

On time-dependent perturbation theory in matrix mechanics and time averaging

Fernando Casas*

Abstract

The time dependent quantum perturbation theory developed by Born, Heisenberg and Jordan in 1926 is revisited. We show that it not only reproduces the standard theory formulated in the interaction picture, but also allows one to construct more accurate approximations if time averaging techniques are employed. The theory can be rendered unitary even if the expansion is truncated by using a transformation previously suggested by Heisenberg. We illustrate the main features of the procedure on a simple example which clearly shows its advantages in comparison with the standard perturbation theory.

Institut de Matemàtiques i Aplicacions de Castelló (IMAC) and
Departament de Matemàtiques, Universitat Jaume I, E-12071 Castellón,
Spain.

1 Introduction

In their formulation of matrix mechanics in the context of the (then) new quantum mechanics, Born, Heisenberg and Jordan [13, 7, 8] were guided by two basic principles. First, the new formalism had to be kept as close to classical Hamiltonian dynamics as possible. Thus, the dynamical equations satisfied by the quantum analogues \hat{q} and \hat{p} of the classical coordinates and momenta, respectively, had the same form as in the classical theory, but the variables no longer satisfied the commutative property of multiplication. Second, the formalism had to be consistent with Bohr's correspondence principle: the quantum theory should reproduce classical mechanics in the limit of large quantum numbers. This led them, in particular, to represent the new variables as (infinite) Hermitian matrices, whereas the eigenvalues of the Hamiltonian $H(\hat{q}, \hat{p})$ (the energy values of the system) allowed them to get the frequencies of transition.

The first guideline is much evident, in particular, in their quantum-mechanical perturbation theory for non-degenerate and degenerate systems. According to Heisenberg, "the analogy with the classical Hamilton–Jacobi technique was the beginning of all efforts; what we did first was just to try to imitate the old

*Email: Fernando.Casas@uji.es

methods as closely as we could” [16, p. 103]. The purpose was then, in analogy with classical perturbation theory, to construct perturbatively a “canonical transformation” S on the variables \hat{q} and \hat{p} to a new set of variables in which the Hamiltonian W of the perturbed system is diagonal. When dealing with time-dependent external forces, an appropriate perturbation scheme was formulated from the preceding one, this time involving an explicitly time-dependent transformation matrix $S(t)$. At each step of the procedure, a choice had to be made for the matrix W and subsequently the transformation S was determined by solving the corresponding equation.

By contrast, the procedure followed in the wave mechanics approach and later in the standard formulation of quantum mechanics was more analytical in character: the purpose is to get the quantum time-evolution operator $U(t)$. If the Hamiltonian H does not depend on time, then one has to solve the time-independent Schrödinger equation $H\psi = E\psi$. This is done by determining both the eigenvalues E_n and the eigenfunctions ψ_n as a power series of the perturbation parameter, whereas if H depends explicitly on time, $U(t)$ is constructed by considering the interaction representation, i.e., as a factorization of the evolution operator corresponding to the unperturbed part of the Hamiltonian operator and the evolution of an appropriately transformed Hamiltonian describing the perturbation, which is then approximated as a power series in the perturbation parameter [17].

Both perturbation theories lead of course to the same results when the Hamiltonian is independent of time [23], although the original connection with classical mechanics is somehow lost in wave mechanics, particularly when dealing with time-dependent problems. Perhaps for this reason, several papers have been published along the years emphasizing this relation, by using techniques of classical mechanics and dynamical systems in a quantum mechanical context: normal forms [1, 12], averaging theory to avoid secular terms [14, 19], unitary transformations in time-dependent perturbation theory [20, 21, 11, 9], etc.

Here we consider once again the time-dependent perturbation theory as originally proposed by Born, Heisenberg and Jordan in [8] and analyze the formalism after expressing it in modern quantum mechanics language. We show that this technique is indeed more general than the standard procedure of wave mechanics, in the sense that different choices for the new Hamiltonian are possible, thus leading to different kinds of approximations. It is with a particular election of W , namely by taking as W the unperturbed part of the original Hamiltonian, that the standard perturbation theory is recovered. We then explore other options that avoids the presence of secular terms by using high-order averaging, a well known technique in classical perturbation theory [3]. Although the approximations resulting from truncating the series are not unitary, they turn out to be more accurate than those arising in the standard treatment over longer time intervals. Moreover, a unitary formalism can be obtained even when truncating the series, by following the initial approach designed by Heisenberg. Here we analyze this alternative and illustrate the different procedures in a simple two-level system whose exact solution can be obtained in closed form.

We believe the approach proposed by Born, Heisenberg and Jordan has sev-

eral advantages when dealing with time-dependent perturbations in quantum mechanics. On the one hand, the procedure is formally similar to the time-independent case. On the other hand, the treatment is obtained by applying standard techniques of classical perturbation theory (transformations of variables, averaging), so that it highlights the connection with classical mechanics, something that might be helpful for students taking a course in quantum mechanics but already familiar with the mathematical aspects of classical and celestial mechanics. In addition, the standard perturbation theory is easily recovered in this framework, if desired, although by applying averaging it is possible to get more accurate results.

2 Standard time-dependent perturbation (STDP) theory in quantum mechanics

In the standard quantum mechanical treatment of a given quantum system interacting with an external environment, such as that created by an external field explicitly depending on time, one considers a constant (non perturbed) Hamiltonian H_0 associated with the system and an additional time-dependent part $\tilde{H}(t, \varepsilon)$, so that the dynamics is governed by the Hamiltonian

$$H(t, \varepsilon) = H_0 + \tilde{H}(t, \varepsilon) \equiv H_0 + \sum_{n=1}^{\infty} \varepsilon^n H_n(t). \quad (1)$$

Here, for generality, we assume that $\tilde{H}(t, \varepsilon)$ depends on a small parameter ε that controls the perturbation in such a way that $\tilde{H}(t, 0) = 0$. In many situations, though, the power series expansion of $\tilde{H}(t, \varepsilon)$ collapses just to the first term: $\tilde{H}(t, \varepsilon) = \varepsilon H_1(t)$.

