Understanding Heisenberg's "magical" paper of July 1925: A new look at the calculational details

Ian J. R. Aitchison[a)]
Department of Physics, Theoretical Physics, University of Oxford, Oxford OX1 3NP, United Kingdom

David A. MacManus
Tripos Receptor Research Ltd., Bude-Stratton Business Park, Bude, Cornwall EX23 8LY, United Kingdom

Thomas M. Snyder
Department of Mathematics and Engineering Sciences, Lincoln Land Community College, Springfield, Illinois 62794-9256

(Received 20 January 2004; accepted 3 June 2004)

In July 1925 Heisenberg published a paper that ushered in the new era of quantum mechanics. This epoch-making paper is generally regarded as being difficult to follow, partly because Heisenberg provided few clues as to how he arrived at his results. We give details of the calculations of the type that Heisenberg might have performed. As an example we consider one of the anharmonic oscillator problems considered by Heisenberg, and use our reconstruction of his approach to solve it up to second order in perturbation theory. The results are precisely those obtained in standard quantum mechanics, and we suggest that a discussion of the approach, which is based on the direct calculation of transition frequencies and amplitudes, could usefully be included in undergraduate courses on quantum mechanics.
[DOI: 10.1119/1.1775243]

I. INTRODUCTION

Heisenberg's paper of July 1925[1] on "Quantum-mechanical reinterpretation of kinematic and mechanical relations,"[2,3] was the breakthrough that quickly led to the first complete formulation of quantum mechanics.[4–6] Despite its undoubtedly crucial historical role, Heisenberg's approach in this paper is not generally followed in undergraduate quantum mechanics courses, in contrast, for example, to Einstein's approach in the teaching of relativity. Indeed Heisenberg's paper is widely regarded as being difficult to understand and of mainly historical interest today. For example, Weinberg[7] has written that "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 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 this sense, Heisenberg's 1925 paper was pure magic."

There have been many discussions aimed at elucidating the main ideas in Heisenberg's paper of which Refs. 3 and 8–18 represent only a partial selection.[19] Of course, it may not be possible to render completely comprehensible the mysterious processes whereby physicists "jump over all intermediate steps to a new insight about nature."[20] In our opinion, however, one of the main barriers to understanding Heisenberg's paper is a more prosaic one: namely, he gave remarkably few details of the calculations he performed.

In Sec. II we briefly review Heisenberg's reasoning in setting up his new calculational method. Then we present in Sec. III the details of a calculation typical of those we conjecture that he performed. Our reconstruction is based on the assumption that, having formulated a method that was capable of determining the relevant physical quantities (the transition frequencies and amplitudes), Heisenberg then applied it to various simple mechanical systems, without any further recourse to the kind of "inspired guesswork" that characterized the old quantum theory. Surprisingly, this point of view appears to be novel. For example, MacKinnon[10] and Mehra and Rechenberg[11] have suggested that Heisenberg arrived at the crucial recursion relations [see Eqs. (33)–(36) in Sec. III B] by essentially guessing the appropriate generalization of their classical counterparts. We are unaware of any evidence that can settle the issue. In any case, our analysis shows that it is possible to read Heisenberg's paper as providing a complete (if limited) calculational method, the results of which are consistent with those of standard quantum mechanics. We also stress both the correctness and the practicality of what we conjecture to be Heisenberg's calculational method. We hope that our reappraisal will stimulate instructors to include at least some discussion of it in their undergraduate courses.

II. HEISENBERG'S TRANSITION AMPLITUDE APPROACH

A. Quantum kinematics

Heisenberg began his paper with a programmatic call[21,22] to "discard all hope of observing hitherto unobservable quantities, such as the position and period of the electron," and instead to "try to establish a theoretical quantum mechanics, analogous to classical mechanics, but in which only relations between observable quantities occur." As an example of such latter quantities, he immediately pointed to the energies $W(n)$ of the Bohr stationary states, together with the associated Einstein–Bohr frequencies[23]

$$\omega(n,n-\alpha)=\frac{1}{\hbar}[W(n)-W(n-\alpha)], \tag{1}$$

and noted that these frequencies, which characterize the radiation emitted in the transition $n \to n-\alpha$, depend on two variables. An example of a quantity he wished to exclude from the new theory is the time-dependent position coordinate $x(t)$. In considering what might replace it, he turned to the probabilities for transitions between stationary states.

Consider a simple one-dimensional model of an atom consisting of an electron undergoing periodic motion, which is the type of system studied by Heisenberg. For a state characterized by the label n, the fundamental frequency $\omega(n)$, and the coordinate $x(n,t)$, we can represent $x(n,t)$ as a Fourier series

$$x(n,t)=\sum_{\alpha=-\infty}^{\infty} X_{\alpha}(n)e^{i\alpha\omega(n)t}, \tag{2}$$

where α is an integer.[24] According to classical theory, the energy emitted per unit time (the power) in a transition corresponding to the αth harmonic $\alpha\omega(n)$ is[25]

$$-\left(\frac{dE}{dt}\right)_{\alpha}=\frac{e^2}{3\pi\epsilon_0 c^3}[\alpha\omega(n)]^4|X_{\alpha}(n)|^2. \tag{3}$$

In the quantum theory, however, the transition frequency corresponding to the classical $\alpha\omega(n)$ is, in general, not a simple multiple of a fundamental frequency, but is given by Eq. (1), so that $\alpha\omega(n)$ is replaced by $\omega(n,n-\alpha)$. Correspondingly, Heisenberg introduced the quantum analogue of $X_{\alpha}(n)$, written (in our notation) as $X(n,n-\alpha)$.[27] Furthermore, the left-hand side of Eq. (3) has to be replaced by the product of the transition probability per unit time, $P(n,n-\alpha)$, and the emitted energy $\hbar\omega(n,n-\alpha)$. Thus Eq. (3) becomes

$$P(n,n-\alpha)=\frac{e^2}{3\pi\epsilon_0\hbar c^3}[\omega(n,n-\alpha)]^3|X(n,n-\alpha)|^2. \tag{4}$$

It is the transition amplitudes $X(n,n-\alpha)$ which Heisenberg took to be "observable;" like the transition frequencies, they depend on two discrete variables.[28]

Equation (4) refers, however, to only one specific transition. For a full description of atomic dynamics (as then conceived), we need to consider all the quantities $X(n,n-\alpha)\exp[i\omega(n,n-\alpha)t]$. In the classical case, the terms $X_{\alpha}(n)\exp[i\alpha\omega(n)t]$ may be combined to yield $x(t)$ via Eq. (2). But in the quantum theory, Heisenberg wrote[29] that a "similar combination of the corresponding quantum-theoretical quantities seems to be impossible in a unique manner and therefore not meaningful, in view of the equal weight of the variables n and $n-\alpha$ [that is, in the amplitude $X(n,n-\alpha)$ and frequency $\omega(n,n-\alpha)$] ... However, one may readily regard the ensemble of quantities $X(n,n-\alpha)\exp[i\omega(n,n-\alpha)t]$ as a representation of the quantity $x(t)$... ." This way of representing $x(t)$, that is, as we would now say, by a matrix, is the first of Heisenberg's "magical jumps," and surely a very large one. Representing $x(t)$ in this way seems to be the sense in which Heisenberg considered that he was offering a "reinterpretation of kinematic relations."

Heisenberg immediately posed the question: how is the quantity $x(t)^2$ to be represented? In classical theory, the answer is straightforward. From Eq. (2) we obtain

$$[x(t)]^2=\sum_{\alpha}\sum_{\alpha'} X_{\alpha}(n)X_{\alpha'}(n)e^{i(\alpha+\alpha')\omega(n)t}. \tag{5}$$

We set $\beta=\alpha+\alpha'$, and rewrite Eq. (5) as

$$[x(t)]^2=\sum_{\beta} Y_{\beta}(n)e^{i\beta\omega(n)t}, \tag{6}$$

where

$$Y_{\beta}(n)=\sum_{\alpha} X_{\alpha}(n)X_{\beta-\alpha}(n). \tag{7}$$

Thus $[x(t)]^2$ is represented classically (via a Fourier series) by the set of quantities $Y_{\beta}(n)\exp[i\beta\omega(n)t]$, the frequency $\beta\omega(n)$ being the simple combination $[\alpha\omega(n)+(\beta-\alpha)\omega(n)]$. In quantum theory, the corresponding representative quantities must be written as $Y(n,n-\beta)\exp[i\omega(n,n-\beta)t]$, and the question is what is the analogue of Eq. (7)?

