La Matematica nella Società e nella Cultura

Rivista dell'Unione Matematica Italiana

Christos Efthymiopoulos

Perturbative methods in Celestial Mechanics and the roots of Quantum Mechanics: a historical note

La Matematica nella Società e nella Cultura. Rivista dell'Unione Matematica Italiana, Serie 1, Vol. **8** (2015), n.2, p. 191–224.

Unione Matematica Italiana

<http://www.bdim.eu/item?id=RIUMI_2015_1_8_2_191_0>

La Matematica nella Società e nella Cultura. Rivista dell'Unione Matematica Italiana, Unione Matematica Italiana, 2015.

La Matematica nella Società e nella Cultura
Rivista dell'Unione Matematica Italiana
Serie I, Vol. VIII, Agosto 2015, 191-224

Perturbative methods in Celestial Mechanics and the roots of Quantum Mechanics: a historical note

CHRISTOS EFTHYMIOPOULOS

1. – The "Three-Man" paper of Quantum Mechanics

It is commonly accepted that the era of modern quantum theory starts by the publication, in 1925, of Werner Heisenberg's landmark paper on *Matrix Mechanics* [32] ("Quantum-Theoretical Re-interpretation of Kinematic and Mechanical Relations"). Heisenberg was at the time postdoctoral collaborator of the leading physicist of his epoch Max Born, in Göttingen, to whom he had returned after a visit to Neils Bohr's renown Institute of Theoretical Physics in Copenhagen. In the history-of-physics webpages of the American Institute of Physics we read [2]: "Heisenberg set himself the task of finding the new quantum mechanics upon returning to Göttingen from Copenhagen in April 1925. Inspired by Bohr and his assistant, H.A. Kramers, in Copenhagen, Pauli in Hamburg, and Born in Göttingen, Heisenberg's intensive struggle over the following months to achieve his goal has been well documented by historians. Since the electron orbits in atoms could not be observed, Heisenberg tried to develop a quantum mechanics without them. He relied instead on what can be observed, namely the light emitted and absorbed by the atoms. By July 1925 Heisenberg had an answer, but the mathematics was so unfamiliar that he was not sure if it made any sense. Heisenberg handed a paper on the derivation to his mentor, Max Born, before leaving on a month-long lecture trip to Holland and England and a camping trip to Scandinavia with his youth-movement group. After puzzling over the derivation, Born finally recognized that the unfamiliar mathematics was related to the mathematics of arrays of numbers known as "matrices". Born sent Heisenberg's paper off for publication. It was the breakthrough to quantum mechanics."

Heisenberg's theory was based upon the idea that the basic kinematic quantities of a particle (e.g. position, momentum, energy) could be represented in terms of what we call today Hermitian square matrices. It seems, however, that the concept of matrix, nowadays presented in high school or early college mathematical curricula, was largely unknown to physicists at the time. Soon after Heisenberg's paper, Born and his assistant Pascual Jordan presented their own paper in Zeitschrift für Physik entitled "On Quantum Mechanics" ([9]). As stated in the abstract, the aim of that paper was to develop the (then, recently) published approach of Heisenberg into a systematic theory of quantum mechanics. Nevertheless, section 1 of that paper is only devoted to explaining basic algebraic properties of matrices, including how to perform matrix multiplication, the non-commutative property, as well as the definitions of inverse and identity matrix. Section 2, instead, deals with true concepts of dynamics, setting the new quantum framework for mechanics. Finally, section 3 introduces the quantization of the electromagnetic field, by considering the electric and magnetic fields as dynamical variables entering in a harmonic-oscillator-type re-expression of Maxwell's equations. This is the key idea behind even present-day quantum field theory.

The cross-fertilization of ideas between Born, Heisenberg and Jordan culminated in a joint publication, in 1926, of the second part of "On Quantum Mechanics" [10]. In the literature, this latter paper is occasionally referred to as as the "three-man" paper of quantum physics (not to be confused with a known "three-man paper" in molecular biology, which coincidently appeared at a similar time as the one of physics). A significant part of the three-man paper is devoted to expressing the newly formed ideas on quantum mechanics in a context that was quite familiar in classical mechanics, namely the so-called *canonical formalism*. In this formalism, the dynamics is expressed in terms of a certain function called the Hamiltonian, which serves as the generator of the equations of motion (via the so-called Hamilton's equations). Furthermore, one may introduce appropriate changes of variables, induced via so-called canonical (or symplectic) transformations. Born, Heisenberg and Jordan (hereafter BHJ) clarified the use of the Hamiltonian formalism as well as how canonical tran-

sformations should be introduced in the newly emerging quantum-theoretical language of matrix mechanics.

At a similar time, the Austrian physicist Erwin Schrödinger was proposing his "wave" version of quantum mechanics [44] (the British physicist Paul Dirac at Cambridge was simultaneously developing an early form of 'axiomatic' quantum theory, which included his celebrated 'bra' and 'ket' representations of quantum states [19]-[20], see the *Appendix* for basic definitions and a short mathematical introduction to quantum theory). In deriving the equation which carries his name, Schrödinger, as well, drew largely upon fundamental (albeit different than in Heisenberg's approach) notions of classical mechanics. In particular, Schrödinger discusses extensively the analogy between trajectories in mechanics and rays in optics, which, in the classical level, is expressed via the well known *Hamilton-Jacobi equation*. ([1]) Initially, it was unclear whether, and up to what extent, the two new quantum theories, i.e., matrix and wave mechanics, could be considered as equivalent, or sharing a common physical content. The issue was settled by Schrödinger himself in a subsequent paper in Annalen der Physik (1926, [45]). Here, Schrödinger clearly explains how the matrix variables of Heisenberg lead to a formal scheme equivalent to the classical canonical scheme, after a re-interpretation of Poisson brackets as equivalent to the quantum *commutators* between two matrix variables. Let us also point out that the use of the term 'canonical', ubiquitous in these early papers of quantum mechanics, in the course of time has been replaced by the use of the term 'unitary', which nowadays commonly appears in quantum mechanics textbooks. Thus, modern textbooks define 'unitary' (instead of canonical) transformations of 'observables', or quantum variables. ([2])

([1]) After introducing, in his own paper, all the basic properties of the Hamilton-Jacobi equation, Schrödinger writes: "Nothing of what has hitherto been said is in any way new. All this was very much better known to Hamilton himself than it is in our day to a good many physicists". This sounds like a comment of a present-day senior professor speaking about younger colleagues. But Schrödinger's paper was published in 1926, and it refers to a paper of Hamilton published in 1834 [31].

([2]) Another common reference is to 'q-valued' (i.e. quantum, meaning, operator-valued) instead of 'c-valued' (i.e. complex scalar, or classical-valued) quantities.

2. – Classical and Quantum Lindstedt series

2.1 – *Heisenberg picture*

Despite their epoch-making role in physics, the original paper by Heisenberg, and to a lesser extent, the subsequent papers with Born and Jordan, are not so oftenly quoted in modern courses of quantum mechanics. In their 'Heisenberg made-easy' attempt to explain Heisenberg's paper, Aitchison et al. [1] emphasize that part of the difficulty to understand Heisenberg's paper lies in the lack of "clues as to how he arrived at the results which he reported". They quote, in respect, the opinion of Nobel physicist S. Weinberg [47]: "If the reader is mystified at what Heisenberg was doing, he or she is not alone. I have tried several times to read the paper that Heisenberg wrote on returning from Heligoland, and, although I think I understand quantum mechanics, I have never understood Heisenberg's motivations for the **mathematical** ([3]) steps in his paper. Theoretical physicists in their most successful work tend to play one of two roles: they are either sages or magicians... It is usually not difficult to understand the papers of sage-physicists, but the papers of magician-physicists are often incomprehensible. In that sense, Heisenberg's 1925 paper was pure magic". Aitchison et al. (2004) continue by noting that "one of the main barriers to understanding Heisenberg's paper, for most people, is a more prosaic one: namely, that he gives remarkably few details of the calculations he actually performed, in order to arrive at his results for the one-dimensional model systems which he treats (anharmonic oscillators and the rigid rotator)". ([4])

Then, Aitchison et al. try themselves to reproduce these detailed calculations for the sake of their readers. Let us recall here that the basic model used by Heisenberg is what we would call today a Duffing (or one-dimensional anharmonic) oscillator, represented by the equation of motion:

$$\ddot{x} + \omega_0^2 x + \lambda x^{b-1} = 0 \ . \tag{1}$$

([3]) The emphasis is ours.

([4]) An instructive presentation (in italian) of the fundamental papers of matrix mechanics, including Heisenberg's first paper, can be found in [12].

This corresponds to the motion of a particle (of unit mass) in a potential well given by

$$V(x) = \frac{1}{2}\omega_0^2 x^2 + \frac{1}{b}\lambda x^b \,.$$

The first term in $V(x)$ is a harmonic oscillator term with proper frequency ω_0. The second term represents the anharmonic correction. The 'anharmonicity parameter' λ is considered as a small parameter.