Under these circumstances the goal consists in determining the time-evolution unitary operator $U_H(t, t_0)$ by solving the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} U_H(t, t_0) = H(t, \varepsilon) U_H(t, t_0), \quad U(t_0, t_0) = I. \quad (2)$$

The state at time t , characterized by the wave function $\psi(t)$, is then related to the initial state $\psi(t_0)$ by $\psi(t) = U(t, t_0)\psi(t_0)$. If $H(t, \varepsilon) = H_0$ the solution of (2) reads $U_{H_0}(t, t_0) = \exp(-i(t - t_0)H_0/\hbar)$, but the general case is not so simple. A convenient way to proceed consists in transforming the problem into the interaction picture [17], i.e. the solution is factorized as

$$U_H(t, t_0) = e^{-i(t-t_0)H_0/\hbar} U_I(t, t_0), \quad (3)$$

where the unknown operator U_I obeys the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} U_I(t, t_0) = H_I(t, \varepsilon) U_I(t, t_0), \quad U_I(t_0, t_0) = I, \quad (4)$$

with the new Hamiltonian given by

$$H_I(t, \varepsilon) = e^{i(t-t_0)H_0/\hbar} \tilde{H}(t, \varepsilon) e^{-i(t-t_0)H_0/\hbar} \equiv \sum_{n=1}^{\infty} \varepsilon^n H_n^I(t). \quad (5)$$

In this way, the transformation defined by (3) allows to incorporate the solvable piece H_0 of the Hamiltonian and focus the problem in the (approximate) integration of the perturbation as given by (5). In general, however, it is not possible to get analytical solutions of (4), and so one must turn to a perturbative analysis, looking for an expansion of U_I in terms of the parameter ε ,

$$U_I(t, t_0) = \sum_{n=0}^{\infty} \varepsilon^n U_n(t, t_0), \quad (6)$$

where $U_n(t, t_0)$ stands for the contribution of order n in ε . This can be obtained by substituting (6) into (4) and then equating terms of the same power in ε . Thus, for the first terms we get

$$\begin{aligned} U_0(t, t_0) &= I, & U_1(t, t_0) &= -\frac{i}{\hbar} \int_{t_0}^t ds H_1^I(s) \\ U_2(t, t_0) &= -\frac{i}{\hbar} \int_{t_0}^t ds H_2^I(s) + \left(-\frac{i}{\hbar}\right)^2 \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 H_1^I(t_1) H_1^I(t_2), \end{aligned}$$

where we have denoted $H_j^I(t) \equiv e^{i(t-t_0)H_0/\hbar} H_j(t) e^{-i(t-t_0)H_0/\hbar}$. If $\tilde{H} = \varepsilon H_1$, then the general expression of U_n in (6) is given by

$$U_n(t, t_0) = \left(-\frac{i}{\hbar}\right)^n \int_{t_0}^t dt_1 \cdots \int_{t_0}^{t_{n-1}} dt_n H_1^I(t_1) H_1^I(t_2) \cdots H_1^I(t_n). \quad (7)$$

Very often, it is sufficient to determine the first few terms to describe transition probabilities between states. The resulting approximations, however, present some undesirable features. First, any truncation of the infinite sum (6) is no longer unitary, and thus the computed transition probability between different quantum states may exceed unity. Second, secular terms in time (i.e., terms of the form t^m) may appear at each order, and so the quality of the approximation degrades considerably with time. In this way, one cannot represent the solution for all t using a finite number of terms in the series, and is only when the infinite series is summed up that the unitary character is restored.

Example. We will use as illustration along the paper the simple two-level quantum system described by the Hamiltonian

$$H(t, \varepsilon) = \frac{1}{2} \hbar \omega_0 \sigma_3 + \varepsilon (\sigma_1 \cos \omega t + \sigma_2 \sin \omega t), \quad (8)$$

where σ_j denote the Pauli matrices and $\varepsilon, \omega_0, \omega \neq \omega_0$ are real parameters. If 1 and 2 stand for the spin up and down states, respectively, the exact transition probability from state 1 to state 2 is given by ($t_0 = 0$)

$$|(U_H(t))_{21}|^2 = \left(\frac{2\varepsilon}{\hbar \omega'} \sin \frac{\omega' t}{2} \right)^2 \quad \text{with} \quad \omega' = \left((\omega - \omega_0)^2 + \frac{4\varepsilon^2}{\hbar^2} \right)^{1/2}. \quad (9)$$

Clearly, in this case $H(t, \varepsilon) = H_0 + \varepsilon H_1(t)$, with

$$H_0 = \frac{1}{2} \hbar \omega_0 \sigma_3, \quad H_1(t) = \sigma_1 \cos \omega t + \sigma_2 \sin \omega t \quad (10)$$

and $H_n = 0$, $n \geq 2$. When the STDP theory is applied to this Hamiltonian we get

$$\begin{aligned} U_1(t) &= -\frac{i}{\hbar(\omega - \omega_0)} (\sin(\omega - \omega_0)t \sigma_1 + (1 - \cos(\omega - \omega_0)t) \sigma_2) \\ U_2(t) &= \frac{1}{\hbar^2(\omega - \omega_0)^2} ((-1 + \cos(\omega - \omega_0)t) I + i((\omega - \omega_0)t - \sin(\omega - \omega_0)t) \sigma_3) \end{aligned}$$

We see then that secular terms already appear at order 2, and so one expect that the truncated (non unitary) approximation obtained from (3)

$$U_H(t) \approx (\cos(\omega_0 t/2) I - i \sin(\omega_0 t/2) \sigma_3) (I + \varepsilon U_1(t) + \varepsilon^2 U_2(t)) \quad (11)$$

will be valid only for short times. The resulting transition probability up to this order is

$$P_t = \frac{2\varepsilon^2(1 - \cos(\omega - \omega_0)t)}{\hbar^2(\omega - \omega_0)^2}. \quad (12)$$

Secular terms appear in P_t at higher orders in ε .

3 Time dependent perturbation theory in matrix mechanics

Historically, however, the first treatment of perturbations in quantum mechanics proceeded in a different way, more related to the usual approach followed in classical Hamiltonian mechanics. The objective there is to construct a canonical transformation (a symplectic change of coordinates in phase space) such that in the new variables the dynamics of the resulting Hamiltonian is easily solved. This transformation can be obtained in particular by solving the corresponding Hamilton–Jacobi equation [3, 15].

