Learning Quantum Mechanics from Heisenberg

L. Doria[1]
[1] TRIUMF, 4004 Wesbrook Mall, Vancouver, B.C. V6T 2A3, Canada and
[1] Douglas College, 700 Royal Ave, New Westminster, BC V3M 5Z5

After 100 years from its discovery, Quantum Mechanics still represents a significant challenge for students facing it for the first time. Often the obstacle is not represented by the formalism but from the lack of understanding of the deep reasons that lead to the new mechanics. Following the suggestion of [1], we try to introduce Quantum Mechanics following the steps which led W. Heisenberg to the very first version of the theory in 1925. The discussion is simplified and requires a basic knowledge of Fourier series, linear algebra and the phenomenology of spectroscopy. We avoid the explicit use of Hamiltonian mechanics, which is usually not taught in introductory courses.

INTRODUCTION

In June 1925, W. Heisenberg published a fundamental paper which paved the way to the construction of Quantum Mechanics [2] and deeply influenced other contemporary physicists. After the initial impact, the paper is nowadays considered quite obscure and difficult to understand. In [1], the authors present a very clear analysis of Heisenberg's paper and argue that it might be pedagogically useful to present its main results to undergraduate students approaching Quantum Mechanics for the first time. Teaching Quantum Mechanics today is of growing importance also outside physics departments, since it is becoming part of our everyday life for example through many engineering applications. Traditionally, Quantum Mechanics is introduced with the Schrödinger formalism. Another path which is becoming popular is to follow a close analogy with classical probability theory and the introducing complex numbers and L_2 norms as the only meaningful generalization (see e.g. [4]). A more mathematical but insightful approach based on axioms as a starting point is provided by the work of L. Hardy [5].

In this manuscript, following the suggestion of [1], we try to present in a simplified way the reasoning which lead Heisenberg to the discovery of "Matrix Mechanics" through a series of deep observations and calculations. Right after the initial breakthrough of Heisenberg, Born and Jordan published a paper where they introduce the matrix formalism [6]. Heisenberg, Born and Jordan finally joined forces and published a paper containing the full formalism of the new mechanics together with a perturbation theory and the first attempt towards the quantization of the electromagnetic field [7]. Soon after, W. Pauli applied the new results to the solution of the hydrogen atom, deriving Balmer's formula for the energy levels [8]. The work of Pauli greatly contributed to the acceptance of the new mechanics. For more historical details, we refer to [9].

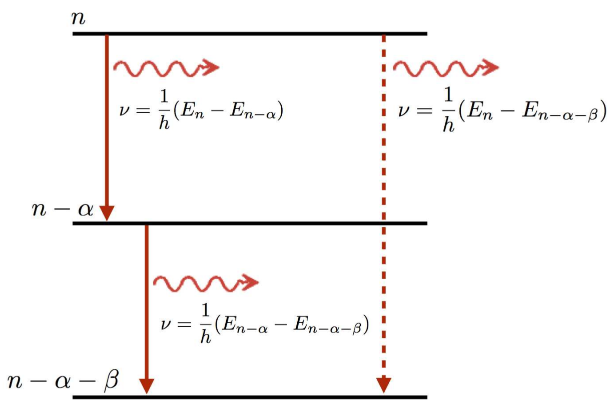

FIG. 1: Energy levels are discrete and arranged in a sequence of states ordered by energy. The Ritz combination principle encodes the level ordering: the frequency of a photon emitted in the $n \to n-\alpha-\beta$ transition must be equal to the sum of the frequencies in the $n \to n-\alpha$ and $n-\alpha \to n-\alpha-\beta$ transitions.

BEFORE JUNE 1925

Planck's and Einstein's results on blackbody radiation and the empirical formulas describing the experimentally observed spectroscopic lines forced a new view about the way in which atoms absorb and emit radiation. In particular, it was realized that the frequency of an emitted or absorbed photon is connected to the difference between two energy levels α and β:

$$\nu = \frac{1}{h}\left(E_\alpha - E_\beta\right) \quad . \tag{1}$$

Other indications were coming from the successful development of a dispersion theory consistent with classical formulas in the case of large quantum numbers (in accordance with Bohr's correspondence principle). [10–12]

COMBINING FREQUENCIES

The vast amount of spectroscopic data available at the time led to the so-called *combination principle* of Ritz