Heisenberg examines two examples, with a cubic ($b = 3$) or quartic ($b = 4$) correction. The state-of-the-art understanding at his time was the following: Classically, the particle (electron) oscillates, in general, between two limiting values of x, while, for b odd, the equilibrium point is off-centered by a constant $\mathcal{O}(\lambda)$. According to the classical theory of electrodynamics, the oscillatory motion of the electron results in that the electron loses energy by emitting a *continuum* of electromagnetic radiation (see, for example, [35]). For an electron in an atomic electrostatic potential due to the positive charge in the nucleus, such loss of energy via electromagnetic radiation would imply that the electron would in-spiral towards the nucleus, eventually falling into it. The so-called 'old quantum theory' of Bohr, Sommerfeld and Kramers (see [11] for a review) was attempting to remedy these predictions by postulating that i) the emission of energy should take place, instead, by *discrete* quanta, and ii) no more emission of energy would be possible after the electron reached its so-called 'ground state', i.e., a state of lowest possible bound energy.

Heisenberg focuses now on formal aspects of the classical theory, and argues on how to properly generalize them in his new quantum theory. According to the classical theory of radiation, the energy emission rate for the electron can be determined once the electron's orbit $x(t; x_0, v_0)$ for given initial condition ($x(0) = x_0, \dot{x}(0) = v(0) = v_0$) has been specified. To this end, one has to express $x(t; x_0, v_0)$ in terms of a *periodic Fourier representation*

$$x(t; x_0, v_0) = \sum_{k=-\infty}^{\infty} X_k(x_0, v_0) e^{ik\omega(x_0, v_0)t} \quad , \tag{2}$$

where X_k are Fourier coefficients and ω is the fundamental frequency of motion. One should stress here that a basic property of classical anharmonic oscillators is that the fundamental frequency ω is a variable quantity depending on the initial conditions. In general one has $\omega \neq \omega_0$, i.e. the frequency differs from the one of the harmonic oscillator limit ($\lambda = 0$ in Eq. (1)). Furthermore, returning to the question of the emitted radiation, one may show that the k–th harmonic (i.e. with frequency $k\omega$) component of the Fourier decomposition of the orbit, given by Eq. (2), contributes to the energy decay rate due to electromagnetic radiation by a quantity proportional to $|\omega k|^4 |X_k|^2$.

Heisenberg invokes now a mixture of classical and quantum arguments in order to define what should be an appropriate generalization of Eq. (2) in the quantum case. The essence of his arguments is as follows:

Classically, the knowledge of the initial conditions (x_0, v_0) would suffice to characterize the so-called *state* of the system, i.e. the value of the phase space variables $(x(t), v(t))$ at any later time t. This would yield a continuum of states inasmuch as we have a continuum of possible initial conditions. Quantum-mechanically, however, we want to incorporate in the formalism the empirical observation that electrons in bound potentials appear to subtend a discrete spectrum of states. This property can be accounted for in the formalism by replacing the 'label' (x_0, v_0) in Eq. (2) by a discrete index, or *quantum number* $n = 0, 1, 2, ...$, to be understood as the information that 'the electron is in the ground state ($n = 0$), or in the states $n = 1$, $n = 2$,' etc. The n–th state corresponds to a value of the energy greater than the energy of the ground state, according to the above-mentioned postulate (ii) of the old quantum theory.

The second change with respect to the classical formula (2) has to do with the quantum interpretation of the quantity $k\omega$. Classically, to every state (x_0, v_0) there corresponds a constant energy value $E(x_0, v_0)$, and also, a constant value of the *action*

$$J(x_0, v_0) = \frac{1}{2\pi} \oint_{C(x_0, v_0)} v dx \ . \tag{3}$$

The integral in (3) is carried over the whole periodic trajectory C with initial conditions (x_0, v_0) (for simplicity we consider the particle's mass equal to 1). One has that the energy is a function only of J. Furthermore, $\omega = dE/dJ$. In the quantum regime, however, we adopt the postulate that the action and the energy can only vary by discrete quanta. Bohr's old theory was postulating that the action variable can only vary by multiples of *Planck's constant* $\hbar$, called the *quantum of action* by Sommerfeld. Based also partially on Bohr's idea, Heisenberg's reasoning now goes as follows: we shall re-interpret both differentials in the classical equation $\omega = dE/dJ$. Thus, for an electron in the n–th quantum state, we replace dE by a *finite difference* ΔE, whose only allowed values are those corresponding to transitions from the n–th state to adjacent states $n - k$, $k = \pm 1, \pm 2, \ldots$. Thus, the permissible values are $\Delta E_{n,k} = E_n - E_{n-k}$, $k = \pm 1, \pm 2, \ldots$. In a similar way, we replace the differential dJ by the finite differences $\Delta J_k = k\hbar$, i.e., by multiples of the quantum of action $\hbar$. Then, the classical quantity $k\omega$ is replaced by

$$k\omega = k\frac{dE}{dJ} \quad \rightarrow \quad k\frac{\Delta E_{n,k}}{\Delta J_k} = \frac{E_n - E_{n-k}}{\hbar} \equiv \omega_{n,n-k} \ . \tag{4}$$

Substituting also the continuous 'state' label (x_0, v_0) with a discrete one $(x_0, v_0) \rightarrow n$, leads to

$$X_k(x_0, v_0) \rightarrow X_{n,n-k} \, . \tag{5}$$

One is now tempted to substitute the corresponding rules (4) and (5) in Eq. (2), thus obtaining a rule for the time evolution of $x_n(t)$, i.e., the presumed quantum analogue of $x(t; (x_0, v_0))$. However, Heisenberg notes that such a substitution "seems to be impossible in a unique manner and therefore makes no sense". Heisenberg's argument is essentially an algebraic one: if, after such a substitution, one regarded Eq. (2) as the quantum version of the Fourier transform of $x_n(t)$, computing $x_n(t)^2$, or, say, $x_n(t)y_n(t)$ (where $y(t)$ is a second position variable) would give results inconsistent with the so-called 'rule of combination of frequencies', i.e.

$$\omega_{n-k,n-k-l} + \omega_{n,n-k} = \omega_{n,n-k-l}$$

which follows directly from the definition (4). Thus, Heisenberg abandons the idea that it is possible to define $x_n(t)$. Instead, he considers the *entire* ensemble of the quantities

$$(6) \quad x_{n,n-k}(t) = X_{n,n-k} e^{i(E_n - E_{n-k})t/\hbar}, \quad n = 0, 1, 2, ..., k = \pm 1, \pm 2, ...$$

as a representation of the quantum variable $x(t)$. As explained in [32], in subsequent computations there appears the need to consider also a representation of quantities involving, e.g. the square or higher powers of $x(t)$. Then Heisenberg proposes that $x(t)^p$, $p = 2, 3, ...$ should also be understood as a collection of quantities $[x(t)^2]_{n,n-k}$, $n = 0, 1, 2, ...$, $k = \pm 1, \pm 2, ...$ Heisenberg proposes a rule for computing, for example, the quantities $[x(t)^2]_{n,n-k}$, in terms of the quantities $x_{n,n-k}(t)$, in such a way so as to respect the frequency combination rule. Although not explicitly recognized by Heisenberg, his proposed rule was equivalent to *matrix multiplication*. In other words, his rules imply that the quantities $x(t)$, $x(t)^2$, etc. should be regarded as *matrices*. ([5])

Note that in Eq.(6) k can in principle take any value from $-\infty$ to $+\infty$. However, as a consequence of the requirement that there be a ground state, it can be shown that some matrix elements $X_{n,n-k}$ in (6) must be set equal to zero, for k outside certain limiting values de-

([5]) In [32], Heisenberg makes no use of the term 'matrix' (or array), while, this term explicitly appears in the subsequent 'explanatory' paper of Born and Jordan [9]. On the other hand, in [32] it is also clear that Heisenberg's calculational scheme was still focusing on the idea of grouping together (and determining, for fixed n) the subsets X_n of matrix elements of $x(t)$ defined via

$$X_n = \{x_{n,n-k}, k = 0, 1, 2, ...\}.$$

The key idea here is that, for fixed n, it is only the matrix elements of the set X_n which will appear in *observables* (i.e. transition amplitudes and frequencies) related to transitions from the n–th state to lower energy states. However, Heisenberg realized that his calculational scheme implied that determining one subset X_n, for a specific value of n, requires to solve relations involving also matrix elements of *different* values of n. Hence the whole set of these relations have to be solved together. Using matrix notation, it turns out that these relations can be written in a unified way, thus yielding finally only one equation involving the matrix variable $x(t)$ and its powers (see, for example, subsection 2.2 below). The beautiful result is that this latter equation formally appears like a classical kinematic equation for $x(t)$.

pending on n. However, one can also remark here that the very existence of a ground state for an arbitrary model has to be either a priori postulated, or shown to be consistent with elementary algebraic properties of the matrix equation of motion of the model considered. ([6])

In modern terms, Heisenberg's $x(t)$ corresponds to the definition of the so-called position 'operator' (quantum variable) in what we nowadays call the quantum 'Heisenberg picture', and in a particular representation (or 'basis') called 'energy representation'. The term Heisenberg picture refers to the picture of quantum mechanics in which state vectors (wavefunctions) do not vary in time, while the time evolution of a system is carried exclusively by its dynamical variables or 'observables' (position, momentum, etc.). It is well known that, mathematically, the Heisenberg picture is equivalent to the more familiar Schrödinger picture, in which the wavefunction varies in time (through Schrödinger's equation), but the definitions of position, momentum, etc., operators are fixed. Finally, the term energy representation means that the label n (or $n - k$) in Eq.(6) refers to states of a constant energy, i.e., to energy *eigenstates*. As will be shown below, the energy values E_n can be computed in an unambiguous way in Heisenberg's picture using the matrix formalism. On the other hand, the most familiar derivation of the energies E_n is in the framework of the Schrödinger picture. Namely, the energies E_n are usually computed as the eigenvalues corresponding to a boundary value problem of the time-independent Schrödinger equation. ([7])

([6]) In the case of a harmonic oscillator, or of electrons' orbitals in atoms and molecules, several routes of reasoning lead to a demonstration of the existence and of the properties of a corresponding ground state. However, in quantum field theory, the properties of the ground state are not always obvious to establish. For example, a Millenium prize is offered by the Clay institute for rigorously establishing the properties of the ground state in Yang-Mills-type theories such as the theory of strong interactions of elementary particle physics.