Analogously, the idea of Born, Heisenberg and Jordan in [8] when developing a perturbation theory in matrix mechanics was to construct a unitary transformation that renders ‘solvable’ the Hamiltonian (see e.g. [4]). In their procedure the (perturbed) system to be solved is described by a Hamiltonian H of the form

$$H = H_0 + \varepsilon H_1 + \varepsilon^2 H_2 + \dots, \quad (13)$$

where the dynamics corresponding to H_0 is known. This means that the Hamiltonian H_0 is diagonal in the momentum $\hat{p}_0(t)$ and coordinate $\hat{q}_0(t)$, so that their time evolution is given by

$$\hat{p}_0(t) = ((p_0)_{mn} e^{i\omega_{mn}t}), \quad \hat{q}_0(t) = ((q_0)_{mn} e^{i\omega_{mn}t}),$$

where $(p_0)_{mn}$, $(q_0)_{mn}$ are the quantum amplitudes of the operators \hat{p}_0 and \hat{q}_0 , respectively, and ω_{mn} is the frequency of transition between two stationary states with energies E_m^0 and E_n^0 , so that $\hbar\omega_{mn} = E_m^0 - E_n^0$. In addition, \hat{q}_0 and \hat{p}_0 satisfy the quantum condition $[\hat{q}_0, \hat{p}_0] = \hat{q}_0 \hat{p}_0 - \hat{p}_0 \hat{q}_0 = i\hbar I$. Thus, if H_0 corresponds, in particular, to the one-dimensional harmonic oscillator,

$$H_0 = \frac{1}{2m_0} \hat{p}^2 + \frac{k}{2} \hat{q}^2,$$

then

$$E_n^0 = \hbar\omega \left(n + \frac{1}{2} \right), \quad n = 0, 1, 2, \dots \quad \omega = \sqrt{k/m_0}.$$

Assume for the time being that time does not enter explicitly in the Hamiltonian H (13). Then the goal in the Born–Heisenberg–Jordan (BHJ) perturbation theory is to construct a “canonical transformation” S such that

$$\hat{p} = S\hat{p}_0S^{-1}, \quad \hat{q} = S\hat{q}_0S^{-1} \quad (14)$$

and the transformed Hamiltonian is a diagonal matrix W ,

$$H(\hat{p}, \hat{q}) = SH(\hat{p}_0, \hat{q}_0)S^{-1} = W. \quad (15)$$

In other words, by means of the unitary transformation S one aims to diagonalize H . Notice that S leaves invariant the fundamental condition: $[\hat{q}, \hat{p}] = [\hat{q}_0, \hat{p}_0] = i\hbar I$. Expanding the transformation matrix S and its inverse S^{-1} in powers of ε ,

$$S = I + \varepsilon S_1 + \varepsilon^2 S_2 + \dots, \quad S^{-1} = I - \varepsilon S_1 + \varepsilon^2 (S_1^2 - S_2) + \dots, \quad (16)$$

inserting these expansions into (15) and collecting together terms of the same power in ε , the following equations are obtained [8]:

$$\begin{aligned} W_0 &= H_0 \\ W_1 &= [S_1, H_0] \\ W_r &= [S_r, H_0] + F_r(H_0, \dots, H_r; S_1, \dots, S_{r-1}) \quad r \geq 2. \end{aligned} \quad (17)$$

Here H_0, H_1, \dots are to be taken as having the arguments \hat{p}_0 and \hat{q}_0 and the matrix S is unitary: $S^{-1} = S^\dagger$. These equations can be solved in sequence, by first building the mean value $\langle F_r \rangle$ to determine the new Hamiltonian, $W_r = \langle F_r \rangle$, and then obtaining the matrix elements of S_r [8] at each order. In this setting, the mean value corresponds to take the diagonal values of F_r . Then, the S series is substituted into (14) to get the evolution of the dynamical variables:

$$\begin{aligned} \hat{q} &= \hat{q}_0 + \varepsilon \hat{q}_1 + \varepsilon^2 \hat{q}_2 + \dots = \hat{q}_0 + \varepsilon [S_1, \hat{q}_0] + \dots \\ \hat{p} &= \hat{p}_0 + \varepsilon \hat{p}_1 + \varepsilon^2 \hat{p}_2 + \dots = \hat{p}_0 + \varepsilon [S_1, \hat{p}_0] + \dots \end{aligned}$$

In this way, it is possible to obtain general formulae for the first terms in the expansion of the eigenvalues of W (i.e., the energy levels of H) even for degenerate systems [8].

What happens if time enters explicitly into H_1, H_2, \dots , but not into H_0 (as it is the case, for instance, when time-dependent external forces act on a system)? Then, according to [8, p. 336], “simple considerations show that for this case the perturbation formulae ensue from those cited earlier” (i.e., equations (17)) on replacing

$$[H_0, S_r] \quad \text{by} \quad [H_0, S_r] - i\hbar \frac{\partial S_r}{\partial t}. \quad (18)$$

Again, W_r is determined as the mean value of F_r and S_r is obtained by solving the resulting matrix equation. Moreover, Born, Heisenberg and Jordan assume that the formalism also applies when the external forces are not periodic in time (even though this assumption was incorporated into the derivation of the formulae) [8].

4 Reformulating the BHJ perturbation theory

At this point, a natural question arises: is this time-dependent perturbation theory equivalent to the standard one developed in section 2? Whereas in the autonomous case the procedure (14)-(17) reproduces the standard perturbative formalism for determining the eigenvalues of H [17], this is by no means obvious in the general case of a time dependency. Our claim here is that the BHJ formalism is indeed more general than the standard treatment of section 2, but reproduces it as a particular case.