(1908) which states that the frequency of a spectroscopic line is equal to the sum or the difference of two other lines' frequencies. This principle was clearly consistent with Eq. 1. Apparently, at the atomic level, only certain frequencies were allowed, as they were arranged in a discrete "ladder". This observation forces us to introduce a new way for summing frequencies together. Let us consider a classical wave associated to a certain state of the system with label n together with all its harmonics. The frequency of a harmonic will be denoted as $\nu(n,\alpha) = \alpha\nu(n)$, since it is an integer multiple of n. Classically, the sum of two harmonics is

$$\nu(n,\alpha) + \nu(n,\beta) = \nu(n,\alpha+\beta) \quad . \tag{2}$$

If we now consider atoms and their "ladder" of states, the sum of frequencies corresponding to different steps (the "harmonics") must add in a very different way: making two consecutive steps must be equivalent to a two-step jump (see Fig. 1):

$$\nu(n,n-\alpha)+\nu(n-\alpha,n-\alpha-\beta) = \nu(n,n-\alpha-\beta) \quad . \tag{3}$$

Another way to rewrite the last equation is

$$\nu(n,n-\alpha)+\nu(n-\alpha,n-\alpha-\beta)+\nu(n-\alpha-\beta,n) = 0 \quad , \tag{4}$$

which led to the conclusion that ν can be written as a difference of two other quantities, consistent with Eq. 1.

Thus, the sum of classical harmonics differs from the "quantum" one. Moreover, Eq. 3 is fully consistent with Eq. 1, while Eq. 2 is not. This is the fist important point which marks a fundamental difference between classical and quantum states.

FOCUS ON OBSERVABLES

In his atomic model, Bohr used classical formulas containing also the position of the electron. Heisenberg argues that experimentally we do not observe electron positions, but only frequencies and intensities of spectral lines. His program was to build a theory based only on observable quantities. To this end, instead of the position $x(n,t)$ of an electron in a state n at time t, he considered its Fourier representation

$$x(n,t) = \sum_{\alpha=-\infty}^{+\infty} X_\alpha(n) e^{2\pi i \nu(n)\alpha t} \quad . \tag{5}$$

In the last equation, the frequency $\nu(n)$ (and its multiples, or harmonics $\alpha\nu$) and the intensities X_α appear explicitly. Accoring ti Heisenberg, we have to focus our attention on them disregarding the unobservable position x (he was not completely right on this point, since in the final quantum theory $x(t)$ can also be an observable.).

The key idea of Heisenberg was to introduce quantum transitions in Eq. 5: this means that now amplitudes and frequencies must depend on *two* numbers:

$$x(n,n-\alpha,t) = \sum_{\alpha=-\infty}^{+\infty} X_\alpha(n,n-\alpha) e^{2\pi i \nu(n,n-\alpha)t} \quad . \tag{6}$$

We thus introduce the new "transition amplitudes" $X(n,n-\alpha)$ from the state n to the state $n-\alpha$.

Heisenberg observes that radiation is commonly treated as a multipole expansion where powers of positions and their derivatives appear. It is therefore necessary to derive expressions for *e.g.* $[x(n,t)]^2$ or higher powers. Classically we have:

$$[x(t)]^2 = \sum_\alpha \sum_\gamma X_\alpha(n) X_\gamma(n) e^{2\pi i \nu(n)(\alpha+\gamma)t} = \tag{7}$$
$$= \sum_\beta Y_\beta(n) e^{2\pi i \nu(n)\beta t} \quad ,$$

where we relabel $\beta = \alpha + \gamma$. Dropping the exponentials, we can write the following relationship among the amplitudes:

$$Y_\beta(n) = \sum_\alpha X_\alpha(n) X_{\beta-\alpha}(n) \quad . \tag{8}$$

Also in this case Heisenberg rewrites the last equation enforcing the validity of the combination principle (Eq. 3):

$$Y(n,n-\beta) = \sum_\alpha X(n,n-\alpha) X(n-\alpha,n-\beta) \quad . \tag{9}$$

This is probably the most important step in Heisenberg's paper. The last equation is the law for multiplying amplitudes in quantum mechanics and he realized right away that there is a radical difference with respect to the classical case. If we consider two quantities $x(t)$ and $y(t)$, the product is classically commutative: $xy = yx$, as it is evident from Eq. 8, but this is not generally the case for Eq. 9 .

Swapping the terms in the righthand side of Eq. 9 will change the ordering of the states and the result of the multiplication will be different. Therefore, Heisenberg discovered that in the quantum theory, observables are not in general commuting quantities and the key reason is that quantum observables depend on *two* indices (they represent *transitions* between two states).