([7]) See the Appendix for a brief definition of the Heisenberg and Schrödinger pictures of quantum mechanics.

2.2 – *The Poincaré - Lindstedt method*

The transition to the matrix formulation was an ingenious conceptual step in Heisenberg's work. However, Heisenberg extended the analogy between the classical and quantum formulations far beyond the level of formal definitions. In the same paper [32], dealing with the example of the anharmonic oscillator, he proposed a perturbative calculational scheme in order to determine the unknown quantities in the definition of the matrix elements of Eq.(6). A careful inspection of his calculations allows us to argue that his scheme can be regarded as a quantum version of a well known method of Celestial Mechanics, i.e. the method of *Lindstedt series*.

The Poincaré - Lindstedt method has been extensively reviewed in the literature (see, for example, [7],[37]). Referring to the example of the anharmonic oscillator, the method can be summarized as follows: One seeks to express the classical solution $x(t)$ as a series in the small parameter λ of the form:

$$x(t;A) = x^{(0)}(t;A) + \lambda x^{(1)}(t;A) + \lambda^2 x^{(2)}(t;A) + ... \tag{7}$$

The functions $x^{(r)}(t;A)$, $r = 0, 1, 2, ...$ are assumed to be polynomials, depending on powers of a certain quantity A, called the 'amplitude' of the oscillation, as well as on trigonometric quantities $\cos(k\omega t)$, $\sin(k\omega t)$ (with k integer), where ω is the fundamental frequency of the oscillation. The frequency ω is itself expressed as a series

$$\omega(A) = \omega_0 + \lambda\omega^{(1)}(A) + \lambda^2\omega^{(2)}(A) + ... \tag{8}$$

Eq.(8) expresses the fact that the frequency depends in general on the amplitude of the oscillation.

Two main approaches to the computation of Lindstedt series are the following:

i) In the simplest approach, one sets at the beginning a fixed value of the amplitude A, which is associated with a given choice of initial conditions. This allows to define the function $x^{(0)}(t;A)$. Then, the functions $\omega^{(r)}(A)$ and $x^{(r)}(t;A)$ at subsequent steps $r = 1, 2, \ldots$ are determined recursively. To this end, we substitute Eqs.(7) and (8) to the equation of motion (e.g. Eq.(1)), and compare terms of equal order in λ.

ii) An alternative approach is to consider, initially, a fixed (pre-selected) value of the *frequency* ω instead of the amplitude A. Now, Eq.(8) is regarded as an equation for the unknown amplitude A, which generates a periodic motion with the given frequency. The equation (7) can again be solved recursively. However, in the equation (8) for the frequency, the terms $\omega^{(r)}$ are now determined by the requirement that no secular terms (i.e. terms of the form $t\cos k\omega t$, $t\sin k\omega t$) survive in the series (see, for example, [28]). In the end of the process, A is specified as a power series in the quantity $\Delta\omega = \omega - \omega_0$, starting with terms linear in $\Delta\omega$. The latter series is found by inversion of the series (8). Substitution of the result into (7) resumes the series computation of the orbit $x(t;A)$. ([8])

In his own series computations, Heisenberg gives no indication of awareness that a well known method of classical mechanics will be employed. In fact, instead of the general expression (7), he proposes a variant taking the form:

$$(9)\quad x(t;A) = a_1(A)\cos\omega t + \lambda[a_0 + a_2(A)\cos 2\omega t] + \lambda^2 a_3(A)\cos 3\omega t + \dots$$

It is easy to see that the usual (i.e. Eq.(7)) and Heisenberg's schemes become equivalent after a resummation of some groups of terms in the series (7) (see [1]; note also that the constant a_0 in Eq.(9) expresses the fact that the average value $\langle x(t;A)\rangle$, i.e. the center of the oscillation, appears shifted with respect to the origin when the total anharmonic potential has no even symmetry). The use of only cosines in the series (9) is without loss of generality, as it simply reflects a choice of the origin of time such that the initial velocity is zero at $t = 0$.

([8]) In books of mechanics one rarely finds an explicit distinction between the two above-mentioned different approaches to performing computations with Lindstedt series. In systems of one degree of freedom, the two approaches give qualitatively similar results, but in systems of more than one degrees of freedom they lead to quite distinct results. The former approach leads to results similar to the so-called 'Birkhoff normalization process'. As discussed below, the resulting Lindstedt series are non-convergent, and exhibit an asymptotic behavior. On the contrary, the latter approach can be viewed as a 'torus fixing' process that leads to results equivalent to the so-called Kolmogorov normalization algorithm (see the tutorial [25]). As shown by Eliasson [27], for particular (diophantine) frequency vectors, the resulting Lindstedt series exhibit cancellations of large terms, yielding eventually a convergent behavior (see also [13]).

Heisenberg's key proposal, now, is that one may use a form analogous to (9) *in the quantum-mechanical case as well*, in order to compute the matrix variable $x(t)$. The method used by Heisenberg could be called a *quantum Lindstedt series*. In particular, instead of a sum $x(t;A)$ over Fourier terms (as in (9)), Heisenberg considers a set $X_n(t)$ formed by a union of particular matrix elements of $x(t)$ according to

$$X_n(t) = \{a_{n,n-1}\cos\omega_{n,n-1}t\}\cup \lambda\{a_{n,n},\ a_{n,n-2}\cos\omega_{n,n-2}t\}\cup\lambda^2\{a_{n,n-3}\cos\omega_{n,n-3}t\}\cup\dots \tag{10}$$

Comparing the expressions (9) and (10) one notices the following: i) the continuously-valued amplitude A in the classical series (9), which serves as a label of a particular 'state' of motion, is substituted in (10) by a discrete label, namely the quantum number n. Thus, all the amplitudes and frequency matrix elements become functions of n. ii) The classical frequency multiples $k\omega$, $k=1,2,...$ are replaced by the frequency matrix elements $\omega_{n,n-k}$. As discussed previously in this section, the frequencies $\omega_{n,n-k}$ correspond to the frequency of emitted radiation $\omega_{n,n-k}=(E_n-E_{n-k})/\hbar$ at the transition from the state n to the state $n-k$.

Heisenberg proceeds now by presenting specific computations of his series in the Duffing model with cubic and quartic anharmonic terms. It is remarkable that he proceeds showing in parallel the classical with his new quantum Lindstedt series. In the latter, all multiplications of matrix elements arising by nonlinear terms in the equation of motion are performed via standard matrix multiplication.

In the cubic case, Heisenberg then shows that the recursive relations of his method lead to:

Classically

$$\omega_0^2 a_0+\frac{a_1^2}{2}=0$$

$$(-4\omega^2+\omega_0^2)a_2+\frac{a_1^2}{2}=0$$

$$(-9\omega^2+\omega_0^2)a_3+a_1a_2=0$$

Quantum – theoretically

$$\omega_0^2 a_{n,n} + \frac{(a_{n+1,n})^2 + (a_{n,n-1})^2}{4} = 0$$

$$(-(\omega_{n,n-2})^2 + \omega_0^2) a_{n,n-2} + \frac{a_{n,n-1} a_{n-1,n-2}}{2} = 0$$

$$(-(\omega_{n,n-3})^2 + \omega_0^2) a_{n,n-3} + \frac{a_{n,n-1} a_{n-1,n-3}}{2} + \frac{a_{n,n-2} a_{n-2,n-3}}{2} = 0 \ .$$

Note that the information on the initial state, i.e., the starting point of the algorithm, is the value of the amplitude a_1, in the classical case, and the corresponding value of $a_{n,n-1}$ in the quantum case. Classically, we have $a_1 = A = (2J/M\omega_0)^{1/2}$, where the variable J (which can be used equivalently as a free parameter in the place of A) represents the value of the action variable in the harmonic oscillator limit. In the 'old' quantum theory we would thus simply assign the quantized values $J = n\hbar$ to the action variable. In looking for an appropriate generalization, Heisenberg remarks that the definition (3) leads to

$$1 = \frac{1}{2\pi} \frac{d}{dJ} \oint_{C(x_0,v_0)} \dot{x}^2 dt \ , \tag{11}$$

where, classically, $\dot{x}(t)$ is found by differentiation of (2) with respect to the time t. Once again, he postulates the existence of a finite difference version of Eq.(11), which he calls the 'additional quantum condition'. Replacing the differential dJ with $\Delta J = J_{n+k} - J_{n-k} = 2k\hbar$, his proposed 'quantum condition' reads:

$$\hbar = 4\pi \sum_{k=0}^{\infty} |X_{n,n+k}|^2 \omega_{n,n+k} - |X_{n,n-k}|^2 \omega_{n,n-k} \, . \tag{12}$$

Albeit not immediately recognized, Eq.(12) would be identified later by Born and Jordan in [9] as equivalent to a basic postulate of quantum mechanics, namely that $xp - px = i\hbar$, where p is the quantum (matrix) momentum variable.