To substantiate this claim, let us first rephrase the BHJ time-dependent perturbation theory in modern quantum mechanics language as follows. Suppose we have a quantum problem defined by the Hamiltonian (1) so that the dynamics associated with H_0 has been solved, i.e., we have computed $U_{H_0}(t, t_0)$. Then, a unitary transformation $S(t, \varepsilon)$ is sought such that the equation satisfied by the new wave function $\Psi(t) = S\psi(t)$, or equivalently, by the new evolution operator $U_W(t, t_0)$, with $\Psi(t) = U_W(t, t_0)\Psi(t_0)$,

$$i\hbar \frac{\partial}{\partial t} U_W(t, t_0) = W(t, \varepsilon) U_W(t, t_0), \quad U_W(t_0, t_0) = I \quad (19)$$

is easier to solve than (2). It is clear that both evolution operators are then related by

$$U_W(t, t_0) = S(t, \varepsilon) U_H(t, t_0) S^{-1}(t_0, \varepsilon), \quad (20)$$

whereas the corresponding Hamiltonians H and W verify [9]

$$W(t, \varepsilon) = S(t, \varepsilon) H(t, \varepsilon) S^{-1}(t, \varepsilon) + i\hbar \frac{\partial S(t, \varepsilon)}{\partial t} S^{-1}(t, \varepsilon). \quad (21)$$

In this setting the new Hamiltonian W has to be fixed and afterwards the transformation $S(t, \varepsilon)$ is determined. This can be done perturbatively by expanding W and S as power series in ε ,

$$\begin{aligned} W(t, \varepsilon) &= W_0 + \widetilde{W}, & \text{with} & \quad \widetilde{W} \equiv \varepsilon W_1(t) + \varepsilon^2 W_2(t) + \dots \\ S(t, \varepsilon) &= I + \widetilde{S}, & \text{with} & \quad \widetilde{S} \equiv \varepsilon S_1(t) + \varepsilon^2 S_2(t) + \dots, \end{aligned} \quad (22)$$

substituting series (22) into (21), or alternatively in $WS = SH + i\hbar \frac{\partial S}{\partial t}$. This results in $W_0 = H_0$ and

$$i\hbar \frac{\partial \widetilde{S}}{\partial t} + [\widetilde{S}, H_0] = \widetilde{W} - \widetilde{H} - \widetilde{S}\widetilde{H} + \widetilde{W}\widetilde{S}. \quad (23)$$

Equating terms with the same power of ε in (23) we get for $n = 1, 2, \dots$

$$i\hbar \dot{S}_n + [S_n, H_0] = W_n - V_n, \quad \text{with} \quad V_n = H_n + \sum_{j=1}^{n-1} (S_{n-j} H_j - W_j S_{n-j}). \quad (24)$$

From equations (23) and (24) we clearly see the origin of the prescription (18) for getting the perturbation formulae as given in [8].

4.1 Recovering the STDP theory

Next we analyze how the present framework allows one to reproduce the standard time-dependent perturbation theory developed in section 2. More specifically, we show that the STDP theory just corresponds to taking the unperturbed part H_0 as the new Hamiltonian. In other words, the expression (3) for the evolution operator U_H with (6) is recovered when the unitary transformation S is designed such that $W = H_0$.

To keep the treatment as simple as possible, we analyze the case $\tilde{H}(t, \varepsilon) = \varepsilon H_1(t)$ in (1) and take $t_0 = 0$. We also drop the second argument t_0 in the definition of U_H , etc. If $W = H_0$, then the solution of (19) is $U_W(t) = \exp(-itH_0/\hbar)$, so that, from (20),

$$U_H(t) = S^{-1}(t, \varepsilon) e^{-itH_0/\hbar}$$

if $S(0, \varepsilon) = I$. To determine the inverse of the transformation S , instead of working with equations (24), it is more convenient to formulate a similar equation for the series

$$S^{-1}(t, \varepsilon) = I + \varepsilon S_1^{-1}(t) + \varepsilon^2 S_1^{-2}(t) + \dots$$

This can be done as before from (21), or equivalently

$$S^{-1}W = HS^{-1} - i\hbar \frac{\partial S^{-1}}{\partial t}. \quad (25)$$

An easy calculation leads to

$$i\hbar \dot{S}_n^{-1} + [S_n^{-1}, H_0] = H_1 S_{n-1}^{-1}, \quad n = 1, 2, \dots$$

with $S_0^{-1} = I$, whose formal solution reads [9]

$$S_n^{-1}(t) = e^{-itH_0/\hbar} g_n(t) e^{itH_0/\hbar}, \quad (26)$$

with

$$g_n(t) = -\frac{i}{\hbar} \int_0^t du e^{iuH_0/\hbar} H_1(u) S_{n-1}^{-1}(u) e^{-iuH_0/\hbar}. \quad (27)$$

Thus, inserting (26) into (27) we get

$$g_n(t) = -\frac{i}{\hbar} \int_0^t H_1^I(u) g_{n-1}(u) du, \quad n = 1, 2, \dots$$

and $g_0 \equiv I$. Notice then that $g_n(t) = U_n(t, 0)$, as given by (7). In consequence,

$$\begin{aligned} U_H(t) &= \left(I + \sum_{n \geq 1} \varepsilon^n S_n^{-1}(t) \right) e^{-itH_0/\hbar} = e^{-itH_0/\hbar} \left(I + \sum_{n \geq 1} \varepsilon^n g_n(t) \right) \\ &= e^{-itH_0/\hbar} \sum_{n=0}^{\infty} \varepsilon^n U_n(t) \end{aligned}$$

and the standard formalism (3)-(7) is recovered.

4.2 Averaging

As a matter of fact, Born, Heisenberg and Jordan pursued a different course of action to deal with equations (24). Inspired by the classical treatment of perturbations in Hamiltonian classical mechanics (see e.g. [3]), they proposed instead to take as the new Hamiltonian W_n the mean value (or time average) of $V_n(t)$ at each order, and subsequently to determine the matrix $S_n(t)$ from (24) [8].

What is the “mean value” of $V_n(t)$? If the function is periodic with period T , then the answer is clear: the mean value is just

$$\langle V_n \rangle = \frac{1}{T} \int_0^T V_n(t) dt. \quad (28)$$

The mean value is also well defined for quasi periodic functions of time. We recall that a function V is said to be *quasi-periodic* with basic frequencies $\omega = (\omega_1, \dots, \omega_r)$ if $V(t) = f(\theta_1, \dots, \theta_r)$, where f is 2π -periodic with respect to $\theta_1, \dots, \theta_r$ and $\theta_j = \omega_j t$ for $j = 1, \dots, r$. The mean value (or *average*) of V is then defined as

$$\langle V \rangle \equiv \lim_{T \rightarrow \infty} \frac{1}{T} \int_a^{a+T} V(t) dt, \quad (29)$$

and this limit does not depend on the particular value of a considered [10]. If V is periodic with period T , (29) reproduces the expression (28).