MATRICES

Back in 1925, Heisenberg was not aware of matrix algebra. M. Born knew it and looking at Heisenberg result realized that Eq. 9 was completely analogous to matrix multiplication, which indeed is not in general a commutative operation. We can rewrite Eq. 9 as

$$Y_{\alpha\beta} = X_{\alpha\gamma} X_{\gamma\beta} \quad . \tag{10}$$

Let us now consider Bohr's quantization condition:

$$\oint pdq = \int_0^{1/\nu} p\dot{q}dt = nh \quad , \tag{11}$$

which was so successful in explaining some of the atomic properties. Following Heisenberg, we express the position q and momentum p as Fourier components

$$q(\tau, t) = \sum_{\tau=-\infty}^{\tau=+\infty} Q_\tau e^{2\pi i\nu\tau t} \quad ; \quad p(\tau, t) = \sum_{\tau=-\infty}^{\tau=+\infty} P_\tau e^{2\pi i\nu\tau t} \tag{12}$$

Heisenberg comments that Eq. 11 arbitrarily forces the dependence of the quantum variables on the quantum number n, while this should be a result of the theory.

The way he removes the dependence is to take the derivative $\partial/\partial n$ on both sides of Eq. 11 in a simplified situation where $p = m\dot{q}$. In [6], Born and Jordan consider a more general case, keeping the two variables q and p separate. The time derivative of q is $\dot{q} = \frac{2\pi i}{\nu}\tau q$. Taking the derivative with respect to n in Eq. 11, transforming the integral in a sum (the states are discrete), and taking into account the integration limits $[0, 1/\nu]$, we obtain

$$1 = \frac{2\pi i}{h} \sum_{\tau=-\infty}^{\tau=+\infty} \tau \frac{\partial}{\partial n}(Q_\tau P_{-\tau}) \quad . \tag{13}$$

The partial derivative is solved by Heisenberg recalling his previous work on dispersion relations with Kramer [11]. Born uses the same technique, which is to replace the derivative with a finite difference:

$$1 = \frac{2\pi i}{h} \sum_{\tau=-\infty}^{\tau=+\infty} Q(n+\tau, n)P(n, n+\tau) - Q(n, n-\tau)P(n-\tau, n) \tag{14}$$

In this way, an expression describing a state is now changed in an expression describing a *transition between states*. Rewriting the last equation in matrix formalism with relabeled indices we obtain

$$Q_{\alpha\gamma}P_{\gamma\alpha} - P_{\gamma\alpha}Q_{\gamma\alpha} = -\frac{h}{2\pi i}\delta_{\alpha\alpha} \quad , \tag{15}$$

which confirms the non-commutativity among quantum mechanical observables, generalizing Heisenberg's results. Still, the latter result refers only to the diagonal elements of the matrix: what about the off-diagonal terms? Born conjectured that these terms must be zero and Jordan finally provided an argument in favor of this intuition (see next section). A more formal proof was soon provided by Dirac [13].

THE NEW QUANTUM THEORY

Looking for an equation describing how quantum observables change in time, we can consider the time derivative of a function $f(q, p)$ where q and p are expressed via their Fourier decompositions. The derivative will look like

$$\dot{f}_{\alpha\beta} = 2\pi i\nu_{\alpha\beta} f_{\alpha\beta} \quad . \tag{16}$$

Since the term $\nu_{\alpha\beta}$ represents the frequency emitted by a transition from two states, we have to assume that $\alpha \neq \beta$ (or that ν is antisymmetric: $\nu_{\alpha\beta} = -\nu_{\beta\alpha}$). If $\alpha = \beta$, then $\dot{f} = 0$. On the other hand, if $\dot{f} = 0$ and $\nu_{\alpha\alpha} = 0$ we conclude that f *must be a diagonal matrix*. We can now rewrite Bohr's relation (Eq. 1) between frequency and energy differences in matrix form and substitute it in Eq. 16:

$$\dot{f} = \frac{2\pi i}{h}\left[(E_\alpha f_{\alpha\beta} - f_{\alpha\beta}E_\beta\right] \quad . \tag{17}$$

The last equation is what we call today Heisenberg's equation of motion, although its first derivation was given by Born and Jordan.