Returning to the series computation, by Eq.(10) one sees that only the coefficients $X_{n,n+1}$, $X_{n,n-1}$ are of order zero in λ. Thus, at order zero, the quantum condition (12) leads to a difference equation for $X_{n,n-1}$, or, equivalently, $a_{n,n-1}$. Solving this equation, Heisenberg finds

$a_{n,n-1} = (2\hbar n/M\omega_0)^{1/2}$ (for a non-unity mass M). This last step resumes his quantum series computation.

An important product of the 'quantum Lindstedt series' is that a theoretical formula for the energy of the n–th state can be computed to all orders in λ. This is realized by substituting the quantum matrix variable $x(t)$, with matrix elements given by (10) for *all* $n = 0, 1, 2, ...$, as well as its time derivative (found by direct differentiation) into the equation for the energy

$$E = \frac{1}{2}M\dot{x}^2 + \frac{1}{2}M\omega_0^2 x^2 + M\lambda\frac{x^b}{b} \quad , \tag{13}$$

where M is the particle's mass. Since x is a matrix variable, E is also a matrix quantity. After substituting the solutions for the matrix elements of (10), one discovers that the time disappears by identity to all orders in λ.([9]) This fact expresses the conservation of energy, which holds in the quantum case as well for energy eigenstates. Finally, one is left with an expression for E which contains only *diagonal* non-zero elements E_n, with $n = 0, 1, 2, \ldots$. This latter fact leads Heisenberg to identify E_n as the energy associated with the n–th state. This is a function of the quantum number n. Heisenberg finds

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega_0 + O(\lambda^2) \tag{14}$$

in the cubic case, while

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega_0 + \frac{3\lambda(n^2 + n + 1/2)\hbar}{8\omega_0^2 M} + O(\lambda^2) \tag{15}$$

in the quartic case. These expressions, together with the corresponding ones for the transition probabilities between different states, allow a complete characterization of the emission spectrum of the systems under study, i.e., of quantities amenable to experimental observations.

([9]) Heisenberg writes: "I could not prove in general that all periodic terms actually vanish, but this was the case for all terms evaluated." Here again, it is remarkable how intuition leads Heisenberg to adopt a result for which no rigorous justification was available at the moment.

3. – Classical and Quantum Lie Series

In [9], Born and Jordan repeated the calculations of Heisenberg, by simultaneously giving many more details on how the matrix structure of canonical variables should be understood within the context of the newly emerging quantum theory. However, in the three-man paper [10], the authors make one further drastic step: they fully introduce quantum perturbation theory computations within the framework of the Hamiltonian, or *canonical* formalism.

The details of how to introduce canonical, i.e., 'unitary' transformations of quantum variables are explained in several books of quantum mechanics (see, for example [39]). On the other hand, as will be argued below, the original method of performing perturbation theory computations by BHJ has been set rather in the margin in later years. A quite distinct method yielding equivalent results has become dominant in the literature, called the *Rayleigh-Schrödinger* method.

Nevertheless, the original BHJ method was essentially re-discovered by a number of authors in more recent years (see below). In this respect, the following remark is worth making: similarly to the case of Heisenberg's quantum analogue to the classical Lindstedt series, the BHJ method can be considered as the quantum analogue of a formal apparatus well known in celestial mechanics, i.e. the method of *normal form* obtained via *Lie series*. ([10])

([10]) Overall, the fact that BHJ start from a number of methods of classical mechanics, which, then, they generalize in quantum mechanics, is a clear indication of the strong background that this group had in classical perturbation theory. Most probably, Born was the leading figure behind this approach. Born's understanding of the formal structure of mechanics was renown among his contemporary physicists (and reflected, for example, in his book 'Atomic Physics' [8]). The influence that perturbative methods had in the overall development of quantum mechanics cannot be overemphasized. In fact, the whole belief of physicists, that an operator-valued mathematical description of mechanical observables was making sense, was strongly consolidated by the various successes of this approach; the latter were largely due, in turn, to the successful use of the quantum perturbative techniques. The story was repeated in the case of quantum field theory. There, the power of perturbative techniques came to recognition after the invention of the calculus based on so-called Feynman diagrams. The equivalence between this calculus and the canonical one (introduced by Dirac, Schwinger and Tomonaga) was demonstrated by F. Dyson [21].

In order to justify the above remarks, let us note the following points:

3.1 – *Hamiltonian normal form*

The classical Lindstedt series of celestial mechanics can be viewed as a *direct* method of perturbation theory, i.e., a method in which one seeks series solutions, like the series (7) and (8), which satisfy directly the equations of motion, like, for example, Eq.(1).

In the canonical formalism, however, the equations of motion are, themselves, regarded as the components of a symplectic flow generated by a particular function of the phase space variables, called the Hamiltonian function. In that sense, the Hamiltonian function carries already the entire information about the equations of dynamics.

One can, now, note the following: for every direct (i.e. based on the equations of motion) method of canonical perturbation theory, there can be constructed an equivalent *indirect* method, which leads to the same perturbative solutions, but is based, instead, on a transformation-of-variables process, implemented to the Hamiltonian function. This is called *Hamiltonian normalization* (or Hamiltonian normal form).

Consider a system of n degrees of freedom with initial Hamiltonian function $H(q,p)$, where (q,p) are n-dimensional canonically conjugate variables (position and momentum). In the normal form method, we seek to find a sequence of canonical transformations

$$(q,p) \equiv (q^{(0)},p^{(0)}) \to (q^{(1)},p^{(1)}) \to (q^{(2)},p^{(2)}) \to \dots \tag{16}$$

such that, after r steps, the Hamiltonian, expressed in the new variables, is decomposed in the form

$$H^{(r)}(q^{(r)},p^{(r)}) = Z^{(r)}(q^{(r)},p^{(r)}) + R^{(r)}(q^{(r)},p^{(r)})\,. \tag{17}$$

The first term $Z^{(r)}$, called the *normal form*, represents a Hamiltonian with an easier-to-study dynamics than in the original Hamiltonian model. For example, $Z^{(r)}$ can have what we call an *integrable* form, or, alternatively, a form allowing to deduce the existence of one or more

particular symmetries of our model. The second term $R^{(r)}$, called the remainder, expresses now the difference between the true dynamics and the dynamics based on the approximation by $Z^{(r)}$ alone. Let us emphasize here that since the total Hamiltonian $H^{(r)}(q^{(r)}, p^{(r)})$ is just a re-expression of the original one in new variables, all conclusions drawn upon the dynamical behavior of the system in the new variables $(q^{(r)}, p^{(r)})$ can be translated to conclusions regarding the dynamical behavior of the system in the original variables.

In most perturbative schemes, the successive normalization steps $r = 1, 2, \ldots$ are solved by recursion. The normalization process is called convergent when the size of the remainder $R^{(r)}$ goes to zero as r tends to infinity. The convergence is crucially affected by the appearance of *small divisors*, i.e. small quantities dividing the successive series terms, whose appearance is inevitable in systems with oscillatory behavior. In fact, the way by which small divisors accumulate in various perturbative schemes determines whether a particular scheme proves eventually to be (or not) convergent (see [25] for a tutorial introduction).

3.2 – *The von Zeipel method*

How do we perform the sequence of canonical transformations (16)? Several methods have been developed in Celestial Mechanics over centuries. A commonly employed method at the time of BHJ was the so-called *von Zeipel* method [46]. This method relies upon the definition of a so-called *generating function* S of a canonical transformation. Let $(q, p) \to (Q, P)$ be a transformation from one set of canonical variables to another. Then, S is a function of the form $S(q, P)$, i.e. depending on the position variables q of the first set and on the momentum variables P of the second. The transformation equations read:

$$Q = \frac{\partial S(q,P)}{\partial P}, \quad p = \frac{\partial S(q,P)}{\partial q}\,. \tag{18}$$

Returning to the Hamiltonian normal form, in the von-Zeipel method we perform a sequence of transformations of the form (18), by

appropriately choosing the form of S in every step. The aim is that, in the final set of variables, the Hamiltonian should take a desired form. For example, if (p, q) are action-angle variables, one option is to look for transformations eliminating the angles from the final Hamiltonian. This allows to obtain an integrable approximation to the original system, since the final Hamiltonian depends only on the actions.

The von Zeipel method was introduced originally in celestial mechanics. However, the method presents a computational disadvantage related to the form of the transformation (18). We can observe that the transformation is *implicit*, i.e., one must solve for the variable P (in terms of q and p), in order to explicitly express all the new canonical variables in terms of all the old ones (or vice versa). Accomplishing this task necessitates, in general, performing a series reversion, which is a cumbersome process.