A formal solution of equation (24), $n \geq 1$, is given by

$$S_n(t) = e^{-itH_0/\hbar} \left(S_n(0) - \frac{i}{\hbar} \int_0^t du e^{iuH_0/\hbar} (W_n - V_n(u)) e^{-iuH_0/\hbar} \right) e^{itH_0/\hbar}, \quad (30)$$

but, as shown in [2], $S_n(t)$ does not contain secular terms (i.e., it is quasi periodic with the same basic frequencies as $V_n(t)$) if one takes

$$\begin{aligned} W_n &= \langle V_n \rangle \\ S_n(t) &= e^{-itH_0/\hbar} G_n(t) e^{itH_0/\hbar}, \end{aligned} \quad (31)$$

where $G_n(t)$ denotes the antiderivative

$$G_n(t) = -\frac{i}{\hbar} \int dt e^{itH_0/\hbar} (\langle V_n \rangle - V_n(t)) e^{-itH_0/\hbar}.$$

Notice that with prescription (31), $S_n(t)$ is compatible with (30) and moreover $S_n(0) = G_n(0)$. In the matrix case, this solution is well defined provided that $V_n(t)$ is quasi-periodic with basic frequencies $\omega = (\omega_1, \dots, \omega_r)$, the distinct eigenvalues of H_0 are E_1^0, \dots, E_s^0 and the quantities $|\omega_{\ell m} + (k, \omega)| > c|k|^{-\gamma}$, $1 \leq \ell, m \leq s$, for all vectors k with integer components. Here c, γ are positive constant, $\omega_{\ell m} = (E_\ell^0 - E_m^0)/\hbar$, $|k| = |k_1| + \dots + |k_r|$ and $(k, \omega) = k_1\omega_1 + \dots + k_r\omega_r$. This inequality prevents of course the presence of small denominators in the expression of $S_n(t)$ [2].

When the procedure is carried out for m steps, the series are truncated at terms of order ε^m , and the resulting transformation defined by the corresponding S is considered, in the new picture the evolution is given by

$$U_W(t) = e^{-itW^{[m]}(\varepsilon)/\hbar}, \quad \text{with} \quad W^{[m]}(\varepsilon) = H_0 + \sum_{j=1}^m \varepsilon^j W_j,$$

and $W_j = \langle V_j \rangle$. Finally, the approximation reads, according to (20),

$$U_H(t) = \left(I + \sum_{j=1}^m \varepsilon^j S_j^{-1}(t) \right) e^{-itW^{[m]}(\varepsilon)/\hbar} \left(I + \sum_{j=1}^m \varepsilon^j S_j(0) \right). \quad (32)$$

Should the series $S(t, \varepsilon)$ for the change of variables converge, the procedure described above would allow us to solve the original perturbed problem.

Notice that with this procedure we have to compute the terms $S_j^{-1}(t)$ for the inverse truncated series to get the actual approximation (32) from the corresponding $S_k(t)$ and this typically involves several matrix products. Thus, in particular, S_1^{-1} and S_2^{-1} are given in (16), whereas $S_3^{-1} = -S_1^3 + S_1 S_2 + S_2^2 - S_3$. It might be then more convenient to reformulate the problem in terms of $S_n^{-1}(t)$ and obtain a solution verifying $S_n(0) = 0$. This can be done as in section 4.1, first obtaining from (25) the equation to be satisfied by $S_n^{-1}(t)$, and then choosing appropriately the solution. In any event, there are no secular terms in the approximation (32) thus obtained. By applying averaging, all these terms are removed from $S(t, \varepsilon)$, just as in the classical perturbation theory of Hamiltonian systems. Still, when the series $S(t, \varepsilon)$ is truncated, the resulting operator is no longer unitary.

Example (cont. I). We next apply the BHJ formalism with averaging to the Hamiltonian (8). In this case, clearly $\langle H_1 \rangle = 0$, so that

$$W_1 = 0 \quad \text{and} \quad S_1^{-1}(t) = \frac{i}{\hbar(\omega - \omega_0)} (\cos \omega t \sigma_2 - \sin \omega t \sigma_1) \quad (33)$$

whereas for the second order one has

$$W_2 = \frac{1}{\hbar(\omega - \omega_0)} \sigma_3, \quad S_2^{-1}(t) = 0, \quad (34)$$

and thus, in this approximation,

$$U_H(t) \approx (I + \varepsilon S_1^{-1}(t)) e^{-it(H_0 + \varepsilon^2 W_2)/\hbar} (I + \varepsilon S_1(0)) \quad (35)$$

with $S_1(0) = -S_1^{-1}(0) = \frac{-i}{\hbar(\omega - \omega_0)} \sigma_2$.

5 Unitary perturbation theory with averaging

At this point it is worth remarking that before the publication of [8], Heisenberg proposed a slightly different perturbation formalism for the time independent case, in terms of a skew-Hermitian operator \mathfrak{S} such that

$$S(t, \varepsilon) = e^{\mathfrak{S}(t, \varepsilon)}, \quad \text{with} \quad \mathfrak{S} = \varepsilon \mathfrak{S}_1 + \varepsilon^2 \mathfrak{S}_2 + \dots \quad (36)$$

Then, equations similar to (17) are obtained, but now involving iterated commutators of the terms \mathfrak{S}_i [22, p. 47]. Although this approach was considered more involved than the one based on a direct power expansion of S (and even wrong at first sight [22, p. 49]), since the corresponding formulae differ from (17) already at order two), we could say that, in a sense, it is qualitatively superior. For the unitary character of the transformation is guaranteed even when the series \mathfrak{S} is truncated, in contrast with the approach (13)-(18), where unitary is only preserved up to the order of the approximation.

