Chapter 2

Heisenberg's Matrix Mechanics and Dirac's Re-creation of it

Matrix mechanics was actually developed somewhat before Schrödinger's work as described in the last chapter and was built upon Heisenberg's early positivistic outlook. Heisenberg reasoned as follows. Physical theory ought to focus on quantities immediately related to observables. Since the Bohr orbits used in the Early Quantum Theory are never directly observable, they should not play a fundamental role in the theory. Now all observable quantities, such as emission frequencies, transition rates, etc., are always related to a pair of Bohr orbits, but never just a single one. Hence Heisenberg sought to build a quantum mechanics based on theoretical constructs linking the totality of all pairs of orbits of a system. Since one usually has an infinity of orbits (stationary states), the most natural mathematical object to use is an infinite matrix:

$$\begin{pmatrix} \times & \times & \times & \dots & \dots \\ \times & \times & \times & \dots & \dots \\ \times & \times & \times & \dots & \dots \\ & \vdots & \vdots & \vdots & \\ & \vdots & \vdots & \vdots & \end{pmatrix} .$$

Heisenberg conjectured that to each Newtonian observable, such as position q or momentum p, there corresponds an infinite matrix, and that the elements of these matrices would actually yield measurable (observable) quantities.

But there was a big problem. Newtonian observables commute; matrices do not in general. This fact caused Heisenberg great anxiety at first. But later it turned out to be the cornerstone of the theory. In fact, as pointed out earlier, it is the mathematical basis of the Uncertainty Principle.

We will not go into the details of exactly how Heisenberg and others (Born and Jordan) constructed matrix mechanics. This was laid out in a series of three seminal papers (Heisenberg 1925; Born and Jordan 1925; Born, Heisenberg and Jordan 1925, English translations of which can be found in Van der Waerden 1968). Instead we will present briefly an elegant account given by Dirac (Dirac 1925). In his paper Dirac observed: "*Heisenberg puts forward a new theory which suggests that it is not the equations of classical mechanics that are in any way at fault, but that the mathematical operations by which physical results are deduced from them require modification. All the information supplied by the classical theory can thus be made use of in the new theory*".

We go back to Hamilton's reformulation of Newtonian mechanics. In the Hamiltonian formulation, one encounters the so-called **Poisson brackets** of the functions of the canonical coordinates and momenta (q, p):

$$\boxed{\{u, v\} \equiv \sum_i \left(\frac{\partial u}{\partial q^i} \frac{\partial v}{\partial p_i} - \frac{\partial u}{\partial p_i} \frac{\partial v}{\partial q^i} \right)} \quad . \tag{2.1}$$

This definition immediately implies that

$$\{q^i, q^j\} = \{p_i, p_j\} = 0 \quad ; \quad \{q^i, p_j\} = \delta^i_j \quad . \tag{2.2}$$

The Poisson brackets are also invariant with respect to **canonical transformations** [cf. (1.19) and (1.20)]:

$$\{u, v\}_{q,p} = \{u, v\}_{Q,P} \quad . \tag{2.3}$$

Two other important properties of the Poisson brackets are

$$\{u, q^i\} = -\frac{\partial u}{\partial p_i} \ , \tag{2.4}$$

$$\{u, p_i\} = \frac{\partial u}{\partial q^i} \ , \tag{2.5}$$

which imply that

$$\{q^i, H\} = \frac{\partial H}{\partial p_i} \ , \tag{2.6}$$

$$\{p_i, H\} = -\frac{\partial H}{\partial q^i} \ , \tag{2.7}$$

where H is the Hamiltonian function.

By Hamilton's equations of motion

$$\frac{dq^i}{dt} = \frac{\partial H}{\partial p_i} \quad , \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q^i} \quad , \tag{2.8}$$

(2.6) and (2.7) give

$$\frac{dq^i}{dt} = \{q^i, H\} \quad , \quad \frac{dp_i}{dt} = \{p_i, H\} \quad . \tag{2.9}$$

The time derivative of a general function $u(p_i, q^i)$ of the canonical coordinates and momenta is then obtained from Hamilton's equations of motion as follows:

$$\frac{du}{dt} = \frac{\partial u}{\partial q^i}\frac{dq^i}{dt} + \frac{\partial u}{\partial p_i}\frac{dp_i}{dt} + \frac{\partial u}{\partial t} = \frac{\partial u}{\partial q^i}\frac{\partial H}{\partial p_i} - \frac{\partial u}{\partial p_i}\frac{\partial H}{\partial q^i} + \frac{\partial u}{\partial t} \quad , \tag{2.10}$$

that is,

$$\boxed{\frac{du}{dt} = \{u, H\} + \frac{\partial u}{\partial t}} \quad . \tag{2.11}$$

Problem 2.1 Using the definition of the Poisson bracket given by (2.1), prove (2.2), (2.4), and (2.5), and the fact that Poisson brackets are invariant under canonical transformations [(2.3)].