The crucial difference in the quantum case is that the frequencies do not combine in the same way as the classical harmonics, but rather in accordance with the Ritz combination principle:

$$\omega(n,n-\alpha)+\omega(n-\alpha,n-\beta)=\omega(n,n-\beta), \tag{8}$$

which is consistent with Eq. (1). Thus in order to end up with the particular frequency $\omega(n,n-\beta)$, it seems "almost necessary" (in Heisenberg's words[30]) to combine the quantum amplitudes in such a way as to ensure the frequency combination Eq. (8), that is, as

$$Y(n,n-\beta)e^{i\omega(n,n-\beta)t}=\sum_{\alpha} X(n,n-\alpha)e^{i\omega(n,n-\alpha)t} \times X(n-\alpha,n-\beta)e^{i\omega(n-\alpha,n-\beta)t}, \tag{9}$$

or

$$Y(n,n-\beta)=\sum_{\alpha} X(n,n-\alpha)X(n-\alpha,n-\beta), \tag{10}$$

which is Heisenberg's rule for multiplying transition amplitudes. Note particularly that the replacements $X_{\alpha}(n)\to X(n,n-\alpha)$, and similarly for $Y_{\beta}(n)$ and $X_{\beta-\alpha}(n)$ in Eq. (7), produce a quite different result.

Heisenberg indicated the simple extension of the rule given in Eq. (10) to higher powers $[x(t)]^n$, but noticed at once[31] that a "significant difficulty arises, however, if we consider two quantities $x(t),y(t)$ and ask after their product $x(t)y(t)$... Whereas in classical theory $x(t)y(t)$ is always equal to $y(t)x(t)$, this is not necessarily the case in quantum theory." Heisenberg used the word "difficulty" three times in referring to this unexpected consequence of his multiplication rule, but it very quickly became clear that the non-commutativity (in general) of kinematical quantities in quantum theory was the essential new idea in the paper.

Born recognized Eq. (10) as matrix multiplication (something unknown to Heisenberg in July 1925), and he and Jordan rapidly produced the first paper[4] to state the fundamental commutation relation (in modern notation)

$$\hat{x}\hat{p}-\hat{p}\hat{x}=i\hbar. \tag{11}$$

Dirac's paper followed soon after,[5] and then the paper of Born, Heisenberg, and Jordan.[6]

The economy and force of Heisenberg's argument in reaching Eq. (10) is remarkable, and it is at least worth considering whether presenting it to undergraduates might help them to understand the ''almost necessity'' of noncommuting quantities in quantum theory.

B. Quantum dynamics

Having identified the transition amplitudes $X(n,n-\alpha)$ and frequencies $\omega(n,n-\alpha)$ as the observables of interest in the new theory, Heisenberg then turned his attention to how they could be determined from the dynamics of the system. In the old quantum theory, this determination would have been done in two stages: by integration of the equation of motion

$$\ddot{x}+f(x)=0, \tag{12}$$

and by determining the constants of the periodic motion through the ''quantum condition''

$$\oint p\,dq=\oint m\dot{x}^2dt=J(=nh), \tag{13}$$

where the integral is evaluated over one period. In regard to Eq. (12), Heisenberg wrote[32] that it is ''very natural'' to take the classical equation of motion over to quantum theory by replacing the classical quantities $x(t)$ and $f(x)$ by their kinematical reinterpretations,[33] as in Sec. II A (or, as we would say today, by taking matrix elements of the corresponding operator equation of motion). He noted that in the classical case a solution can be obtained by expressing $x(t)$ as a Fourier series, substitution of which into the equation of motion leads (in special cases) to a set of recursion relations for the Fourier coefficients. In the quantum theory, Heisenberg wrote that[32] ''we are at present forced to adopt this method of solving equation Eq. (12) [his Eq. (H11)] ... since it was not possible to define a quantum-theoretical function analogous to the [classical] function $x(n,t)$.'' In Sec. III we shall consider the simple example (the first of those chosen by Heisenberg) $f(x)=\omega_0^2x+\lambda x^2$, and obtain the appropriate recursion relations in the classical and the quantum cases.

A quantum-theoretical reinterpretation of Eq. (13) is similarly required in terms of the transition amplitudes $X(n,n-\alpha)$. In the classical case, the substitution of Eq. (2) into Eq. (13) gives

$$\oint m\dot{x}^2dt=2\pi m\sum_{\alpha=-\infty}^{\infty}|X_\alpha(n)|^2\alpha^2\omega(n)=nh, \tag{14}$$

using $X_\alpha(n)=[X_{-\alpha}(n)]^*$. Heisenberg argued that Eq. (14) appeared arbitrary in the sense of the correspondence principle, because the latter determined J only up to an additive constant (times h). He therefore replaced Eq. (14) by the derivative form [Eq. (H15)]

$$h=2\pi m\sum_{\alpha=-\infty}^{\infty}\alpha\frac{d}{dn}(\alpha|X_\alpha(n)|^2\omega(n)). \tag{15}$$

The summation can alternatively be written as over positive values of α, replacing $2\pi m$ by $4\pi m$. In another crucial jump, Heisenberg then replaced the differential in Eq. (15) by a difference, giving

$$h=4\pi m\sum_{\alpha=0}^{\infty}[|X(n+\alpha,n)|^2\omega(n+\alpha,n)-|X(n,n-\alpha)|^2\omega(n,n-\alpha)], \tag{16}$$

which is Eq. (H16) in our notation.[34] As he later recalled, he had noticed that ''if I wrote down this [presumably Eq. (15)] and tried to translate it according to the scheme of dispersion theory, I got the Thomas-Kuhn sum rule [Eq. (16)[35,36]]. And that is the point. Then I thought, That is apparently how it is done.''[37]

By ''the scheme of dispersion theory,'' Heisenberg referred to what Jammer[38] calls Born's correspondence rule, namely[39]

$$\alpha\frac{\partial\Phi(n)}{\partial n}\leftrightarrow\Phi(n)-\Phi(n-\alpha), \tag{17}$$

or rather to its iteration to the form[40]

$$\alpha\frac{\partial\Phi(n,\alpha)}{\partial n}\leftrightarrow\Phi(n+\alpha,n)-\Phi(n,n-\alpha), \tag{18}$$

as used in the Kramers–Heisenberg theory of dispersion.[41,42] It took Born only a few days to show that Heisenberg's quantum condition, Eq. (16), was the diagonal matrix element of Eq. (11), and to guess[43] that the off-diagonal elements of $\hat{x}\hat{p}-\hat{p}\hat{x}$ were zero, a result that was shown to be compatible with the equations of motion by Born and Jordan.[4]

At this point it is appropriate to emphasize that Heisenberg's transition amplitude $X(n,n-\alpha)$ is the same as the quantum-mechanical matrix element $\langle n-\alpha|\hat{x}|n\rangle$, where $|n\rangle$ is the eigenstate with energy $W(n)$. The relation of Eq. (16) to the fundamental commutator Eq. (11) is discussed briefly in Appendix A.

Heisenberg noted[44] that the undetermined constant still contained in the quantities X of Eq. (16) [assuming the frequencies known from Eq. (12)] would be determined by the condition that a ground state should exist, from which no radiation is emitted [see Eqs. (51) and (52) below]. He therefore summarized the state of affairs thus far by the statement[44] that Eqs. (12) and (16) ''if soluble, contain a complete determination not only of frequencies and energy values, but also of quantum-theoretical transition probabilities.'' We draw attention to the strong claim here: that he has arrived at a new calculational method, which will completely determine the observable quantities. Let us now see in detail how this method works, for a harmonic oscillator perturbed by an anharmonic force of the form λx^2 per unit mass.[45]

III. HEISENBERG'S CALCULATIONAL METHOD AND ITS APPLICATION TO THE ANHARMONIC OSCILLATOR

A. Recursion relations in the quantum case

The classical equation of motion is

$$\ddot{x}+\omega_0^2x+\lambda x^2=0. \tag{19}$$

We depart from the order of Heisenberg's presentation and begin by showing how—as he stated—Eq. (19) leads to recursion relations for the transition amplitudes $X(n,n-\alpha)$. The $(n,n-\alpha)$ representative[46] of the first two terms in Eq. (19) is straightforward, being

$$[-\omega^2(n,n-\alpha)+\omega_0^2]X(n,n-\alpha)e^{i\omega(n,n-\alpha)t}, \tag{20}$$

while that of the third term is, by Eq. (10),

$$\lambda \sum_{\beta} X(n,n-\beta)X(n-\beta,n-\alpha)e^{i\omega(n,n-\alpha)t}. \tag{21}$$

The $(n,n-\alpha)$ representative of Eq. (19) therefore yields[47]

$$[\omega_0^2-\omega^2(n,n-\alpha)]X(n,n-\alpha)+\lambda \sum_{\beta} X(n,n-\beta) \times X(n-\beta,n-\alpha)=0, \tag{22}$$

which generates a recursion relation for each value of α ($\alpha=0,\pm 1,\pm 2,\ldots$). For example, for $\alpha=0$ we obtain

$$\omega_0^2 X(n,n)+\lambda[X(n,n)X(n,n)+X(n,n-1)X(n-1,n) + X(n,n+1)X(n+1,n)+\cdots]=0. \tag{23}$$

No general solution for this infinite set of nonlinear algebraic equations seems to be possible, so, following Heisenberg, we turn to a perturbative approach.