When this approach is applied to the explicitly time dependent case, equation (21) can be rewritten in terms of the exponent \mathfrak{S} as

$$W = e^{\mathfrak{S}} H e^{-\mathfrak{S}} + i\hbar \text{dexp}_{\mathfrak{S}}(\dot{\mathfrak{S}}), \quad (37)$$

where the symbol $\text{dexp}_{\mathfrak{S}}(C)$ stands for the (everywhere convergent) power series

$$\text{dexp}_{\mathfrak{S}}(C) = \sum_{k=0}^{\infty} \frac{1}{(k+1)!} \text{ad}_{\mathfrak{S}}^k(C) \equiv \frac{\exp(\text{ad}_{\mathfrak{S}}) - I}{\text{ad}_{\mathfrak{S}}}(C)$$

which naturally arises when differentiating the exponential of a non constant matrix [6]. Here $\text{ad}_{\mathfrak{S}}(C)$ is a short-hand notation for the (iterated) commutator

$$\text{ad}_{\mathfrak{S}}(C) = [\mathfrak{S}, C], \quad \text{ad}_{\mathfrak{S}}^j(C) = [\mathfrak{S}, \text{ad}_{\mathfrak{S}}^{j-1}(C)], \quad \text{ad}_{\mathfrak{S}}^0(C) = C.$$

Inverting the operator $\text{dexp}_{\mathfrak{S}}(C)$, one has

$$i\hbar \dot{\mathfrak{S}} = \text{dexp}_{\mathfrak{S}}^{-1} \left(W - e^{\mathfrak{S}} H e^{-\mathfrak{S}} \right), \quad (38)$$

where

$$\text{dexp}_{\mathfrak{S}}^{-1}(C) = \frac{\text{ad}_{\mathfrak{S}}}{\exp(\text{ad}_{\mathfrak{S}}) - I}(C) = \sum_{k=0}^{\infty} \frac{B_k}{k!} \text{ad}_{\mathfrak{S}}^k(C)$$

and B_k are the Bernoulli numbers [18]. Inserting the corresponding series for H , W and \mathfrak{S} into (38) and collecting terms of the same power in ε one arrives at $W_0 = H_0$ and

$$i\hbar \dot{\mathfrak{S}}_n + [\mathfrak{S}_n, H_0] = W_n - F_n, \quad n = 1, 2, \dots, \quad (39)$$

with

$$F_1 = H_1, \quad F_2 = H_2 + \frac{1}{2}[\mathfrak{S}_1, W_1 + H_1] \quad (40)$$

and in general $F_n = H_n + \tilde{F}_n$, where \tilde{F}_n is a linear combination of nested commutators involving $\mathfrak{S}_1, \dots, \mathfrak{S}_{n-1}, H_1, \dots, H_{n-1}, W_1, \dots, W_{n-1}$ whose expression can be determined using the same techniques as in e.g. [5]. By following a similar approach as in section 4.2, we get the solution of (39) such that $\mathfrak{S}_n(t)$ is free of secular terms as

$$\begin{aligned} W_n &= \langle F_n \rangle \\ \mathfrak{S}_n(t) &= e^{-itH_0/\hbar} J_n(t) e^{itH_0/\hbar} \end{aligned} \quad (41)$$

with

$$J_n(t) = -\frac{i}{\hbar} \int dt e^{itH_0/\hbar} (\langle F_n \rangle - F_n(t)) e^{-itH_0/\hbar}.$$

Finally, the approximation is constructed as

$$U_H(t) \approx e^{-\mathfrak{S}^{[m]}(t,\varepsilon)} e^{-itW^{[m]}(\varepsilon)/\hbar} e^{\mathfrak{S}^{[m]}(0,\varepsilon)}, \quad (42)$$

where

$$\mathfrak{S}^{[m]}(t,\varepsilon) = \sum_{n=1}^m \varepsilon^n \mathfrak{S}_n(t), \quad W^{[m]}(\varepsilon) = \sum_{n=0}^m \varepsilon^n W_n.$$

We thus get formally the same expressions as in section 4.2, but there are important differences, though. First, equations (39) for determining \mathfrak{S}_n are formulated only in terms of skew-Hermitian operators and commutators. If, as it is often the case, the original problem is defined in a Lie algebra of finite dimension, one can use the constants of structure to compute the commutators appearing in F_n and thus render a more efficient algorithm. Second, the procedure renders unitary approximations by construction even if the series $\mathfrak{S}(t,\varepsilon)$ is truncated after m terms. Third, computing the inverse transformation $S^{-1}(t,\varepsilon)$ is trivial once \mathfrak{S} has been obtained: we only have to change the sign of \mathfrak{S} and evaluate $S^{-1}(t,\varepsilon) = \exp(-\mathfrak{S}(t,\varepsilon))$ (or the corresponding truncation of \mathfrak{S}).

A variation of this procedure, considered in the context of quantum averaging [21] and the design of unitary transformations [9] is the following: introduce a skew-Hermitian operator $L(t,\varepsilon)$ such that $S(t,\varepsilon)$ is the formal solution of the operator differential equation

$$\frac{\partial}{\partial \varepsilon} S(t,\varepsilon) = L(t,\varepsilon) S(t,\varepsilon), \quad S(t,0) = I \quad (43)$$

with $L(t,\varepsilon) = L_1(t) + \varepsilon L_2(t) + \dots$. In other words, the required unitary transformation $S(t,\varepsilon)$ is obtained by shifting a “time” ε along the trajectories of the differential equation (43). The operator L can be seen as the generator of the transformation. This is in complete analogy with the procedure proposed by Deprit in classical mechanics [15], and leads to some computational simplifications in the procedure.

Example (cont. II). For the two-level system (8) the first orders obtained by applying this procedure are

$$\begin{aligned} W_1 = 0, \quad \mathfrak{S}_1(t) &= \frac{-i}{\hbar(\omega - \omega_0)} (\cos \omega t \sigma_2 - \sin \omega t \sigma_1) \\ W_2 = \frac{-1}{\hbar(\omega - \omega_0)} \sigma_3, \quad \mathfrak{S}_2(t) &= 0 \end{aligned} \quad (44)$$

and so the approximation up to ε^2 reads

$$U_H(t) \approx \exp(-\varepsilon \mathfrak{S}_1(t)) \exp((-i/\hbar)(H_0 + \varepsilon^2 W_2)) \exp(\varepsilon \mathfrak{S}_1(0)) \quad (45)$$

Notice that, by expanding the exponentials $\exp(-\varepsilon \mathfrak{S}_1(t))$ and $\exp(\varepsilon \mathfrak{S}_1(0))$ up to order ε , we recover the result (35) obtained with the BHJ formalism.

In Figure 1 we represent the exact result for the transition probability between states (solid line) together with the second order approximations achieved by the different procedures: approximations (11) (dot-dashed curve), (35) (dashed line) and (45) (dotted curved) when $\omega_0 = 1$, $\omega = 3$, $\varepsilon = 0.5$ (and $\hbar = 1$). Even when there are no secular terms in (12), we see that the result furnished by the STDP theory is clearly less accurate than the approximations rendered by both the BHJ and the unitary schemes.