Despite these difficulties, with the advent of modern computers the von Zeipel method played a key role in the development not only of celestial mechanics but also of the classical theory of dynamical systems in general. For example, it was the method employed in the first computer-algebraic symbolic computations of Hamiltonian normal forms in models like the celebrated *Hénon-Heiles* [33]:

$$H = \frac{1}{2}(p_x^2 + p_y^2 + x^2 + y^2) + \varepsilon\left(x^2 y - \frac{1}{3}y^3\right). \tag{19}$$

The system (19) has served as an archetypical nonlinear Hamiltonian dynamical system of two degrees of freedom. It physically corresponds to two oscillators with a nonlinear coupling. In 1960, Contopoulos [14] had computed approximate integrals of motion in a similar system using a 'direct' method called the 'third integral'. The higher order terms of the third integral were calculated by Contopoulos using a computer-algebraic method [15]. Then, Gustavson [29] used a computer-generated algorithm of the von Zeipel method, and successfully computed formal integrals of motion in the model (19). These integrals were by-products of a resonant normal form computed by the von Zeipel method. In the case of coupled oscillators, the method had been previously explored theoretically by Birkhoff [6], and it is known as the

method of 'Birkhoff series'. These early computations of Contopoulos, Hénon-Heiles and Gustavson established the role of regular (quasi-periodic) orbits in nonlinear systems of coupled oscillators. The importance of these works lies in that they corroborated the original point of view of Poincaré, that systems with non-linear couplings should not, in general, be completely ergodic, as some degree of *order* (i.e. regular motions) should co-exist along with some degree of *chaos* (see [16]). In fact, chaos becomes strong in the domain where the normal form series computations fail to practically converge. ([11])

3.3 – *Lie series in classical mechanics*

The von Zeipel method has been superceded in recent years by a more powerful method of determination of normal form canonical transformations, based on *Lie series* [34],[18]. The method of Lie series stems from an elementary remark, proved by Poincaré [42]: the Hamiltonian flow $\dot{q} = \partial\chi/\partial p$, $\dot{p} = -\partial\chi/\partial q$ under any arbitrary function $\chi(q,p)$ generates a canonical transformation by the corresponding time evolution of the canonical variables (q,p). This means that, starting from some initial condition $q = q(0), p = p(0)$, and computing $q_t = q(t)$, $p_t = p(t)$ under the flow of χ, the mapping $(q,p) \to (q_t, p_t)$ is a canonical transformation. The time t serves here as a parameter of the transformation.

([11]) It is well known that the Birkhoff series are, in general, non-convergent. However, they can be useful in practice, since they exhibit an *asymptotic* behavior: the remainder at the r–th order has a size $r!^a\varepsilon^{br}$, where a, b are positive constants and ε represents the series's small parameter. This generic behavior, predicted already by H. Poincaré [42], implies that up to a maximum order $r_{max} \sim \varepsilon^{-b/a}$ the series appear as 'pseudo-convergent'. Thus, a finite truncation at the order $r = r_{max}$ allows to obtain a quite useful representation of regular orbits. A similar phenomenon appears in quantum field theory. Namely, the perturbative computations of the S-matrix based on Feynman's diagrams can be shown to be asymptotic [22] (the role of the small parameter ε is played here by the fine structure constant). As a result, the so-called (by Feynman) 'jewel of physics' computations (e.g. the Lamb shift or the anomalous magnetic moment of the electron), albeit precise up to eleven digits or so, are still based on a finite truncation of an otherwise asymptotic series. Founders of the quantum field theory like F. Dyson were considering the non-convergence of the QFT series, even after re-normalization, as the most unsatisfactory feature of standard QFT [23][38].

Setting $(Q,P)=(q(0),p(0))$, taking Hamilton's equations $\dot{q}=\{q,\chi\}$, $\dot{p}=\{p,\chi\}$, expanding $q(t)$ and $p(t)$ in Taylor series around $t=0$, and setting finally $t=1$, one arrives at a formula for the so-called *Lie series* canonical transformation $(Q,P)\rightarrow(q,p)$

$$q=\exp(L_\chi)Q,\quad p=\exp(L_\chi)P \tag{20}$$

where $L_\chi=\{\cdot,\chi\}$ denotes the Poisson bracket operator. The function χ, in this context, is called a Lie generating function (see [25] for a detailed introduction to Lie series).

The main advantage of the Lie method over the von Zeipel method is that the transformation equations (20) are explicit, thus necessitate no series inversion as in the case of Eqs.(16). As a result, the method of Lie series offers great algorithmic simplicity. In fact, the following property of the Lie series can be readily shown: Let $F(q,p)$ be an arbitrary function of the canonical variables (q,p), and $F'(Q,P)=F(q(Q,P),p(Q,P))$ its image under the canonical transformation (20). Then, F' can be found directly by

$$F'=\exp(L_{\chi(Q,P)})F(Q,P)\,. \tag{21}$$

Eq.(21) is particularly easy to implement, since it only involves the computation of derivatives and avoids any function composition, which is a cumbersome procedure.

Due to the above advantages, the use of Lie series as proposed in [34] and [18] has nowadays become nearly standard in implementations of classical perturbation theory. However, as we will see now, an equivalent scheme of perturbation theory, that we hereafter call 'quantum Lie series', was proposed by BHJ in the three-man paper of quantum mechanics [10].

3.4 – *Lie series in quantum mechanics*

In order to motivate the comparison between classical and quantum Lie series, let us consider first how we would express in modern terms the quantum evolution of a system with an arbitrary 'Hamiltonian' (i.e. Lie generating function) $\chi(q,p)$.

In the familiar Schrödinger picture (see Appendix), the time evolution is carried by the state vector $|\Psi(t)\rangle$, which obeys Schrödinger's equation

$$i\hbar\frac{d|\Psi(t)\rangle}{dt} = \chi|\Psi(t)\rangle\,. \tag{22}$$

In Eq.(22), χ is meant now as a 'matrix' quantity, i.e. an operator composed by the operators q and p. The formal solution of (22), viewed as an initial value problem with initial state $|\Psi(0)\rangle$, is

$$|\Psi(t)\rangle = \exp\left(-\frac{i\chi t}{\hbar}\right)|\Psi(0)\rangle\,. \tag{23}$$

The operator

$$S_\chi(t) = \exp\left(-\frac{i\chi t}{\hbar}\right) \tag{24}$$

is called the 'unitary evolution operator' under the Hamiltonian $\chi(q,p)$. Consider now, as in the classical case, an arbitrary function of the phase-space variables $F(q,p)$. The function F may represent a mechanical quantity like the energy, angular momentum, etc. Since (q,p) are operator-valued, F is also operator-valued. This should be understood as the statement that, in the quantum regime, we can only assign a quantum probability for F being measured to one of its eigenvalues at any fixed time t. In the Schrödinger picture, this probability evolves in time depending only on the time evolution of the state vector $|\Psi(t)\rangle$. The probability distribution $P(f)$ that F is measured at the value f can be determined unambiguously once all the moments of $P(f)$ are known. According to quantum mechanics (see Appendix), the k–th moment

$$\langle f^k\rangle = \int_D f^k P(f)df$$

(where D means the support of $P(f)$) is given by

$$\langle f^k\rangle = \langle\Psi(t)|F^k|\Psi(t)\rangle\,. \tag{25}$$

Taking into account Eq.(23), we have, for example, for the mean value of f (i.e. $k = 1$):

$$\langle f \rangle = \langle \Psi(t)|F|\Psi(t)\rangle = \langle \Psi(0)| \exp\left(\frac{i\chi t}{\hbar}\right) F \exp\left(-\frac{i\chi t}{\hbar}\right) |\Psi(0)\rangle \,. \tag{26}$$

In the Heisenberg picture, now, we consider state vectors as non-varying in time, and we assign the time evolution to the operator-valued quantity F itself. Thus, we re-interpret Eq.(26), by setting $|\Psi(t)\rangle = |\Psi(0)\rangle = |\Psi\rangle$ (time invariant), while, the original definition of F is now set as an *initial condition* $F \equiv F(0)$, taking F to evolve in time according to

$$F(t) = \exp\left(\frac{i\chi t}{\hbar}\right) F(0) \exp\left(-\frac{i\chi t}{\hbar}\right) . \tag{27}$$

Then, since $|\Psi(t)\rangle = |\Psi(0)\rangle = |\Psi\rangle$, Eq.(25) takes the equivalent form:

$$\langle f^k \rangle = \langle \Psi(0)|F(t)^k|\Psi(0)\rangle = \langle \Psi|F(t)^k|\Psi\rangle \,. \tag{28}$$

Hence, $P(f)$ arises the same in the Heisenberg and Schrödinger pictures, i.e., the two pictures are equivalent regarding their empirical predictions.

Which differential equation of motion should the operator $F(t)$ satisfy in order that the solution be given by Eq.(27)? It can be shown (see, for example [39]) that Eq.(27) represents the formal solution of the initial value problem for the operator-valued differential equation

$$\frac{dF}{dt} = \frac{1}{i\hbar}[F, \chi] \quad , \tag{29}$$

where $[F, \chi]$ is the *commutator* $F\chi - \chi F$. Eq.(29) is called *Heisenberg's* equation of quantum motion.