B. Perturbation theory

To make the presentation self-contained, we need to discuss several ancillary results. Heisenberg began by considering the perturbative solution of the classical equation (12). He wrote the solution in the form

$$x(t)=\lambda a_0+a_1\cos\omega t+\lambda a_2\cos 2\omega t+\lambda^2 a_3\cos 3\omega t +\cdots+\lambda^{\alpha-1}a_\alpha \cos\alpha\omega t+\cdots, \tag{24}$$

where the coefficients a_α, and ω, are to be expanded as a power series in λ, the first terms of which are independent of λ:[48]

$$a_0=a_0^{(0)}+\lambda a_0^{(1)}+\lambda^2 a_0^{(2)}+\cdots, \tag{25a}$$

$$a_1=a_1^{(0)}+\lambda a_1^{(1)}+\lambda^2 a_1^{(2)}+\cdots, \tag{25b}$$

and

$$\omega=\omega_0+\lambda\omega^{(1)}+\lambda^2\omega^{(2)}+\cdots. \tag{26}$$

We substitute Eq. (24) into Eq. (12), use standard trigonometric identities, and equate to zero the terms that are constant and which multiply $\cos\omega t$, $\cos 2\omega t$, etc., to obtain

$$\lambda\{\omega_0^2 a_0+\tfrac{1}{2}a_1^2+[\lambda^2(a_0^2+\tfrac{1}{2}a_2^2)+\cdots]\}=0, \tag{27a}$$

$$(-\omega^2+\omega_0^2)a_1+[\lambda^2(a_1a_2+2a_0a_1)+\cdots]=0, \tag{27b}$$

$$\lambda\{(-4\omega^2+\omega_0^2)a_2+\tfrac{1}{2}a_1^2+[\lambda^2(a_1a_3+2a_0a_2)+\cdots]\} =0, \tag{27c}$$

$$\lambda^2\{(-9\omega^2+\omega_0^2)a_3+a_1a_2+[\lambda^2(a_1a_4+2a_0a_3)+\cdots]\} =0, \tag{27d}$$

where the dots stand for higher powers of λ. If we drop the terms of order λ^2 (and higher powers), and cancel overall factors of λ, Eq. (27) becomes (for $\lambda\neq 0$ and $a_1\neq 0$)

$$\omega_0^2 a_0+\tfrac{1}{2}a_1^2=0, \tag{28a}$$

$$(-\omega^2+\omega_0^2)=0, \tag{28b}$$

$$(-4\omega^2+\omega_0^2)a_2+\tfrac{1}{2}a_1^2=0, \tag{28c}$$

$$(-9\omega^2+\omega_0^2)a_3+a_1a_2=0, \tag{28d}$$

which is the same as Eq. (H18).[49] The lowest order in λ solution is obtained from Eq. (28) by setting $\omega=\omega_0$, and replacing each a_α by the corresponding one with a superscript $^{(0)}$ [see Eq. (25)].

In the quantum case, Heisenberg proposed to seek a solution analogous to Eq. (24). Of course, it is now a matter of using the representation of $x(t)$ in terms of the quantities $X(n,n-\alpha)\exp[i\omega(n,n-\alpha)t]$. But it seems reasonable to assume that, as the index α increases from zero in integer steps, each successive amplitude will (to leading order in λ) be suppressed by an additional power of λ, as in the classical case. Thus Heisenberg suggested that, in the quantum case, $x(t)$ should be represented by terms of the form

$$\lambda a(n,n),\ \ a(n,n-1)\cos\omega(n,n-1)t, \lambda a(n,n-2)\cos\omega(n,n-2)t,\ldots, \lambda^{\alpha-1}a(n,n-\alpha)\cos\omega(n,n-\alpha)t,\ldots, \tag{29}$$

where, as in Eqs. (25) and (26),

$$a(n,n)=a^{(0)}(n,n)+\lambda a^{(1)}(n,n)+\lambda^2 a^{(2)}(n,n)+\cdots, \tag{30}$$

$$a(n,n-1)=a^{(0)}(n,n-1)+\lambda a^{(1)}(n,n-1) +\lambda^2 a^{(2)}(n,n-1)+\cdots, \tag{31}$$

and

$$\omega(n,n-\alpha)=\omega^{(0)}(n,n-\alpha)+\lambda\omega^{(1)}(n,n-\alpha) +\lambda^2\omega^{(2)}(n,n-\alpha)+\cdots. \tag{32}$$

As Born and Jordan pointed out,[4] some use of correspondence arguments has been made here in assuming that as $\lambda\to 0$, only transitions between adjacent states are possible. We shall return to this point in Sec. III C.

Heisenberg then simply wrote down what he asserted to be the quantum version of Eq. (28), namely[50]

$$\omega_0^2 a(n,n)+\tfrac{1}{4}[a^2(n+1,n)+a^2(n,n-1)]=0 \tag{33}$$

$$-\omega^2(n,n-1)+\omega_0^2=0, \tag{34}$$

$$[-\omega^2(n,n-2)+\omega_0^2]a(n,n-2)+\tfrac{1}{2}[a(n,n-1) \times a(n-1,n-2)]=0, \tag{35}$$

$$[-\omega^2(n,n-3)+\omega_0^2]a(n,n-3)+\tfrac{1}{2}a(n,n-1) \times a(n-1,n-3)+\tfrac{1}{2}a(n,n-2)a(n-2,n-3)=0. \tag{36}$$

The question we now address is how did Heisenberg arrive at Eqs. (33)–(36)?

We shall show that these equations can be straightforwardly derived from Eq. (22) using the ansatz (29), and we suggest that this is what Heisenberg did. This seems to be a novel proposal. Tomonaga[8] derived Eq. (22) but then discussed only the $\lambda\to 0$ limit, that is, the simple harmonic oscillator, a special case to which we shall return in Sec. III C. The only other authors, to our knowledge, who have discussed the presumed details of Heisenberg's calculations are[51] Mehra and Rechenberg.[11] They suggest that Heisenberg guessed how to "translate," "reinterpret," or "reformulate" (their words) the classical equation (28) into the quantum ones, Eqs. (33)–(36), in a way that was consistent with his multiplication rule, Eq. (10). Although such "inspired guesswork" was undoubtedly necessary in the stages leading up to Heisenberg's paper,[1] it seems more plausible to us that by the time of the paper's final formulation, Heisenberg realized that he had a calculational method in which guesswork was no longer necessary, and in which Eqs. (33)–(36), in particular, could be derived.

Unfortunately, we know of no documentary evidence that directly proves (or disproves) this suggestion, but we think there is some internal evidence for it. In the passage to which attention was drawn earlier,[44] Heisenberg asserted that his formalism constituted a complete method for calculating everything that needs to be calculated. It is difficult to believe that Heisenberg did not realize that his method led directly to Eqs. (33)–(36), without the need for any "translations" of the classical relations.