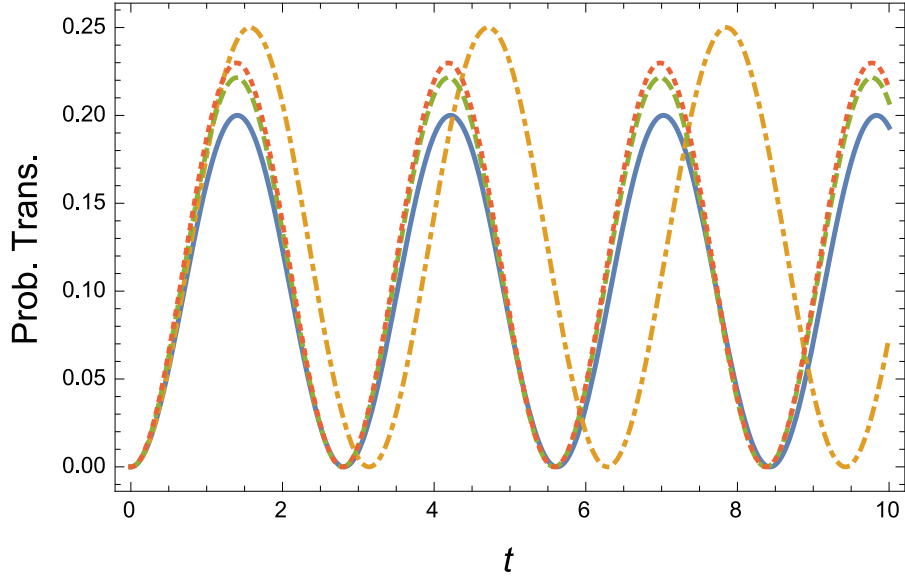


Figure 1: Transition probability when $\omega_0 = 1$, $\omega = 3$, $\varepsilon = 0.5$ obtained by the different perturbation schemes up to order ε^2 analyzed in this paper together with the exact result (9) (solid line): STDP theory (dot-dashed curve), BHJ theory (dashed line) and the unitary version (dotted curved).

To illustrate how the different approximations behave for longer times, in Figure 2 we collect the errors in the transition probability obtained with the STDP theory (solid line), the BHJ formalism with averaging (dashed line) and the unitary version (dotted line), for the same value of the parameters but now up to order ε^6 after 100 periods of the exact result $T = 2\pi/\omega'$. Whereas BHJ and its unitary version provide almost undistinguishable results (although the former is not unitary), those achieved the standard perturbation theory are completely useless for large times.

To clarify the nature of the error in the BHJ and its unitary variant, in Figure 3 we depict the transition probability in the interval $t \in [280, 283]$, where now the solid line corresponds to the exact result (the result given by STDP theory is by large out of the scale of the graph). Only the result obtained by BHJ theory is depicted, since the curve corresponding to the the unitary version is almost identical. Notice that the amplitude of the approximations is quite similar to the exact result, and the error is mainly due to a shift in the graph of the function describing the transition probability.

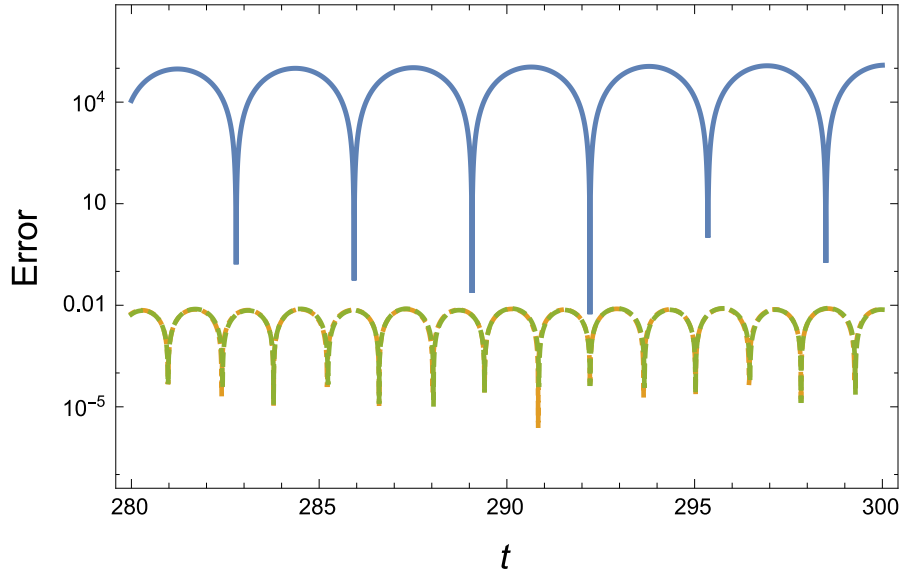


Figure 2: Error in the transition probability when $\omega_0 = 1$, $\omega = 3$, $\varepsilon = 0.5$ after 100 periods of the exact result furnished by the STDP theory (solid line), the BHJ formalism (dashed line) and the unitary version (dotted curve) up to ε^6 .

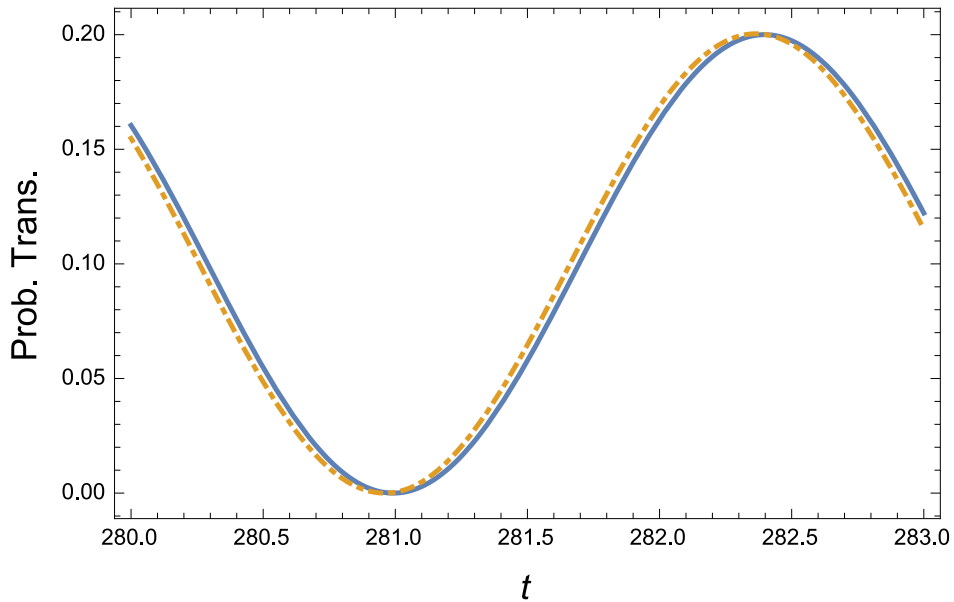


Figure 3: Transition probability when $\omega_0 = 1$, $\omega = 3$, $\varepsilon = 0.5$: exact result (solid line) and approximation furnished by the BHJ formalism (dashed-dotted line) up to ε^6 (the result achieved by the unitary version is almost identical to BHJ; for this reason it is not collected here).