The key remark is now the following: if we define the quantum operator

$$\mathcal{L}_\chi = \frac{1}{i\hbar}[\cdot, \chi] \tag{30}$$

and set, as in the classical case, $t = 1$, we can view Eq.(27) as generating a (quantum) canonical (i.e. 'unitary') transformation $F \to F'$ given

by the formal solution $F' = F(t)$ of (29) for $t = 1$, i.e.:

$$F' = \exp(\mathcal{L}_\chi) F = \mathcal{S}_\chi F \mathcal{S}_\chi^{-1} \quad , \tag{31}$$

where $\mathcal{S}_\chi = S_\chi(t = 1)$. We note the formal equivalence between the 'quantum' (Eq.(31)) and the classical (Eq.(21)) transformation equations. In particular, implementing Eq.(31) to the quantum variables $(Q, P) \to (q, p)$, we arrive at the expressions

$$\begin{aligned} q &= \exp(\mathcal{L}_\chi) Q = \mathcal{S}_\chi Q \mathcal{S}_\chi^{-1} \\ p &= \exp(\mathcal{L}_\chi) P = \mathcal{S}_\chi P \mathcal{S}_\chi^{-1} \, . \end{aligned} \tag{32}$$

However, the last term in Eq.(32) is *formally identical to the definition of canonical transformations as given by BHJ* in ([10]) (see their equation (17)). Via the operator $\mathcal{L}_\chi$, we see that this definition is actually a quantum Lie series. The whole section 3 of the three-man paper is devoted to the implementation of such Lie series in order to perform computations in quantum canonical perturbation theory. This has preceded the classical implementation of Lie series by about 40 years. ([12])

([12]) It is worth noting that the use of the Lie method for computing a quantum normal form by BHJ as early as in 1926 passed nearly unnoticed in later developments on the same subject. In fact, a different approach to the same problem was devised over the years and used, in connection mainly to problems of chemical dynamics. In this, so-called 'semi-classical' approximation, one first computes a classical (Birkhoff-Gustavson) normal form, as in Eq.(17). One then quantizes the function Z, thus obtaining a quantum model corresponding to the classical normal form. The semi-classical approach has been used extensively since the 70's (see for example [43] and references therein). Its two main problems, compared to the original quantum normal form approach of BHJ, are: i) the method lacks a clear theoretical justification, and its validation necessitates a posteriori comparison with numerical results (as e.g. the numerical determination of the spectrum of a system), ii) operator-ordering problems appear in the process of quantizing the normal form function Z [43]. This occurs even when such problems do not exist in the original Hamiltonian. In order to overcome such problems, several authors proposed a 'quantum normal form and its equivalents' [3], or the 'Birkhoff-Gustavson normal form in (classical and) quantum mechanics' [24], or the 'proper quantum analogue of the Birkhoff-Gustavson normal form' [17], (see also [41]). However, these works propose essentially the same quantum normal form algorithm as originally proposed by BHJ [10].

In their analysis, BHJ give all the basic steps regarding their scheme of quantum perturbation theory. Transcribing them to our notation, given a Hamiltonian function of the form

$$H = H_0 + \lambda H_1 + \lambda^2 H_2 + ... \tag{33}$$

one seeks to determine a sequence of functions χ_1, χ_2, ..., or, equivalently, $\mathcal{S}_1$, $\mathcal{S}_2$, ..., such, that, after implementing the corresponding canonical transformations (32), the Hamiltonian in the new variables (Q, P) takes a normal form

$$Z = Z_0 + \lambda Z_1 + \lambda^2 Z_2 + ... \tag{34}$$

with $Z_0 = H_0$. In order to determine the functions S_r, Z_r, $r = 1, 2, ...$, BHJ give the general equation of determination of the r-th order terms (their Eq.(25)):

$$[\mathcal{S}_r, H_0] + F_r(H_0, ..., H_r, \mathcal{S}_0, ..., \mathcal{S}_{r-1}) = Z_r \quad , \tag{35}$$

where F_r are quantities defined recursively. Note that here as well all quantities are meant as matrices. Eq.(35) is an exact quantum analogue of what is called, in the classical case, the *homological equation* (see [25]). In chapter 2 of [10], BHJ discuss extensively the case of many degrees of freedom, in which, they show that the solutions of Eq.(35) generate *small divisors* in the successive series terms. Furthermore they discuss separately the cases of 'non-degenerate' or 'degenerate' systems, i.e., systems without or with resonances. This chapter ends with a discussion about the convergence issues, related to the presence of small divisors in the series. The authors leave the question of the convergence open. In their words, they aimed at "solving the basic quantum theoretical equations in a manner as closely parallel to classical theory as possible".

4. – Classical and Quantum (near-)integrability

In his seminal work *Les Méthodes Nouvelles de la Mécanique Céleste* (1892), H. Poincaré [42] defined the following as constituting *"the fundamental problem of dynamics"*: study of the motion in a Hamiltonian system, expressed in action angle variables (I, ϕ), of the

following form:

$$H(\phi, I) = H_0(I) + \varepsilon H_1(\phi, I; \varepsilon) \quad , \tag{36}$$

where the function H_1 is analytic in the 'small' parameter ε, i.e., H_1 admits a convergent series expansion in powers of ε. In modern Hamiltonian nonlinear dynamical systems theory, such systems are characterized as 'nearly integrable'. Understanding their behavior has played a key role in the development of Celestial Mechanics, but also of atomic or molecular dynamics, the dynamics of lattices, accelerators, the motions of satellites, space physics, and the physics of astronomical systems like stars and galaxies. The use of action-angle variables refers to practically every system in which the motions exhibit some degree of periodicity, or, in the case of many degrees of freedom, multiple-periodicity (i.e. motion with more than one non-commensurable frequencies).

One of the main insights to the classical dynamical behavior of systems of the form (36), offered by the work of Poincaré himself, was to realize that, in such systems, the ordered (i.e. multiply-periodic) motions co-exist, in general, with chaotic motions. In the middle of the 20th century, the celebrated Kolmogorov-Arnold-Moser (KAM) theorem ([36], [4], [40]) established rigorously the proof that multiply-periodic motions exist, but their location in the phase space forms a Cantor set of initial conditions, whose measure is $1 - \mathcal{O}(\varepsilon)$, while its complement forms the web of chaotic motions. The interplay between order and chaos in such systems (called of 'mixed' phase space) has been one of the most expanded domains of research in dynamical systems in the last five decades.

What happens in the quantum realm regarding the question of order or chaos in systems of the form (36)? One may make immediately the following remark: from the very definition of the matrix position variables as in Eq.(6) (and similarly for the momentum variables), we see that, in the Heisenberg picture, we seek solutions (i.e. 'trajectories') which (for bound systems) should be *by definition multiply-periodic*. Indeed, the matrix function $x(t)$ has a structure depending only trigonometrically on time via the set of frequencies $\omega_n = E_n/\hbar$. Generically, the frequencies ω_n satisfy no full set of commensurabili-

ties (i.e. a number of independent rational relations among the frequencies, equal to the number of degrees of freedom minus one). Thus, the resulting motion is quasi-periodic. Let us emphasize that a choice of definition as in Eq.(6) is not arbitrary, but nearly dictated by the discrete nature of the spectrum in bound systems, which is an experimental fact. Indeed, as summarized in the previous sections, the essence of Heisenberg's quantum mechanics is to demonstrate that a definition of the position (and similarly, the rest of mechanical variables) as in Eq.(6), when substituted back into the equations of motion, leads to a selection rule for the energies E_n, i.e., to a discrete spectrum. Precisely the same result is arrived at in the Schrödinger picture by exploiting, instead, the boundary conditions for the wavefunction; in the case of bound systems, the requirement that the wavefunction should satisfy smooth vanishing conditions at the boundary leads to a unique determination of a complete and discrete set of energy eigenvalues and eigenfunctions of the time-independent Schrödinger equation. Also, in Schrödinger's picture the dynamical evolution is encoded in the time evolution of the wavefunction. The latter, however, is produced by the time-dependent Schrödinger equation, i.e., a linear differential equation whose solutions are manifestly multiply-periodic.

Similar remarks have led a number of authors to question the possibility to define 'quantum chaos' (see for example [5]). ([13]) In fact, the notion of 'quantum chaos' that currently prevails in literature refers to an ensemble of phenomenological effects manifested in quantum systems when the corresponding *classical* system is chaotic. Without entering into further analysis, let us just name some of these effects: i) level repulsion and the transition from Poisson to Wigner distributions of the differences of energy levels, ii) quantum 'scars', iii) decoherence of the Wigner function, iv) anomalous localization of the

([13]) Definitions of quantum chaos based on alternative interpretations of quantum mechanics also exist. For example, the so-called de Broglie - Bohm pilot wave theory allows to define quantum trajectories that can be either regular or chaotic (see [26] for a review). However, we presently do not examine such approaches, but restrict ourselves to a characterization of dynamics only in the 'orthodox' (i.e. Heisenberg or Schrödinger) picture of quantum mechanics.

maxima of the energy states etc. (see [30] for definitions and a review of the use of the above notions).