To apply the ansatz of Eq. (29) to Eq. (22), we need to relate the amplitudes $X(n,n-\alpha)$ to the corresponding quantities $\lambda^{\alpha-1}a(n,n-\alpha)$. We first note that in the classical case,

$$X_\alpha(n)=X^*_{-\alpha}(n), \tag{37}$$

because $x(t)$ in Eq. (2) has to be real. Consider, without loss of generality, the case $\alpha>0$. Then the quantum-theoretical analogue of the left-hand side of Eq. (37) is $X(n,n-\alpha)$, and that of the right-hand side is $X^*(n-\alpha,n)$ (see Ref. 27). Hence the quantum-theoretical analogue of Eq. (37) is

$$X(n,n-\alpha)=X^*(n-\alpha,n), \tag{38}$$

which is nothing but the relation $\langle n-\alpha|\hat{x}|n\rangle=\langle n|\hat{x}|n-\alpha\rangle^*$ for the Hermitian observable $\hat{x}$. Although $X(n,n-\alpha)$ can in principle be complex (and Heisenberg twice discussed the significance of the phases of such amplitudes), Heisenberg seems to have assumed (as is certainly plausible) that in the context of the classical cosine expansion in Eq. (24) and the corresponding quantum terms in Eq. (29), the $X(n,n-\alpha)$'s should be chosen to be real, so that Eq. (38) becomes

$$X(n,n-\alpha)=X(n-\alpha,n), \tag{39}$$

that is, the matrix with elements $\{X(n,n-\alpha)\}$ is symmetric. Consider a typical term of Eq. (29),

$$\begin{aligned}\lambda^{\alpha-1}&a(n,n-\alpha)\cos[\omega(n,n-\alpha)t]\\&=\frac{\lambda^{\alpha-1}}{2}a(n,n-\alpha)[e^{i\omega(n,n-\alpha)t}+e^{-i\omega(n,n-\alpha)t}]\\&=\frac{\lambda^{\alpha-1}}{2}a(n,n-\alpha)[e^{i\omega(n,n-\alpha)t}+e^{i\omega(n-\alpha,n)t}],\end{aligned} \tag{40}$$

using $\omega(n,n-\alpha)=-\omega(n-\alpha,n)$ from Eq. (1). If we assume that $a(n,n-\alpha)=a(n-\alpha,n)$ as discussed for Eq. (39), we see that it is consistent to write

$$X(n,n-\alpha)=\frac{\lambda^{\alpha-1}}{2}a(n,n-\alpha)\quad(\alpha>0) \tag{41}$$

and in general

$$X(n,n-\alpha)=\frac{\lambda^{|\alpha|-1}}{2}a(n,n-\alpha)\quad(\alpha\neq 0). \tag{42}$$

The case $\alpha=0$ is clearly special, with $X(n,n)=\lambda a(n,n)$.

We may now write out the recurrence relations Eq. (22) explicitly for $\alpha=0,1,2,\ldots$, in terms of $a(n,n-\alpha)$ rather than $X(n,n-\alpha)$. We shall include terms up to and including terms of order λ^2. For $\alpha=0$ we obtain

$$\begin{aligned}\lambda\{\omega_0^2a(n,n)&+\tfrac{1}{4}[a^2(n+1,n)+a^2(n,n-1)]+\lambda^2[a^2(n,n)\\&+\tfrac{1}{4}(a^2(n+2,n)+a^2(n,n-2))]\}=0.\end{aligned} \tag{43}$$

We note the connection with Eq. (27a), and that Eq. (43) reduces to Eq. (33) when the λ^2 term is dropped and an overall factor of λ is canceled. Similarly, for $\alpha=1$ we obtain

$$\begin{aligned}(-\omega^2(n,n-1)&+\omega_0^2)a(n,n-1)+\lambda^2\{a(n,n)a(n,n-1)\\&+a(n,n-1)a(n-1,n-1)+\tfrac{1}{2}[a(n,n+1)\\&\times a(n+1,n-1)+a(n,n-2)a(n-2,n-1)]\}=0\end{aligned} \tag{44}$$

[see Eq. (27b)]. For $\alpha=2$ we have

$$\begin{aligned}\lambda\{(-\omega^2(n,n-2)&+\omega_0^2)a(n,n-2)+\tfrac{1}{2}a(n,n-1)\\&\times a(n-1,n-2)+\lambda^2[a(n,n)a(n,n-2)+a(n,n-2)\\&\times a(n-2,n-2)+\tfrac{1}{2}a(n,n+1)a(n+1,n-2)\\&+\tfrac{1}{2}a(n,n-3)a(n-3,n-2)]\}=0\end{aligned} \tag{45}$$

[see Eq. (27c)]. For $\alpha=3$ [see Eq. (27d)] we obtain

$$\begin{aligned}\lambda^2\{(-\omega^2(n,n-3)&+\omega_0^2)a(n,n-3)+\tfrac{1}{2}[a(n,n-1)\\&\times a(n-1,n-3)+a(n,n-2)a(n-2,n-3)]\\&+\lambda^2[a(n,n)a(n,n-3)+a(n,n-3)a(n-3,n-3)\\&+\tfrac{1}{2}a(n,n+1)a(n+1,n-3)+\tfrac{1}{2}a(n,n-4)\\&\times a(n-4,n-3)]\}=0.\end{aligned} \tag{46}$$

If we drop the terms multiplied by λ^2, Eqs. (43)–(46) reduce to Eqs. (33)–(36). This appears to be the first published derivation of the latter equations.

In addition to these recurrence relations which follow from the equations of motion, we also need the perturbative version of the quantum condition Eq. (16).[52] We include terms of order λ^2, consistent with Eqs. (43)–(46), so that Eq. (16) becomes

$$\begin{aligned}\frac{h}{\pi m}&=a^2(n+1,n)\omega(n+1,n)-a^2(n,n-1)\omega(n,n-1)\\&\quad+\lambda^2[a^2(n+2,n)\omega(n+2,n)-a^2(n,n-2)\\&\quad\times\omega(n,n-2)].\end{aligned} \tag{47}$$

We are now ready to obtain the solutions.

C. The lowest-order solutions for the amplitudes and frequencies

We begin by considering the lowest-order solutions in which all λ^2 terms are dropped from Eqs. (43) to (47), and all quantities (a's and ω's) are replaced by the corresponding ones with a superscript $^{(0)}$ [compare Eqs. (30)–(32)].[53] In this case, Eq. (44) reduces to

$$[-(\omega^{(0)}(n,n-1))^2+\omega_0^2]a^{(0)}(n,n-1)=0, \tag{48}$$

so that assuming $a^{(0)}(n,n-1)\neq 0$, we obtain

$$\omega^{(0)}(n,n-1)=\omega_0 \tag{49}$$

for all n. If we substitute Eq. (49) into the lowest-order version of Eq. (47), we find

$$\frac{h}{\pi m\omega_0}=[a^{(0)}(n+1,n)]^2-[a^{(0)}(n,n-1)]^2. \tag{50}$$

The solution of this difference equation is

$$[a^{(0)}(n,n-1)]^2=\frac{h}{\pi m\omega_0}(n+\text{constant}), \tag{51}$$

as given in Eq. (H20).[53] To determine the value of the constant, Heisenberg used the idea that in the ground state there can be no transition to a lower state. Thus

$$[a^{(0)}(0,-1)]^2=0, \tag{52}$$

and the constant in Eq. (51) is determined to be zero. Equation (51) then gives (up to a convention as to sign)

$$a^{(0)}(n,n-1)=\beta\sqrt{n}, \tag{53}$$

where

$$\beta=(h/\pi m\omega_0)^{1/2}. \tag{54}$$

Equations (49) and (53) were Heisenberg's first results, and they pertain to the simple (unperturbed) oscillator. We can check Eq. (53) against the usual quantum mechanical calculation via

$$a^{(0)}(n,n-1)=2X^{(0)}(n,n-1)=2\,{}_0\langle n-1|\hat{x}|n\rangle_0, \tag{55}$$

where the states $|n\rangle_0$ are unperturbed oscillator eigenstates. It is well known that[54]

$${}_0\langle n-1|\hat{x}|n\rangle_0=\left(\frac{\hbar}{2m\omega_0}\right)^{1/2}\sqrt{n}, \tag{56}$$

which agrees with Eq. (53), using Eq. (54). A similar treatment of Eq. (43) leads to

$$a^{(0)}(n,n)=-\frac{\beta^2}{4\omega_0^2}(2n+1). \tag{57}$$

Turning next to Eq. (45), the lowest-order form is

$$(-[\omega^{(0)}(n,n-2)]^2+\omega_0^2)a^{(0)}(n,n-2)$$
$$+\tfrac{1}{2}a^{(0)}(n,n-1)a^{(0)}(n-1,n-2)=0. \tag{58}$$

Because the combination law Eq. (8) must be true for the lowest-order frequencies, we have

$$\omega^{(0)}(n,n-2)=\omega^{(0)}(n,n-1)+\omega^{(0)}(n-1,n-2)=2\omega_0, \tag{59}$$

where we have used Eq. (49), and in general

$$\omega^{(0)}(n,n-\alpha)=\alpha\omega_0 \quad (\alpha=1,2,3,\ldots). \tag{60}$$

If we use Eqs. (53), (59), and (60), we obtain

$$a^{(0)}(n,n-2)=\frac{\beta^2}{6\omega_0^2}\sqrt{n(n-1)}. \tag{61}$$

A similar treatment of Eq. (46) yields

$$a^{(0)}(n,n-3)=\frac{\beta^3}{48\omega_0^4}\sqrt{n(n-1)(n-2)}. \tag{62}$$

Consideration of the lowest-order term in Eq. (22) leads to

$$a^{(0)}(n,n-\alpha)=A_\alpha\frac{\beta^\alpha}{\omega_0^{2(\alpha-1)}}\sqrt{\frac{n!}{(n-\alpha)!}}, \tag{63}$$

where A_α is a numerical factor depending on α; Eq. (63) is equivalent to Eq. (H21).