Dirac gave the following recipe to convert (2.11) into a quantum mechanical equation of motion for dynamical quantities:

$$\boxed{\{\, a \,,\, b \,\} \longrightarrow \frac{[\, a \,,\, b \,]}{i\hbar}} \quad , \tag{2.12}$$

where $[a, b] \equiv ab - ba$ is the commutator of the operators (or matrices) a and b. Thus the correct quantum mechanical equation of motion is

$$\boxed{\frac{du}{dt} = \frac{[u, H]}{i\hbar} + \frac{\partial u}{\partial t}} \quad , \tag{2.13}$$

where u and H are in general infinite matrices. This is known as **Heisenberg's equation of motion**.

The recipe (2.12), together with (2.2), immediately yields the quantum conditions

$$[\, q^i \,,\, p_j \,] = i\hbar\, \delta^i_j \quad , \quad [\, q^i \,,\, q^j \,] = [\, p_i \,,\, p_j \,] = 0 \quad , \tag{2.14}$$

which are the same as (1.37), derived from quite a different route.

Dirac justifies recipe (2.12) by the following arguments (see Dirac 1967, pp. 86, 87). First observe that the classical Poisson brackets $\{\ ,\ \}$ satisfy the following properties:

$$\{\, u \,,\, v \,\} = -\{\, v \,,\, u \,\} \quad , \tag{2.15}$$

$$\begin{aligned} \{ u_1 + u_2 \, , \, v \} &= \{ u_1 \, , \, v \} + \{ u_2 \, , \, v \} \, , \\ \{ u \, , \, v_1 + v_2 \} &= \{ u \, , \, v_1 \} + \{ u \, , \, v_2 \} \, , \end{aligned} \tag{2.16}$$

$$\begin{aligned} \{ u_1 u_2 \, , \, v \} &= \{ u_1 \, , \, v \} \, u_2 + u_1 \{ u_2 \, , \, v \} \, , \\ \{ u \, , \, v_1 v_2 \} &= \{ u \, , \, v_1 \} \, v_2 + v_1 \{ u \, , \, v_2 \} \, , \end{aligned} \tag{2.17}$$

as well as the **Jacobi identity**

$$\boxed{\{ u \, , \{ v \, , \, w \}\} + \{ v \, , \{ w \, , \, u \}\} + \{ w \, , \{ u \, , \, v \}\} = 0} \quad . \tag{2.18}$$

Problem 2.2 Using the definition (2.1), verify the properties of the Poisson bracket as given by Eqs. (2.15) to (2.18).

Now evaluate $\{ u_1 u_2 \, , \, v_1 v_2 \}$ in two different ways, using one and then the other of the two formulas in (2.17) first:

$$\begin{aligned} &\{ u_1 u_2 \, , \, v_1 v_2 \} = \{ u_1 \, , \, v_1 v_2 \} \, u_2 + u_1 \{ u_2 \, , \, v_1 v_2 \} \\ &= (\{ u_1 \, , \, v_1 \} \, v_2 + v_1 \{ u_1 \, , \, v_2 \}) \, u_2 + u_1 (\{ u_2 \, , \, v_1 \} \, v_2 + v_1 \{ u_2 \, , \, v_2 \}) \\ &= \{ u_1 \, , \, v_1 \}, v_2 u_2 + v_1 \{ u_1 \, , \, v_2 \} \, u_2 + u_1 \{ u_2 \, , \, v_1 \} \, v_2 + u_1 v_1 \{ u_2 \, , \, v_2 \} \quad ; \end{aligned}$$

$$\begin{aligned} &\{ u_1 u_2 \, , \, v_1 v_2 \} = \{ u_1 u_2 \, , \, v_1 \} \, v_2 + v_1 \{ u_1 u_2 \, , \, v_2 \} \\ &= (\{ u_1 \, , \, v_1 \} \, u_2 + u_1 \{ u_2 \, , \, v_1 \}) \, v_2 + v_1 (\{ u_1 \, , \, v_2 \} \, u_2 + u_1 \{ u_2 \, , \, v_2 \}) \\ &= \{ u_1 \, , \, v_1 \}, u_2 v_2 + u_1 \{ u_2 \, , \, v_1 \} \, v_2 + v_1 \{ u_1 \, , \, v_2 \} \, u_2 + v_1 u_1 \{ u_2 \, , \, v_2 \} \quad . \end{aligned}$$

