizations of equation (3) have been derived subsequently for these quantities. We note that any distribution of outcomes is completely known *only* if, apart from the mean and variance, all non-vanishing higher moments are also known.

for every possible normalized state of the particle. This is the position-momentum uncertainty relationship (*Box* 6). Analogous relations hold good for the other two Cartesian components of the position and momentum of the particle.

The actual value of the uncertainty product (the LHS in (3)) depends on the state of the system. Minimum uncertainty states (states for which $\Delta X \, \Delta P_x = \hbar/2$) are especially interesting, because they are closest to classical states. (Recall that the classical limit corresponds to $\hbar \to 0$.) A 'coherent state' is an example of a minimum uncertainty state. Ideal single-mode laser light provides a physical realization of such a state.

The Kennard inequality was generalized in 1929 by H P Robertson to any pair of observables A and B of a system. He showed that

$$\Delta A \, \Delta B \geq \frac{1}{2} \left| \langle [A, B] \rangle \right| \qquad (4)$$

A 'coherent state' is an example of a minimum uncertainty state. Ideal single-mode laser light provides a physical realization of such a state.

for any state of the system. The relation was generalized further by Schrödinger in 1930 to

$$(\Delta A)^2 (\Delta B)^2 \geq$$

$$\frac{1}{4} \left| \langle [A \;\; B] \rangle \right|^2 + \frac{1}{4} \left| \left\langle \left\{ (A - \langle A \rangle) \;\; (B - \langle B \rangle) \right\} \right\rangle \right|^2 \qquad (5)$$

Box 7. Uncertainty Principle and Cells in Phase Space

The state of a classical particle moving freely inside a box is fully known if its position and momentum are given. But then an infinite number of such states are available, from which it follows that the *a priori* probability that the particle is in any particular state vanishes! This is absurd, as the particle is definitely inside the box. The resolution comes through the Uncertainty Principle which associates spreads like ΔX and ΔP_x with every component of the position and momentum. This reduces the number of accessible states to a finite number, so that the corresponding probabilities are non-zero. This is the origin of the concept of cells in phase space underlying the counting procedures used in statistical physics.

Here $\{A, B\}$ denotes the anti-commutator $(AB + BA)$, and all averages are with respect to a given state. For the special case in which A is the position and B is the momentum, it can be checked that, for a coherent state, the second term in the RHS of (5) vanishes, and we obtain the minimum uncertainty relation $\Delta X \Delta P = \hbar/2$.

Equations (3)-(5) are mathematically unambiguous. They are therefore the precise statements of what Heisenberg attempted to establish semi-quantitatively through his *gedanken* experiment. They have far-reaching implications. For instance, the position-momentum uncertainty relation has played a pivotal role in statistical physics (*Box* 7). Similarly, the energy-time uncertainty relation has played a central role in quantum field theory – a subject that deals, among others, with the interactions between elementary particles (*Box* 8).

Box 8. Energy-Time Uncertainty Principle

The energy-time uncertainty relation states that the product of the energy spread ΔE and the spread in time Δt of a process also satisfies $\Delta E \, \Delta t \geq \hbar/2$. However, unlike position, momentum, energy, etc., time is a parameter and not an observable in quantum mechanics. The interpretation of the energy-time uncertainty relation is therefore somewhat different from that of other uncertainty relations. When applied to an unstable particle, this uncertainty relation yields information on the half-life of the particle. In quantum field theory, it allows for the so-called virtual processes that are required to explain experimental results on reactions involving elementary particles.

Heisenberg has thus played a major role in the march of modern physics. To quote an extract from the Nobel presentation speech by H Pleigel, "· Your quantum mechanics has created new concepts, and has led physics into fresh trains of thought, which have now already proved of fundamental importance for our knowledge of the phenomena of physics. The Royal Academy of Sciences has awarded you the Nobel Prize for Physics for 1932 in recognition of these studies."

Heisenberg's fellow campaigners in the march, Schrödinger and Dirac, shared the Physics Nobel Prize for 1933.

Suggested Reading

[1] **G Gamow, *Thirty Years That Shook Physics: The Story of Quantum Theory*, Dover Publications, New York, 1966.**

[2] **J Mehra and H Rechenberg, *The Discovery of Quantum Mechanics*, Springer-Verlag, New York, 1982.**

[3] **M Jammer, *The Philosophy of Quantum Mechanics*, Wiley, New York, 1974.**

Address for Correspondence
S Lakshmibala
Department of Physics
Indian Institute of Technology
Madras, Chennai 600 036
India.
Email:slbala@physics.iitm.ac.in

To what extent are we bound by tradition in the selection of our problems? ... Looking back upon history in this way, we see that we apparently have little freedom in the selection of our problems. We are bound up with the historical process, our lives are part of this process, and our choice seems to be restricted to the decision whether or not we want to participate in a development that takes place in our time, with or without our contribution.... One may say that a fruitful period is characterized by the fact that the problems are given, that we need not invent them. This seems to be true in science as well as in art.

–Werner Heisenberg