It is instructive to comment on the relation of the above results to those that would be obtained in standard quantum-mechanical perturbation theory. At first sight, it is surprising to see nonzero amplitudes for two-quantum [Eq. (61)], three-quantum [Eq. (62)], or α-quantum [Eq. (63)] transitions appearing at lowest order. But we have to remember that in Heisenberg's perturbative ansatz, Eq. (29), the α-quantum amplitude appears multiplied by a factor $\lambda^{\alpha-1}$. Thus, for example, the lowest order two-quantum amplitude is really $\lambda a^{(0)}(n,n-2)$, not just $a^{(0)}(n,n-2)$. Indeed, such a transition is to be expected precisely at order λ^1 in conventional perturbation theory. The amplitude is $\langle n-2|\hat{x}|n\rangle$ where, to order λ,

$$|n\rangle=|n\rangle_0+\frac{1}{3}m\lambda\sum_{k\neq n}\frac{{}_0\langle k|\hat{x}^3|n\rangle_0}{(n-k)\hbar\omega_0}|k\rangle_0. \tag{64}$$

The operator $\hat{x}^3$ connects $|n\rangle_0$ to $|n+3\rangle_0, |n+1\rangle_0, |n-1\rangle_0$, and $|n-3\rangle_0$, and similar connections occur for ${}_0\langle n-2|$, so that a nonzero $O(\lambda)$ amplitude is generated in $\langle n-2|\hat{x}|n\rangle$.

It is straightforward to check that Eq. (61) is indeed correct quantum-mechanically, but it is more tedious to check Eq. (62), and distinctly unpromising to contemplate checking Eq. (63) by doing a conventional perturbation calculation to order $\alpha-1$. For this particular problem, the improved perturbation theory represented by Eq. (29) is clearly very useful.

After having calculated the amplitudes for this problem to lowest order, Heisenberg next considered the energy. Unfortunately he again gave no details of his calculation, beyond saying that he used the classical expression for the energy, namely

$$W=\tfrac{1}{2}m\dot{x}^2+\tfrac{1}{2}m\omega_0^2x^2+\tfrac{1}{3}m\lambda x^3. \tag{65}$$

It seems a reasonable conjecture, however, that he replaced each term in Eq. (65) by its corresponding matrix, as discussed in Sec. II A. Thus x^2, for example, is represented by a matrix whose $(n,n-\alpha)$ element is

$$\sum_\beta X(n,n-\beta)X(n-\beta,n-\alpha)e^{i\omega(n,n-\alpha)t}, \tag{66}$$

according to his multiplication rule, Eq. (10). A similar replacement is made for x^3, and $\dot{x}^2$ is replaced by

$$\sum_\beta i\omega(n,n-\beta)X(n,n-\beta)e^{i\omega(n,n-\beta)t}$$
$$\times i\omega(n-\beta,n-\alpha)X(n-\beta,n-\alpha)e^{i\omega(n-\beta,n-\alpha)t}$$
$$=\sum_\beta \omega(n,n-\beta)\omega(n-\alpha,n-\beta)X(n,n-\beta)$$
$$\times X(n-\beta,n-\alpha)e^{i\omega(n,n-\alpha)t}, \tag{67}$$

using $\omega(n,m)=-\omega(m,n)$. The total energy is represented by the matrix with elements

$$W(n,n-\alpha)e^{i\omega(n,n-\alpha)t}. \tag{68}$$

It follows that if energy is to be conserved (that is, time-independent) the off-diagonal elements must vanish:

$$W(n,n-\alpha)=0. \quad (\alpha\neq 0). \tag{69}$$

The term $\alpha=0$ is time-independent, and may be taken to be the energy in the state n. The crucial importance of checking the condition Eq. (69) was clearly appreciated by Heisenberg.

To lowest order in λ, the last term in Eq. (65) may be dropped. Furthermore, referring to Eq. (29), the only λ-independent terms in the X-amplitudes are those involving one-quantum jumps such as $n\rightarrow n-1$, corresponding in lowest order to amplitudes such as $X^{(0)}(n,n-1)=\frac{1}{2}a^{(0)}(n,n-1)$. It then follows from Eqs. (66) and (67) that the elements $W(n,n)$, $W(n,n-2)$ and $W(n,n+2)$, and only these elements, are independent of λ when evaluated to lowest order. In Appendix B we show that $W(n,n-2)$ vanishes to lowest order, and $W(n,n+2)$ vanishes similarly. Thus, to lowest order in λ, the energy is indeed conserved (as Heisenberg noted), and is given [using Eq. (66) and Eq. (67) with $\alpha=0$ and $\beta=\pm 1$] by

$$\begin{aligned} W(n,n)=&\ \tfrac{1}{2}m[\omega^{(0)}(n,n-1)]^2[X^{(0)}(n,n-1)]^2\\ &+\tfrac{1}{2}m[\omega^{(0)}(n+1,n)]^2[X^{(0)}(n+1,n)]^2\\ &+\tfrac{1}{2}m\omega_0^2[X^{(0)}(n,n-1)]^2+\tfrac{1}{2}m\omega_0^2\\ &\times[X^{(0)}(n+1,n)]^2\\ =&\ (n+\tfrac{1}{2})\hbar\omega_0, \end{aligned} \tag{70}$$

where we have used Eqs. (49), (53), and (54). Equation (70) is the result given by Heisenberg in Eq. (H23).

These lowest order results are the only ones Heisenberg reported for the λx^2 term. We do not know whether he carried out higher-order calculations for this case or not. What he wrote next[55] is that the "more precise calculation, taking into account higher order approximations in W, a, ω will now be carried out for the simpler example of an anharmonic oscillator $\ddot{x}+\omega_0^2+\lambda x^3=0$." This case is slightly simpler because in the expression corresponding to the ansatz (29) only the odd terms are present, that is, $a_1,\lambda a_3,\lambda^2 a_5$, etc.

The results Heisenberg stated for the λx^3 problem include terms up to order λ in the amplitudes, and terms up to order λ^2 in the frequency $\omega(n,n-1)$ and in the energy W. Once again, he gave no details of how he did the calculations. We believe there can be little doubt that he went through the algebra of solving the appropriate recurrence relations up to order λ^2 in the requisite quantities. As far as we know, the details of such a calculation have not been given before, and we believe that it is worth giving them here, as they are of both pedagogical and historical interest. In the following section we shall obtain the solutions for the λx^2 term (up to order λ^2) which we have been considering, rather than start afresh with the λx^3 term. The procedure is the same for both.

Before leaving the lowest order calculations, we address a question that may have occurred to the reader. Given that, at this stage in his paper, the main results actually relate to the simple harmonic oscillator rather than to the anharmonic one, why did Heisenberg not begin his discussion of toy models with the simplest one of all, namely the simple harmonic oscillator? And indeed, is it not possible to apply his procedure to the simple harmonic oscillator without going through the apparent device of introducing a perturbation, and then retaining only those parts of the solution that survive as the perturbation vanishes?

For the simple harmonic oscillator, the equation of motion is $\ddot{x}+\omega_0^2 x=0$, which yields

$$[\omega_0^2-\omega^2(n,n-\alpha)]X(n,n-\alpha)=0 \tag{71}$$

for the amplitudes X and frequencies ω. It is reasonable to retain the quantum condition, Eq. (16), because this condition is supposed to hold for any force law. If we assume that the only nonvanishing amplitudes are those involving adjacent states (because, for example, in the classical case only a single harmonic is present[56]), then because $X(n,n-1)=\frac{1}{2}a(n,n-1)$, Eqs. (16) and (71) reduce to Eqs. (50) and (48), respectively, and we quickly recover our previous results. This is indeed an efficient way to solve the quantum simple harmonic oscillator.[57] For completeness, however, it would be desirable not to have to make the adjacent states assumption. Born and Jordan[4] showed how this could be done, but their argument is somewhat involved. Soon thereafter, of course, the wave mechanics of Schrödinger and the operator approach of Dirac provided the derivations used ever since.