Equating these two results, we obtain

$$\{ u_1 \, , \, v_1 \} \, [u_2 \, , \, v_2] = [u_1 \, , \, v_1] \, \{ u_2 \, , \, v_2 \} \quad .$$

Since $u_1 \, , \, u_2 \, , \, v_1 \, , \, v_2$ are all independent of each other, the above condition implies

$$[u_1 \, , \, v_1] = i\hbar \, \{ u_1 \, , \, v_1 \} \quad , \quad [u_2 \, , \, v_2] = i\hbar \, \{ u_2 \, , \, v_2 \} \quad , \tag{2.19}$$

where $\hbar$ is a real number independent of the u's and v's, and has the dimensions of action (energy $\times$ time). The appearance of the imaginary unit i in (2.19) is of the utmost significance. Its introduction is dictated by the fact that the u's and v's must be **hermitian matrices** if they are to represent **dynamical variables** (**observables**) in quantum theory, where the mathematical property of **hermiticity** of a matrix is defined by the condition

$$u = u^\dagger \equiv (u^*)^T \quad .$$

In the above equation, $*$ denotes complex conjugation and T denotes the transpose of a matrix. $u^\dagger$ is called the **hermitian adjoint** of u.

The important mathematical requirement that observables be represented by hermitian matrices (operators) in the formalism of matrix mechanics is due to the physical requirement that measured values of observable quantities must be real. According to the formalism (discussed in Chapter 8), these are given by the eigenvalues of the matrices (or the **spectrum** of the corresponding operators). The property of hermiticity of a matrix guarantees that its eigenvalues must be real. (This will also be shown in Chapter 8.)

Comparing (2.11) and (2.13) we see that in the transition from classical to quantum mechanics, classical Poisson brackets, which are classical observables and are real, must be replaced by "quantum Poisson brackets", which are required to be hermitian. Note that $[\,u\,,\,v\,]$ is not necessarily hermitian even if u and v are. On the other hand, it is straightforward to see that, if u and v are hermitian, then $\pm i[\,u\,,\,v\,]$ are both hermitian:

$$\begin{aligned}(\pm\, i[\,u\,,\,v\,])^\dagger &= \mp\, i(uv - vu)^\dagger = \mp\, i(v^\dagger u^\dagger - u^\dagger v^\dagger) \\ &= \mp\, i(vu - uv) = \pm\, i(uv - vu) = \pm\, i[\,u\,,\,v\,] \quad .\end{aligned}$$

In (2.19) the choice $+i$ has been picked by convention.

Suppose u does not depend explicitly on time in (2.13), so that $\partial u/\partial t = 0$. Then the Heisenberg equation of motion reads

$$\frac{dq}{dt} = \frac{[\,q\,,\,H\,]}{i\hbar} \tag{2.20}$$

for a system with one degree of freedom. Equation (2.20) is actually a matrix equation, and thus represents (since q is an infinite matrix in general) an infinite set of coupled equations:

$$\frac{dq_{nm}}{dt} = \frac{(qH)_{nm} - (Hq)_{nm}}{i\hbar} = \sum_k \frac{(q_{nk}H_{km} - H_{nk}q_{km})}{i\hbar} \quad . \tag{2.21}$$

If we can find a **representation** in which the Hamiltonian H is diagonal, that is, $H_{nk} = \delta_{nk}H_{nn}$, then the equations become decoupled:

$$\begin{aligned}\frac{dq_{nm}}{dt} &= \sum_k \left(q_{nk}\delta_{km}H_{mm} - H_{nn}\delta_{nk}q_{km}\right) / \,(i\hbar) \\ &= q_{nm}(H_{mm} - H_{nn}) \,/\, (i\hbar) \\ &= -i\omega_{mn}q_{nm} = i\omega_{nm}q_{nm} \quad ,\end{aligned} \tag{2.22}$$

where

$$\omega_{mn} \equiv \frac{(H_{mm} - H_{nn})}{\hbar} \quad . \tag{2.23}$$

Equation (22) can be solved easily to obtain

$$q_{nm}(t) = q_{nm}(0)\, e^{i\omega_{nm}t} \quad . \tag{2.24}$$

Hence the fundamental problem of Heisenberg's matrix mechanics is to find infinite matrices q^i and p_i such that the quantum conditions (2.14) hold, and such that the Hamiltonian $H(q^1, \ldots, q^N; p_1, \ldots, p_N)$ becomes a diagonal matrix.

On the face of it, this seems to have nothing to do with Schrödinger's approach discussed in the last chapter. It is indeed one of the most remarkable results in the development of the quantum theory that the two approaches, Schrödinger's wave mechanics and Heisenberg's matrix mechanics, are in fact equivalent. This will be further discussed in Chapters 4 and 5.