6 Discussion

In this paper we have reformulated in modern quantum mechanical language the original time dependent theory developed by Born, Heisenberg and Jordan

in matrix mechanics. The formalism is based on designing a unitary transformation S such that in the new representation the dynamics associated with the corresponding Hamiltonian W is easier to solve than the original problem. In the case of autonomous problems, W is taken as diagonal. When time enters explicitly, the new Hamiltonian is chosen as constant. Both S and W are constructed as power series in the perturbation parameter. We have shown that the choice $W = H_0$ leads directly to the STDP theory, whereas the application of averaging eliminates secular terms and allows one to construct approximations that are reasonably accurate for longer time intervals. This is so although the approximations are not unitary when the series are truncated. Nevertheless, by constructing S as the exponential of a anti-Hermitian operator \mathfrak{S} , as Heisenberg first proposed, the theory can be rendered unitary even after truncation.

Here we have proceeded on a purely formal level, and several important issues should be further studied, in particular the convergence of the procedure and the conditions to be satisfied by the operators so that the formal solution (30) is well defined in a functional analytic framework. On the other hand, when only matrices are involved, this formalism can be seen as a particular case of application of high order averaging to linear systems [3].

Several studies exist in the literature where quantum mechanical analogues of classical Hamiltonian perturbation methods have been proposed, involving averaging techniques and even several unitary transformations [11, 19, 20, 21]. These transformations can be designed to be unitary by applying an analogy of the Deprit technique in classical mechanics [9]. The novelty of the treatment exposed here is that, in one way or another, all these contributions can be obtained from the original Born–Heisenberg–Jordan treatment, and that this formalism constitutes indeed an effective alternative to the usual time dependent perturbation theory when dealing with practical problems in quantum mechanics.

Acknowledgments

This work has been partially supported by Ministerio de Economía y Competitividad (Spain) through the coordinated project MTM2013-46553-C3.

References

- [1] M.K. Ali, The quantum normal form and its equivalents. *J. Math. Phys.* **26** (1985), 2565–2572.
- [2] A. Arnal, F. Casas, and C. Chiralt, A perturbative algorithm for quasi-periodic linear systems close to constant coefficients. (Submitted).
- [3] V.I. Arnold, V. Kozlov, and A.I. Neishtadt, *Mathematical Aspects of Classical and Celestial Mechanics*, 3rd edn. Springer (Berlin), 2006.
- [4] G. Birtwistle, *The New Quantum Mechanics*. Cambridge University Press (Cambridge), 1928.

- [5] S. Blanes and F. Casas, On the convergence and optimization of the Baker–Campbell–Hausdorff formula. *Lin. Alg. Appl.* **378** (2004), 135-158.
- [6] S. Blanes, F. Casas, J.A. Oteo, and J. Ros, The Magnus expansion and some of its applications. *Phys. Rep.* **470** (2009), 151–238.
- [7] M. Born and P. Jordan, Zur Quantenmechanik. *Z. Phys.* **34** (1925), 858-888. Also in [22], pages 277-306.
- [8] M. Born, W. Heisenberg, and P. Jordan, Zur Quantenmechanik II. *Z. Phys.* **35** (1926), 557-615. Also in [22], pages 321-385.
- [9] F. Casas, J.A. Oteo, and J. Ros, Unitary transformations depending on a small parameter. *Proc. R. Soc. A* **468** (2012), 685-700.
- [10] C. Corduneanu, *Almost periodic functions*. John Wiley (New York), 1968.
- [11] D. Daems, S. Guérin, H.R. Jauslin, A. Keller, and O. Atabek, Optimized time-dependent perturbation theory for pulse-driven quantum dynamics in atomic or molecular systems. *Phys. Rev. A* **68** (2003), 051402(R) (4 pages).
- [12] B. Eckhardt, Birkhoff–Gustavson normal form in classical and quantum mechanics. *J. Phys. A: Math. Gen.* **19** (1986), 2961–2972.
- [13] W. Heisenberg, Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen. *Z. Phys.* **33** (1925), 879-893. Also in [22], pages 261-276.
- [14] M. Kummer, How to avoid “secular” terms in classical and quantum mechanics. *Il Nuovo Cimento* **B1** (1971), 123–148.
- [15] A.J. Lichtenberg and M.A. Leiberman, *Regular and Stochastic Motion*. Springer-Verlag (New York), 1983.
- [16] J. Mehra and H. Rechenberg, *The Historical Development of Quantum Theory. Vol. 3: The Formulation of Matrix Mechanics and Its Modifications*. Springer-Verlag (New York), 1982.
- [17] A. Messiah, *Quantum Mechanics*. Dover (New York), 1999.
- [18] F.W.J. Over, D.W. Looser, R.F. Boisvert, and C.W. Clark, *NIST Handbook of Mathematical Functions*. Cambridge University Press (Cambridge), 2010.
- [19] W. Scherer, Quantum averaging: I. Poincaré–von Zeipel is Rayleigh–Schrödinger. *J. Phys. A: Math. Gen.* **27** (1994), 8231–8246.
- [20] W. Scherer, New perturbation algorithms for time-dependent quantum systems. *Phys. Lett. A* **233** (1997), 1-6.
- [21] W. Scherer, Quantum averaging: III. time-dependent systems. *J. Math. Phys.* **39** (1998), 2597–2634.

- [22] B.L. Van der Waerden (Ed.), *Sources of Quantum Mechanics*. Dover (New York), 1968.
- [23] T.-Y. Wu, *Quantum Mechanics*. World Scientific (Singapore), 1986.