D. The solutions up to and including λ^2 terms

We now turn to the higher order corrections for the λx^2 term. Consider Eq. (44) and retain terms of order λ. We set [see Eqs. (25) and (26)]

$$\omega(n,n-1)=\omega_0+\lambda\omega^{(1)}(n,n-1), \tag{72}$$

$$a(n,n-1)=a^{(0)}(n,n-1)+\lambda a^{(1)}(n,n-1), \tag{73}$$

and find

$$2\lambda\omega_0\omega^{(1)}(n,n-1)a^{(0)}(n,n-1)=0, \tag{74}$$

so that

$$\omega^{(1)}(n,n-1)=0. \tag{75}$$

If we consider Eq. (44) up to terms of order λ^2 and employ Eqs. (53), (57), and (61) for the zeroth-order amplitudes, we obtain the $O(\lambda^2)$ correction to $\omega(n,n-1)$ [see Eq. (26)]:

$$\omega^{(2)}(n,n-1)=-\frac{5\beta^2}{12\omega_0^3}n. \tag{76}$$

The corresponding corrections to $a(n,n-1)$ are found from the quantum condition Eq. (16). To order λ we set

$$a(n+1,n)=a^{(0)}(n+1,n)+\lambda a^{(1)}(n+1,n), \tag{77}$$

as in Eq. (73), and find

$$\sqrt{n+1}a^{(1)}(n+1,n)-\sqrt{n}a^{(1)}(n,n-1)=0. \tag{78}$$

Equation (78) has the solution $a^{(1)}(n,n-1)=\text{constant}/\sqrt{n}$, but the condition $a^{(1)}(0,-1)=0$ [see Eq. (52)] implies that the constant must be zero, and so

$$a^{(1)}(n,n-1)=0. \tag{79}$$

In a similar way, we obtain to order λ^2

$$\sqrt{n+1}a^{(2)}(n+1,n)-\sqrt{n}a^{(2)}(n,n-1)=\frac{11\beta^3}{72\omega_0^4}(2n+1), \tag{80}$$

which has the solution

$$a^{(2)}(n,n-1)=\frac{11\beta^3}{72\omega_0^4}n\sqrt{n}. \qquad (81)$$

We now find the higher order corrections to $a(n,n)$ by considering Eq. (43). We obtain $a^{(1)}(n,n)=0$ and

$$a^{(2)}(n,n)=-\frac{\beta^4}{72\omega_0^6}(30n^2+30n+11). \qquad (82)$$

Similarly, we find from Eq. (45) $a^{(1)}(n,n-2)=0$ and

$$a^{(2)}(n,n-2)=\frac{3\beta^4}{32\omega_0^6}(2n-1)\sqrt{n(n-1)}, \qquad (83)$$

where we have used

$$\omega^{(2)}(n,n-2)=\omega^{(2)}(n,n-1)+\omega^{(2)}(n-1,n-2)$$
$$=-\frac{5\beta^2}{12\omega_0^3}(2n-1). \qquad (84)$$

These results suffice for our purpose. If n is large, they agree with those obtained for the classical λx^2 anharmonic oscillator using the method of successive approximations.[58]

As an indirect check of their quantum mechanical validity, we now turn to the energy evaluated to order λ^2. Consider first the (n,n) element of $\frac{1}{2}m\omega_0^2\hat{x}^2$. This matrix element is given to order λ^2, by

$$\frac{1}{2}m\omega_0^2\left\{\frac{1}{4}[(a^{(0)}(n,n-1))^2+(a^{(0)}(n,n+1))^2]\right.$$
$$+\frac{\lambda^2}{4}[4(a^{(0)}(n,n))^2+2a^{(2)}(n,n-1)a^{(0)}(n-1,n)$$
$$+2a^{(2)}(n,n+1)a^{(0)}(n+1,n)+(a^{(0)}(n,n-2))^2$$
$$\left.+(a^{(0)}(n,n+2))^2]\right\}=\frac{1}{2}m\omega_0^2\left[\frac{\beta^2}{2}\left(n+\frac{1}{2}\right)\right.$$
$$\left.+\frac{5\beta^4\lambda^2}{12\omega_0^4}(n^2+n+11/30)\right]. \qquad (85)$$

Similarly, using Eq. (67) up to order λ^2, with $\alpha=0$, the (n,n) element of $\frac{1}{2}m\dot{\hat{x}}^2$ is found to be

$$\frac{1}{2}m\omega_0^2\left[\frac{\beta^2}{2}\left(n+\frac{1}{2}\right)-\frac{5\beta^4\lambda^2}{24\omega_0^4}(n^2+n+11/30)\right]. \qquad (86)$$

Finally we consider the (n,n) element of the potential energy $\frac{1}{3}m\lambda\hat{x}^3$. To obtain the result to order λ^2, we need to calculate the (n,n) element of $\hat{x}^3$ only to order λ. If we use

$$\hat{x}^3(n,n)=\sum_\alpha\sum_\beta X(n,n-\alpha)X(n-\alpha,n-\beta)$$
$$\times X(n-\beta,n), \qquad (87)$$

we find that there are no zeroth-order terms, but twelve terms of order λ [recall that amplitudes such as $X(n,n)$ and $X(n,n-2)$ each carry one power of λ]. We evaluate these terms using Eqs. (53), (57), and (61), and obtain

$$-\frac{5m\lambda^2\beta^4}{24\omega_0^2}(n^2+n+11/30) \qquad (88)$$

for this term in the energy. If we combine Eqs. (85), (86), and (88), we obtain the energy up to order λ^2,

$$W(n,n)=\left(n+\frac{1}{2}\right)\hbar\omega_0-\frac{5\lambda^2\hbar^2}{12m\omega_0^4}(n^2+n+11/30), \qquad (89)$$

a result[59] that agrees with classical perturbation theory when n is large,[60] and is in agreement with standard second-order perturbation theory in quantum mechanics.[61]

As mentioned, Heisenberg did not give results for the λx^2 term beyond zeroth order. He did, however, give the results for the λx^3 term up to and including λ^2 terms in the energy, and λ terms in the amplitudes. By "the energy" we mean, as usual, the (n,n) element of the energy matrix, which as noted in Sec. III C is independent of time. We also should check that the off-diagonal elements $W(n,n-\alpha)$ vanish [see Eq. (69)]. These are the terms that would (if nonzero) carry a periodic time-dependence, and Heisenberg wrote[62] that "I could not prove in general that all periodic terms actually vanish, but this was the case for all the terms evaluated." We do not know how many off-diagonal terms $W(n,n-\alpha)$ he evaluated, but he clearly regarded their vanishing as a crucial test of the formalism. In Appendix B we outline the calculation of all off-diagonal terms for the λx^2 term up to order λ, as an example of the kind of calculation Heisenberg probably did, finishing it late one night on Heligoland.[63]

IV. CONCLUSION

We have tried to remove some of the barriers to understanding Heisenberg's 1925 paper by providing the details of calculations of the type we believe he performed. We hope that more people will thereby be encouraged to appreciate this remarkable paper.

The fact is that Heisenberg's "amplitude calculus" works, at least for the simple one-dimensional problems to which he applied it. It is an eminently practical procedure, requiring no sophisticated mathematical knowledge to implement. Because it uses the correct equations of motion and incorporates the fundamental commutator, Eq. (11), via the quantum condition, Eq. (16), the answers obtained are correct, in agreement with conventional quantum mechanics.

We believe that Heisenberg's approach, as applied to simple dynamical systems, has much pedagogical value, and could usefully be included in undergraduate courses on quantum mechanics. The multiplication rule, Eq. (10), has a convincing physical rationale, even for those who (like Heisenberg) do not recognize it as matrix multiplication. Indeed, this piece of quantum physics could provide an exciting application for those learning about matrices in a concurrent mathematics course. The simple examples of Eq. (10), in equations such as Eq. (22) or the analogous one for the $\lambda\hat{x}^3$ term, introduce students directly to the fundamental quantum idea that a transition from one state to another occurs via all possible intermediate states, something that can take time to emerge in the traditional wave-mechanical approach. The solution of the quantum simple harmonic oscillator, sketched at the end of Sec III D, is simple in comparison with the standard methods. Finally, the type of perturbation theory employed here provides an instructive introduction to the technique, being more easily related to the classical analysis than is conventional quantum-mechanical perturbation theory (which students tend to find very formal).