On the other hand, one may wonder whether there could be a sensible definition of 'quantum near-integrability' (as the opposite limit to quantum chaos), in a way analogous to classical near-integrability. In this respect, one may remark that, despite the above mentioned fact that the quantum representation of phase-space variables in bound systems is, by definition, multiply-periodic, the computation of the frequencies/energies $\omega_n = E_n/\hbar$ of Eq.(6) relies, in general, in a solution of an eigenvalue problem, which is a *nonlinear* problem. Such a solution is, in general, found by numerical means. However, the original Lindstedt series method of Heisenberg, as well as the canonical perturbative method of BHJ, provide the means to obtain an *analytical* determination of the energy levels in terms of *quantum numbers*. For example, in Eqs. (14) and (15), the energy levels of the anharmonic oscillator are represented as functions of a quantum number n. Thus, to the degree of approximation of the series expansions, one obtains a characterization of the spectrum in terms of one or more (depending on the number of degrees of freedom) quantum numbers. The importance of such characterization cannot be overemphasized. For example, the whole classification of chemical elements in the periodic table is based on the possibility to define 'orbitals', i.e. assign quantum numbers to the states of the outer electrons in the atom. ([14])

In modern applications, in the quantum perturbation theory use is made of a convenient set of quantum operators defined by

$$a = \frac{x + ip}{\sqrt{2}}, \quad a^+ = \frac{x - ip}{\sqrt{2}}, \tag{37}$$

where x is a position variable and p its conjugate momentum. These are known as the 'anihilation' and 'creation' operators, since it can be shown that their action on the n–th energy eigenstate produces the

([14]) Chemists speak in general about the posibility to characterize the spectrum of an atom or molecule using integer numbers as the property that the system possesses 'good quantum numbers'.

$n-1$ and $n+1$ energy eigenstates respectively. Their classical counterparts have been used in perturbation theory already by Birkhoff [6]. In terms of the creation and anihilation operators, the harmonic oscillator Hamiltonian takes the form

$$H_{osc} = \left(\frac{1}{2} + a^{+}a\right)\hbar\omega\,. \tag{38}$$

In a system of N oscillators coupled with nonlinear terms, the normal form of BHJ takes the form

$$Z = \sum_{i=1}^{N}\left(\frac{1}{2} + a_i^{+}a_i\right)\hbar\omega_i + Z_2(a_1^{+}a_1, a_2^{+}a_2, ..., a_N^{+}a_N) + Z_3(a_1^{+}a_1, a_2^{+}a_2, ..., a_N^{+}a_N) + \ldots \tag{39}$$

where Z_2, Z_3, etc. are of second, third, etc. degree in the products $a_i^{+}a_i$. The operators $\hat{n}_i = a_i^{+}a_i$ commute with Z, thus they constitute quantum integrals of motion. These are called number operators, and they are the quantum analogues of the classical action variables. A series representation of the energy spectrum is found by substituting, in the expression (39), the number operators by their corresponding eigenvalues, i.e., the quantum numbers. Thus, we find

$$E(n_1, n_2, ..., n_N) = \sum_{i=1}^{N}\left(\frac{1}{2} + n_i\right)\hbar\omega_i + Z_2(n_1, n_2, ..., n_N) + Z_3(n_1, n_2, ..., n_N) + \ldots \tag{40}$$

The expression obtained via Eq.(40) is identical to the one found in the Schrödinger picture via the Rayleigh-Schrödinger perturbation theory. However, Eqs. (39) and (40) can also be viewed as the quantum analogue of the classical Birkhoff normal form

$$Z = \sum_{i=1}^{N}\omega_i J_i + Z_2(J_1, J_2, ..., J_N) + Z_3(J_1, J_2, ..., J_N) + \ldots \tag{41}$$

In the classical case, in the near-integrable regime a finite truncation of the series (41) provides a good approximation to the dynamics in

terms of a set of good action variables. On the basis of this analogy, a quantum system can be called 'nearly-integrable' to the extent that a finite truncation of the series (40) provides an approximation of the energy spectrum in terms of a set of good quantum numbers.

Appendix: Basic mathematical definitions in Quantum Mechanics

Let f be a mechanical quantity (e.g. position, momentum, energy, etc.) of a certain physical system (e.g. a particle in a potential). Let A_f be a measuring apparatus able to determine the value of f in a concrete measurement of the system. Let $f_1, f_2, ...$ be the possible outcome values of the measurement. If, *after* passing through A_f, the system is measured at the value f_i, $i = 1, 2, ...$, we say that the system after the measurement is in the *state* $|f_i\rangle$.

According to Quantum Mechanics, before the measurement, the system can be in a state called *superposition* of more than one states $|f_i\rangle$. Mathematically, the most general state is written as

$$|\Psi\rangle = c_1|f_1\rangle + c_2|f_2\rangle + ... \tag{42}$$

where the c_i are complex numbers. The physical content of Eq.(42) is the following: if infinitely many replicas of the system are prepared in the state $|\Psi\rangle$, and each one of them is passed through the apparatus A_f, then the probability that the apparatus measures the value f_i is equal to $P(f_i) = |c_i|^2$. The quantity c_i is called the 'probability amplitude' of the system being in the state i.

The definition (42) allows to view the 'states' $|\Psi\rangle$ as vectors in a complex Hilbert space $\mathcal{H}$, called the space of state vectors. The vectors $|f_i\rangle$, $i = 1, 2, ...$ form a basis in $\mathcal{H}$.

We also consider linear transformations $\langle\Phi|: \mathcal{H} \to \mathcal{C}$, which act on the state vectors $|\Psi\rangle$ so that $\langle\Phi|\Psi\rangle$ is a complex number. Such transformations are called 'bra-vectors'. We define the elementary transformations $\langle f_i|$ by the rule $\langle f_j|f_i\rangle = \delta_{i,j}$. Furthermore, to every ket vector $|\Psi\rangle$ of the form (42) we associate a corresponding (dual) bra-vector, according to:

$$\langle\Psi| = c_1^*\langle f_1| + c_2^*\langle f_2| + ... \tag{43}$$

Since the sum of all probabilities should be normalized to unity, one has $|c_1|^2 + |c_2|^2 + ... = 1$. Then, from the definitions (42) and (43) it follows that $\langle \Psi | \Psi \rangle = 1$.

Consider now the linear Hermitian operator F: $\mathcal{H} \rightarrow \mathcal{H}$ which maps state vectors to state vectors according to the elementary rules:

$$F|f_i\rangle = f_i|f_i\rangle \quad . \tag{44}$$

Since every ket-vector can be written as a linear combination of the basis vectors $|f_i\rangle$, it follows that Eq.(44) defines completely the operator F. The operator F is then called the quantum (operator-valued) *observable* associated to the physical quantity f.

As an example, we can have a position operator X acting on the states $|x\rangle$ according to $X|x\rangle = x|x\rangle$. According to the above definitions, $|x\rangle$ is the quantum state, after the measurement, of a particle which, passing through a position-measuring apparatus, was measured at the position x.

The definitions (42), (43), and (44), although having a clear physical meaning, allow for no quantitative computations up to the point at which we define a so-called *representation* of state vectors. This is equivalent to assigning coordinates to a vector. We say that we work at the representation of the magnitude f, if we identify the vectors $|f_1\rangle$, $|f_2\rangle$, ..., with the column vectors $|f_1\rangle = (1, 0, 0, ...)^T$, $|f_2\rangle = (0, 1, 0, ...)^T$, etc. After such an identification, the state vector $|\Psi\rangle$ in Eq.(42) can by written as $|\Psi\rangle = (c_1, c_2, ...)^T$. One then easily finds that in the f–representation the dual bra-vector $\langle \Psi |$ corresponds to the line $(c_1^*, c_2^*, ...)$. Also, the operator F becomes a *matrix*, given by $F = diag(f_1, f_2, ...)$.

When studying a particular quantum system, we first have to fix a representation. For example, most elementary books of quantum mechanics deal in detail mostly with the *position* representation. In this representation, the position operator X is diagonal. Since the spectrum of possible measured position values $x_1, x_2, ...$ is continuous, the definition $X = diag(x_1, x_2, ...)$ does not strictly apply. Instead, following Dirac [19]) we can think of X as a "matrix with continuous indices". Similarly, in the definition of the state vector $|\Psi\rangle$, the sum of Eq.(42) should be substituted by

an integral

$$|\Psi\rangle = \int \psi(x)|x\rangle dx\,. \tag{45}$$

We observe that instead of the probability amplitude $c_i = \langle f_i|\Psi\rangle$ of Eq.(42), which depends on the discrete index i, here we have a continuous function $\psi(x) = \langle x|\Psi\rangle$. This is called the *wavefunction*. It should be noted, however, that we can define similar functions for all quantities f which vary continuously. For example, we can have a momentum wavefunction $\tilde{\psi}(p) = \langle p|\Psi\rangle$.

The main 'algorithm' for performing computations in quantum mechanics goes now as follows:

i) Choose a representation. This determines automatically the form of the operator F associated with the magnitude f whose representation was chosen.

ii) Express all other mechanical magnitudes of interest in the same representation. In this step, one first has to find the form of the operator G which corresponds to the physical magnitude g being canonically conjugate to f. For fixing G, one solves the *commutation relation* $FG - GF \equiv [F, G] = i\hbar I$, where I is the identity operator. The commutation relation can be viewed as an equation with F known and G unknown. After expressing G, one can then find the operator form of any other mechanical magnitude depending on (F, G).

Example: in the position representation, solve the equation $[X, P] = i\hbar I$. Answer: $P = -i\hbar d/dx$.

iii) Find the form of the Hamiltonian operator $H(X, P)$ in the chosen representation.