In the last section, we calculated the diagonal term of the matrix $A = qp - pq$ (Eq. 15). We can further prove that the off-diagonal terms of A are zero, proving that A is diagonal, which is equivalent to $\dot{A} = 0$. Born and Jordan give a quite general argument based on the derivation of functions of products of powers of q and p and the Hamiltonian formalism. A less general and simplified argument is the following. Let us consider the time derivative of the matrix A:

$$\dot{A} = \dot{q}p + q\dot{p} - \dot{p}q - p\dot{q} \quad . \tag{18}$$

In a very simple dynamic situation, if p is the momentum, q the position, m the mass and $U(q)$ a potential, we have $\dot{q} = p/m$ and $\dot{q} = -\partial U/\partial q$. Substituting into the last equation:

$$\dot{A} = -\frac{\partial U}{\partial q}q + \frac{p^2}{m} - \frac{p^2}{m} + q\frac{\partial U}{\partial q} = 0 \quad , \tag{19}$$

where we used the fact that $\partial U/\partial q$ commutes with q. This (not rigorous) calculation supports the conclusion that $\dot{A} = 0$ and therefore the matrix $A = qp - pq$ is diagonal. The final expression is one of the central results of Quantum Mechanics:

$$qp - pq = -\frac{h}{2\pi i}I \quad , \tag{20}$$

where I is the unit matrix. The last equation has far reaching consequences in our ability to perform measurements, as the fathers of Quantum Mechanics realized quickly by simple considerations about the mathematical properties of matrices. Since energy is conserved, $\dot{E} = 0$ and E must be a diagonal matrix:

$$E = \delta_{\alpha\beta}E_\alpha \quad \text{and} \quad h\nu_{\alpha\beta} = E_{\alpha\alpha} - E_{\beta\beta} \quad . \tag{21}$$

We can assume that the eigenvalues (or equivalently the diagonal elements if the matrix is diagonal) of a matrix representing the observable are the values we measure. But if two matrices do not commute, we cannot diagonalize them at the same time with the same transformation. This leads to the conclusion that if we diagonalize q, we cannot do it for p and thus if we measure precisely the position of a particle, we cannot measure its momentum with the same precision. Eq. 20 represents Heisenberg's uncertainty principle. Its final form was derived by Born and Jordan [6].

CONCLUSIONS

Following the papers of Heisenberg, Born, and Jordan we showed in a simplified way how two central results of Quantum Mechanics (the equation of motion and the uncertainty principle) were deduced for the first time. The key ideas behind these breakthroughs were the following:

- Heisenberg focused his analysis only on what spectroscopy experiments were actually measuring: intensities and frequencies. This led him to represent observable quantities with Fourier series.

- He then enforced the Ritz combination principle into the Fourier decomposition: frequencies and amplitudes were now dependent on two numbers instead of one. This is fundamental for introducing the idea of transition between two states in the formalism.

- Investigating the algebra of the new quantities led him to the non-commutativity of the amplitudes, which has its roots exactly in the combination principle.

- Born, Jordan and Dirac generalized the formalism recognizing the isomorphisms between Heisenberg's new quantities and matrix algebra. The first form of Quantum Mechanics was born: Matrix Mechanics.

The first picture of Quantum Mechanics was born 100 years ago with the seminal paper of W. Heisenberg. In the modern teaching of Quantum Mechanics, his line of reasoning is somewhat lost in favor of the more practical Schrödinger picture. Heisenberg's original approach, though, has the pedagogical advantage of making a clear connection with experimental facts. Moreover, it is of great historical interest, and it is an excellent example of physics reasoning and intuition, which highlights how new ideas and advancements in physics can take place.

[1] I.J.R. Aitchinson *et al.*, Am. J. Phys. 72, 1370 (2004).
[2] W. Heisenberg, Zeitschrift fr Physik, 33, 879-893 (1925).
[3] W. Heisenberg, Annalen der Physik **379**(15), 577 (1924).
[4] S. Aaronson, Quantum Computing Since Democritus, Cambridge University Press (2013)
[5] L. Hardy, arXiv:0101012 (2001)
[6] M. Born, P. Jordan, Zeitschrift fr Physik **34**, 858-888 (1925).
[7] M. Born, W. Heisenberg, P. Jordan, Zeitschrift fr Physik **35**(8-9) 557-615 (1925).
[8] W. Pauli, Z. Physik **36**, 336-363 (1926).
[9] B.L. van der Waerden, Sources of Quantum Mechanics, ISBN 0-486-61881-1.
[10] H.A. Kramers, Nature **113**, 673 (1924).
[11] H.A. Kramers, W. Heisenberg, ZS. f. Phys., **31**, 381 (1925)
[12] M. Born, ZS. f. Phys., **26**, 379 (1924)
[13] P.A.M Dirac, Proc. Roy. Soc. A**109**, 642, 653 (1926)