It is true that many important problems in quantum mechanics are much more conveniently handled in the wave-mechanical formalism: unbound problems are an obvious example, but even the Coulomb problem required a famous *tour de force* by Pauli.[64] Nevertheless, a useful seed may be sown, so that when students meet problems involving a finite number of discrete states—for example, in the treatment of spin—the introduction of matrices will come as less of a shock. And they may enjoy the realization that the somewhat mysteriously named ''matrix elements'' of wave mechanics are indeed the elements of Heisenberg's matrices.

APPENDIX A: THE QUANTUM CONDITION, EQ. (16), AND $\hat{x}\hat{p}-\hat{p}\hat{x}=i\hbar$

Consider the (n,n) element of $(\hat{x}\dot{\hat{x}}-\dot{\hat{x}}\hat{x})$, which is

$$\sum_\alpha X(n,n-\alpha)i\omega(n-\alpha,n)X(n-\alpha,n) - \sum_\alpha i\omega(n,n-\alpha)X(n,n-\alpha)X(n-\alpha,n). \tag{A1}$$

In the first term of Eq. (A1), the sum over $\alpha>0$ may be rewritten as

$$-i\sum_{\alpha>0}\omega(n,n-\alpha)|X(n,n-\alpha)|^2 \tag{A2}$$

using $\omega(n,n-\alpha)=-\omega(n-\alpha,n)$ from Eq. (1) and $X(n-\alpha,n)=X^*(n,n-\alpha)$ from Eq. (38). Similarly, the sum over $\alpha<0$ becomes

$$i\sum_{\alpha>0}\omega(n+\alpha,n)|X(n+\alpha,n)|^2 \tag{A3}$$

on changing α to $-\alpha$. Similar steps for the second term of Eq. (A1) lead to the result

$$\begin{aligned}(\hat{x}\dot{\hat{x}}-\dot{\hat{x}}\hat{x})(n,n)&=2i\sum_{\alpha>0}[\omega(n+\alpha,n)|X(n+\alpha,n)|^2\\&\quad-\omega(n,n-\alpha)|X(n,n-\alpha)|^2]\\&=2ih/(4\pi m),\end{aligned} \tag{A4}$$

where the last step follows from Eq. (16). We set $\hat{p}=m\dot{\hat{x}}$ and find

$$(\hat{x}\hat{p}-\hat{p}\hat{x})(n,n)=i\hbar \tag{A5}$$

for all values of n. Equation (A5) was found by Born[43] shortly after reading Heisenberg's paper. In further developments the value of the fundamental commutator $\hat{x}\hat{p}-\hat{p}\hat{x}$, namely $i\hbar$, was taken to be a basic postulate. The sum rule in Eq. (16) is then derived by taking the (n,n) matrix element of the relation $[\hat{x},[\hat{H},\hat{x}]]=\hbar^2/m$.

APPENDIX B: CALCULATION OF THE OFF-DIAGONAL MATRIX ELEMENTS OF THE ENERGY $W(n,n-\alpha)$ FOR THE λx^2 TERM

We shall show that, for $\alpha\neq 0$, all the elements $(n,n-\alpha)$ of the energy operator $\frac{1}{2}m\dot{\hat{x}}^2+\frac{1}{2}m\omega_0^2\hat{x}^2+\frac{1}{3}\lambda m\hat{x}^3$ vanish up to order λ. We begin by noting that at any given order in λ, only a limited number of elements $W(n,n-1),W(n,n-2),\ldots$ will contribute, because the amplitudes $X(n,n-\alpha)$ are suppressed by increasing powers of λ as α increases. In fact, for $\alpha\geqslant 2$ the leading power of λ in $W(n,n-\alpha)$ is $\lambda^{\alpha-2}$, which arises from terms such as $X(n,n-1)X(n-1,n-\alpha)$ and $\lambda X(n,n-1)X(n-1,n-2)X(n-2,n-\alpha)$. Thus to order λ, we need to calculate only $W(n,n-1),W(n,n-2),W(n,n-3)$.

(a) $W(n,n-1)$. There are four $O(\lambda)$ contributions to the $(n,n-1)$ element of $\frac{1}{2}m\omega_0^2\hat{x}^2$:

$$\begin{aligned}&\tfrac{1}{4}m\omega_0^2\lambda\{a^{(0)}(n,n)a^{(0)}(n,n-1)+a^{(0)}(n,n-1)\\&\quad\times a^{(0)}(n-1,n-1)+\tfrac{1}{2}[a^{(0)}(n,n+1)a^{(0)}(n+1,n-1)\\&\quad+a^{(0)}(n,n-2)a^{(0)}(n-2,n-1)]\}\\&\quad=-\tfrac{5}{24}m\lambda\beta^3 n\sqrt{n}.\end{aligned} \tag{B1}$$

There are two $O(\lambda)$ contributions to the $(n,n-1)$ element of $\frac{1}{2}m\dot{\hat{x}}^2$:

$$\begin{aligned}&-\tfrac{1}{8}\lambda m\{\omega^{(0)}(n,n+1)\omega^{(0)}(n+1,n-1)a^{(0)}(n,n+1)\\&\quad\times a^{(0)}(n+1,n-1)+\omega^{(0)}(n,n-2)\omega^{(0)}(n-2,n-1)\\&\quad\times a^{(0)}(n,n-2)a^{(0)}(n-2,n-1)\}=\tfrac{1}{12}m\lambda\beta^3 n\sqrt{n}.\end{aligned} \tag{B2}$$

There are three $O(\lambda)$ contributions to the $(n,n-1)$ element of $\frac{1}{3}m\lambda\hat{x}^3$:

$$\begin{aligned}&\tfrac{1}{24}m\lambda\{a^{(0)}(n,n-1)a^{(0)}(n-1,n)a^{(0)}(n,n-1)\\&\quad+a^{(0)}(n,n-1)a^{(0)}(n-1,n-2)a^{(0)}(n-2,n-1)\\&\quad+a^{(0)}(n,n+1)a^{(0)}(n+1,n)a^{(0)}(n,n-1)\}\\&\quad=\tfrac{1}{8}m\lambda\beta^3 n\sqrt{n}.\end{aligned} \tag{B3}$$

The sum of Eqs. (B1)–(B3) vanishes, as required.

(b) $W(n,n-2)$. The leading contribution is independent of λ. From the term $\frac{1}{2}m\omega_0^2\hat{x}^2$, it is

$$\tfrac{1}{8}m\omega_0^2 a^{(0)}(n,n-1)a^{(0)}(n-1,n-2), \tag{B4}$$

which is canceled by the corresponding term from $\frac{1}{2}m\dot{\hat{x}}^2$. The next terms are $O(\lambda^2)$, for example from the leading term in the $(n,n-2)$ element of $\frac{1}{3}\lambda m\hat{x}^3$.

(c) $W(n,n-3)$. There are two $O(\lambda)$ contributions from $\frac{1}{2}m\omega_0^2\hat{x}^2$:

$$\begin{aligned}&\tfrac{1}{8}m\omega_0^2\lambda\{a^{(0)}(n,n-1)a^{(0)}(n-1,n-3)\\&\quad+a^{(0)}(n,n-2)a^{(0)}(n-2,n-3)\}\\&\quad=\tfrac{1}{24}m\lambda\beta^3\sqrt{n(n-1)(n-2)}.\end{aligned} \tag{B5}$$

There are two $O(\lambda)$ contributions from $\frac{1}{2}m\dot{\hat{x}}^2$:

$$\begin{aligned}&-\tfrac{1}{8}m\lambda\{\omega^{(0)}(n,n-1)a^{(0)}(n,n-1)\omega^{(0)}(n-1,n-3)\\&\quad\times a^{(0)}(n-1,n-3)+\omega^{(0)}(n,n-2)\\&\quad\times a^{(0)}(n,n-2)\omega^{(0)}(n-2,n-3)a^{(0)}(n-2,n-3)\}\\&\quad=-\tfrac{1}{12}\lambda m\beta^3\sqrt{n(n-1)(n-2)}.\end{aligned} \tag{B6}$$

There is only one $O(\lambda)$ contribution from $\frac{1}{3}m\lambda\hat{x}^3$:

$$\frac{1}{24} m\lambda a^{(0)}(n,n-1)a^{(0)}(n-1,n-2)a^{(0)}(n-2,n-3) = \frac{1}{24}\lambda m \beta^3 \sqrt{n(n-1)(n-2)}. \tag{B7}$$

The sum of Eqs. (B5)–(B7) vanishes, as required.

[a)]Electronic mail: i.aitchisonl@physics.oxford.ac.uk

[1]W. Heisenberg, ''Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen,'' Z. Phys. **33**, 879–893 (1925).

[2]This is the title of the English translation, which is paper 12 in Ref. 3, pp. 261–276. We shall refer exclusively to this translation, and to the equations in it as (H1), (H2), ...