After these purely algebraic steps, one is ready to compute how evolve in time the probability amplitudes $c(v)$ to measure the system in a value v, where $v(f, g)$ is any quantity expressed in terms of the canonically conjugate quantities (f, g). To this end, one must be given (or assume) an initial state $|\Psi(0)\rangle$ for the system. Then one may proceed in two alternative ways, also called *pictures* of quantum mechanics:

I) *Schrödinger picture*: compute the operator $V(F, G)$. Compute the complete spectrum of *eigenvalues* v and *eigenvectors* $|v\rangle$ of V. Finally, compute the time evolution $|\Psi(t)\rangle$ of the initial state vector $|\Psi(0)\rangle$ by solving Schrödinger's equation:

$$i\hbar \frac{d|\Psi\rangle}{dt} = H\Psi \ .$$

The amplitudes $c(v, t)$ are then given by

$$c(v, t) = \langle v|\Psi(t)\rangle \, .$$

II) *Heisenberg picture*: compute the operator $V(F, G)$. Evolve the operator V in time according to Heisenberg equation

$$i\hbar \frac{dV}{dt} = [V, H] \equiv VH - HV \, .$$

Compute the complete spectrum of *eigenvalues* v and *eigenvectors* $|v\rangle_t$ of $V(t)$. The amplitudes $c(v, t)$ are then given by

$$c(v, t) = \langle v_t|\Psi\rangle \, .$$

Acknowledgements. – I would like to thank Profs. G. Contopoulos, for stimulating my interest in the subject of the comparison between classical and quantum perturbation theories, L. Galgani, for making a constructive criticism and correction to my first draft of the present review, and A. Celletti, for a very detailed reading with many helpful suggestions.

REFERENCES

[1] AITCHISON, I., MACMANUS, D.A., and SNYDER, T.M., 2004, *Understanding Heisenberg's 'magical' paper of July 1925: a new look at the calculational details*, American J. Phys. **72**, 1370.
[2] https://www.aip.org/history/heisenberg/p07.htm
[3] ALI, M.K., 1985, *The quantum normal form and its equivalents*, J. Math. Phys. **26**, 2565.
[4] ARNOLD, V.I., 1963, *Small denominators and the problem of stability of motion in classical and celestial mechanics*, Russ. Math. Surveys, **18**, 9.
[5] BERRY, M., 1989, *Quantum chaology, not quantum chaos*, Phys. Scr. **40**, 335.
[6] BIRKHOFF, G., 1927, *Dynamical Systems.* American Mathematical Society. Providence.

[7] BOCCALETTI, D., and PUCACCO, G., 1998 *Theory of Orbits Vol II. Perturbative and Geometric Methods*, Springer, Berlin.
[8] BORN, M., 1935, *Atomic Physics*, Blackie & Son, Glasgow.
[9] BORN, M., and JORDAN, P., 1925, *Zur Quantenmechanik*, Zeitschrift für Physik **34**, 858; Reprinted in translation in B. L. van der Waerden (Ed.), *Sources of Quantum Mechanics*, North-Holland, Amsterdam, 1967.
[10] BORN, M., HEISENBERG, W., and JORDAN, P., 1926, *Zur Quantenmechanik II*, Zeitschrift für Physik **35**, 557; Reprinted in translation in B. L. van der Waerden (Ed.), *Sources of Quantum Mechanics*, North-Holland, Amsterdam, 1967.
[11] BUCHER, M., 2008, *The rise and premature fall of the old quantum theory*, ArXiv. 0802.1366.
[12] CARATI, A., and GALGANI, L., *Fondamenti della meccanica quantistica: uno studio storico critico*. electronic page http:/www.mat.unimi.it/users/galgani
[13] CHRISTODOULIDI, H., and EFTHYMIOPOULOS, C., 2013, *Low-dimensional q-Tori in FPU Lattices: Dynamics and Localization Properties*, Physica D **261**, 92.
[14] CONTOPOULOS, G., 1960, *A third integral of motion in a galaxy*, Z. Astrophys., **49**, 273.
[15] CONTOPOULOS, G., 1966, *Tables of the third integral*, Astrophys. J. Suppl., **13**, 503.
[16] CONTOPOULOS, G., 2002, *Order and Chaos in Dynamical Astronomy*, Springer, Berlin.
[17] CREHAN, P, 1990, *The proper quantum analogue of the Birkhoff-Gustavson method of normal forms*, J. Phys. A Math. Gen. **23**, 5815.
[18] DEPRIT, A., 1969, *Canonical Transformations Depending on a Small Parameter*, Cel. Mech., **1**, 12.
[19] DIRAC, P.A.M., 1926, *On the theory of Quantum Mechanics*, Proc. Roy. Soc. A **109**, 642.
[20] DIRAC, P.A.M., 1930, *The Principles of Quantum Mechanics*, Oxford Science Publications.
[21] DYSON, F.J., 1949, *The Radiation Theories of Tomonaga, Schwinger, and Feynman*, Phys. Rev. **75**, 486.
[22] DYSON, F.J., 1952, *Divergence of Perturbation Theory in Quantum Electrodynamics*, Phys. Rev. **85**, 631.
[23] *Attempts to make quantum electrodynamics into a completely solvable theory* http://www.webofstories.com / play / freeman.dyson / 92; jsessionid= DFD9E857211FDDC7D5727D72EC3DF00A.
[24] ECKHARDT, B., 1986, *Birkhoff-Gustavson normal form in classical and quantum mechanics*, J. Phys. A Math. Gen. **19**, 2961.
[25] EFTHYMIOPOULOS, C., 2012, *Canonical perturbation theory; stability and diffusion in Hamiltonian systems: applications in dynamical astronomy*, in P.M. Cincotta, C.M. Giordano, and C. Efthymiopoulos (Eds), Third La Plata International School on Astronomy and Geophysics, Asociación Argentina de Astronomía, Workshop Series, Vol. 3, p.3-146.
[26] EFTHYMIOPOULOS, C., and CONTOPOULOS, G., 2006, *Chaos in Bohmian Quantum Mechanics*, J. Phys. A. Math Gen **39**, 1819.
[27] ELIASSON, L., 1996, *Absolutely Convergent Series Expansions for Quasi Periodic Motions*, Math. Phys. Electron. J. **2**, 1.
[28] GIORGILLI, A, 1998, *Small denominators and exponential stability: from Poincaré to the present time*, Rend. Sem. Mat. Fis. Milano, LXVIII, 19.
[29] GUSTAVSON, F.G., 1966, *On Constructing Formal Integrals of a Hamiltonian System. Near an Equilibrium Point*, Astron. J. **71**, 670.
[30] GUTZWILLER, M.C., 1990, *Chaos in Classical and Quantum Mechanics*, Springer, New York.

[31] HAMILTON, W.R., 1834, *On a general method of dynamics; by which the study of the motions of all free systems of attracting or repelling points is reduced to the search and differentiation of one central relation, or characteristic function*, Phil. Transactions, Part 2, pp. 247-308 (*Mathematical Papers 2, 103*).
[32] HEISENBERG, W., 1925, *Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen*, Zeitschrift für Physik **33**, 879.
[33] HÉNON, M., and HEILES, C., 1964, *The applicability of the third integral of motion: Some numerical experiments*, Astron. J. **69**, 73.
[34] HORI, G.I., 1966, *Theory of General Perturbation with Unspecified Canonical Variables*, Publ. Astron. Soc. Japan **18**, 287.
[35] JACKSON, J.D., 1962, *Classical Electrodynamics*, John Wiley and Son.
[36] KOLMOGOROV, A.N., 1954, *On the preservation of quasi-periodic motions under a small variation of Hamilton's function*, Dokl. Akad. Nauk SSSR, **98**, 527.
[37] LICHTENBERG, A.J., and LIEBERMAN, M.A., 1992, *Regular and Chaotic Dynamics*, Springer, New York.
[38] LYNDEN-BELL, D., private communication.
[39] MESSIAH, A., 1961, *Quantum Mechanics*, Dover Publications.
[40] MOSER, J, 1962, *On Invariant Curves of Area-Preserving Mappings of an Annulus*, Nachr. Akad. Wiss. Göttingen, Math. Phys. Kl II, 1-20.
[41] NIKOLAEV, A.S., 1996, *On the diagonalization of quantum Birkhoff-Gustavson normal form*, J. Math. Phys. **37**, 2643.
[42] POINCARÉ, H, 1892, *Les Méthodes Nouvelles de la Mécanique Céleste*, Gauthier-Villars, Paris.
[43] ROBNIK, M., 1984, *The algebraic quantisation of the Birkhoff-Gustavson normal form*, J. Phys. A Math. Gen. **17**, 109.
[44] SCHRÖDINGER, E., 1926, *An Undulatory Theory of the Mechanics of Atoms and Molecules*, Phys. Rev. **28**, 1049.
[45] SCHRÖDINGER, E., 1926, *Über das Verhältnis der Heisenberg Born Jordanischen Quantenmechanik zu der meinen*, Annalen der Physik **79**, 45.
[46] VON ZEIPEL, H., 1916, *Recherches sur le mouvement de petites planètes*, Arkiv. Mat. Astron. Physik, II, Nos, 1,3,7,9, 12.
[47] WEINBERG, S., 1992, *Dreams of a Final Theory*, Pantheon Books, New York.
[48] WOOD, W.R., and ALI, M.K., 1987, *On the limitations of the Birkhoff-Gustavson normal form approach*, J. Phys. A Math. Gen. **20**, 351.

Christos Efthymiopoulos
Research Center for Astronomy and Applied Mathematics,
Academy of Athens
E-mail: cefthim@academyofathens.gr