[3]*Sources of Quantum Mechanics*, edited by B. L. van der Waerden (North-Holland, Amsterdam, 1967). A collection of reprints in translation.

[4]M. Born and P. Jordan, ''Zur Quantenmechanik,'' Z. Phys. **34**, 858–888 (1925), paper 13 in Ref. 3.

[5]P. A. M. Dirac, ''The fundamental equations of quantum mechanics,'' Proc. R. Soc. London, Ser. A **109**, 642–653 (1926), paper 14 in Ref. 3.

[6]M. Born, W. Heisenberg, and P. Jordan, ''Zur Quantenmechanik II,'' Z. Phys. **35**, 557–615 (1926), paper 15 in Ref. 3.

[7]S. Weinberg, *Dreams of a Final Theory* (Pantheon, New York, 1992), pp. 53–54. Weinberg goes on to say that ''Perhaps we should not look too closely at Heisenberg's first paper ... .'' We will not follow his suggestion here.

[8]S.-I. Tomonaga, *Quantum Mechanics: Old Quantum Theory* (North-Holland, Amsterdam, 1962), Vol. 1.

[9]M. Jammer, *The Conceptual Development of Quantum Mechanics* (McGraw–Hill, New York, 1966).

[10]E. MacKinnon, ''Heisenberg, models and the rise of matrix mechanics,'' Hist. Stud. Phys. Sci. **8**, 137–188 (1977).

[11]J. Mehra and H. Rechenberg, *The Historical Development of Quantum Theory* (Springer-Verlag, New York, 1982), Vol 2.

[12]J. Hendry, *The Creation of Quantum Mechanics and the Bohr-Pauli Dialogue* (Reidel, Dordrecht, 1984).

[13]T.-Y. Wu, *Quantum Mechanics* (World Scientific, Singapore, 1986).

[14]M. Taketani and M. Nagasaki, *The Formation and Logic of Quantum Mechanics* (World Scientific, Singapore, 2002), Vol. 3.

[15]G. Birtwistle, *The New Quantum Mechanics* (Cambridge U.P., Cambridge, 1928).

[16]M. Born, *Atomic Physics* (Dover, New York, 1989).

[17]J. Lacki, ''Observability, Anschaulichkeit and abstraction: A journey into Werner Heisenberg's science and philosophy,'' Fortschr. Phys. **50**, 440–458 (2002).

[18]J. Mehra, *The Golden Age of Theoretical Physics* (World Scientific, Singapore, 2001), Vol. 2.

[19]Of these the most detailed are Ref. 3, pp. 28–35, Ref. 8, pp. 204–224, Ref. 10, pp. 161–188, and Ref. 11, Chap. IV.

[20]Reference 7, p. 53.

[21]Reference 2, p. 262.

[22]All quotations are from the English translation of Ref. 2.

[23]We use ω rather than Heisenberg's ν.

[24]We depart from the notation of Refs. 1 and 2, preferring that of Ref. 8, pp. 204–224. Our $X_\alpha(n)$ is Heisenberg's $a_\alpha(n)$.

[25]The reader may find it helpful at this point to consult Ref. 26, which provides a clear account of the connection between the classical analysis of an electron's periodic motion and simple quantum versions. See also J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975), 2nd ed., Sec. 9.2.

[26]W. A. Fedak and J. J. Prentis, ''Quantum jumps and classical harmonics,'' Am. J. Phys. **70**, 332–344 (2002).

[27]The association $X_\alpha(n) \leftrightarrow X(n,n-\alpha)$ is generally true only for non-negative α. For negative values of α, a general term in the classical Fourier series is $X_{-|\alpha|}(n)\exp[-i\omega(n)|\alpha|t]$. If we replace $-\omega(n)|\alpha|$ by $-\omega(n,n-|\alpha|)$, which is equal to $\omega(n-|\alpha|,n)$ using Eq. (1), we see that $X_{-|\alpha|}(n) \leftrightarrow X(n-|\alpha|,n)$. The association $X_{-|\alpha|} \leftrightarrow X(n,n+|\alpha|)$ would not be correct because $\omega(n,n+|\alpha|)$ is not the same, in general, as $\omega(n-|\alpha|,n)$.

[28]Conventional notation, subsequent to Ref. 4, would replace $n-\alpha$ by a second index m, say. We prefer to remain as close as possible to the notation of Heisenberg's paper.

[29]Reference 2, p. 264.

[30]Reference 2, p. 265.

[31]Reference 2, p. 266.

[32]Reference 2 p. 267.

[33]This step apparently did not occur to him immediately. See Ref. 11, p. 231.

[34]Actually not quite. We have taken the liberty of changing the order of the arguments in the first terms in the braces; this (correct) order is as given in the equation Heisenberg wrote before Eq. (H20).

[35]W. Thomas, ''Über die Zahl der Dispersionelektronen, die einem stationären Zustande zugeordnet sind (Vorläufige Mitteilung),'' Naturwissenschaften **13**, 627 (1925).

[36]W. Kuhn, ''Über die Gesamtstärke der von einem Zustande ausgehenden Absorptionslinien,'' Z. Phys. **33**, 408–412 (1925), paper 11 in Ref. 3.

[37]W. Heisenberg, as discussed in Ref. 11, pp. 243 ff.

[38]Reference 9, p. 193; Φ is any function defined for stationary states.

[39]M. Born, ''Über Quantenmechanik,'' Z. Phys. **26**, 379–395 (1924), paper 7 in Ref. 3.

[40]Reference 9, p. 202.

[41]H. A. Kramers and W. Heisenberg, ''Über die Streuung von Strahlen durch Atome,'' Z. Phys. **31**, 681–707 (1925), paper 10 in Ref. 3.

[42]For considerable further detail on dispersion theory, sum rules, and the ''discretization'' rules, see Ref. 8, pp. 142–147, 206–208, and Ref. 9, Sec. 4.3.

[43]See Ref. 3, p. 37.

[44]Reference 2, p. 268.

[45]For an interesting discussion of the possible reasons why he chose to try out his method on the anharmonic oscillator, see Ref. 11, pp. 232–235; and Ref. 3, p. 22. Curiously, most of the commentators—with the notable exception of Tomonaga (Ref. 8)—seem to lose interest in the details of the calculations at this point.

[46]The $(n,n-\alpha)$ matrix element, in the standard terminology.

[47]Equation (22) is not in Heisenberg's paper, although it is given by Tomonaga, Ref. 8, Eq. (32.20′).

[48]Note that this means that, in $x(t)$, all the terms which are of order λ^p arise from many different terms in Eq. (24).

[49]Except that Heisenberg relabeled most of the a_α's as $a_\alpha(n)$.

[50]Because of an oversight, he wrote $a_0(n)$ in place of $a(n,n)$ in Eq. (33).

[51]MacKinnon (Ref. 10) suggests how, in terms of concepts from the ''virtual oscillator'' model, Eq. (28) may be transformed into Eqs. (33)–(36). We do not agree with MacKinnon (Ref. 10, footnote 62) regarding ''mistakes'' in Eqs. (35) and (36).

[52]In the version of the quantum condition that Heisenberg gave just before Eq. (H20), he unfortunately used the same symbol for the transition amplitudes as in (H16)–see Eq. (16)—but replaced $4\pi m$ by πm, not explaining where the factor 1/4 came from [see Eq. (42)]; he also omitted the λ's.

[53]Heisenberg omitted the superscripts.

[54]See, for example, L. I. Schiff, *Quantum Mechanics* (McGraw–Hill, New York, 1968), 3rd ed., p. 72.

[55]Reference 2, p. 272.

[56]This is the justification suggested in Ref. 4, p. 297 of paper 13 in Ref. 3.

[57]It is essentially that given by L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Pergamon, Oxford, 1977), 3rd ed., pp. 67–68.

[58]L. D. Landau and E. M. Lifshitz, *Mechanics* (Pergamon, Oxford, 1976), 3rd ed., pp. 86–87.

[59]This equation corresponds to Eq. (88) of Ref. 4, in which there appears to be a misprint of 17/30 instead of 11/30.

[60]See Ref. 4, p. 305.

[61]Reference 57, p. 136.

[62]Reference 2, pp. 272–3.

[63]See W. Heisenberg, *Physics and Beyond* (Allen & Unwin, London, 1971), p. 61.

[64]W. Pauli, ''Über das Wasserstoffspektrum vom Standpunkt der neuen Quantenmechanik,'' Z. Phys. **36**, 336–363 (1926), paper 16 in Ref. 3.