

# Unitary transformations depending on a small parameter

BY F. CASAS<sup>1</sup>, J. A. OTEO<sup>2</sup> AND J. ROS<sup>2,3,\*</sup>

<sup>1</sup>*Departament de Matemàtiques and IMAC, Universitat Jaume I, E-12071 Castellón, Spain*

<sup>2</sup>*Departament de Física Teòrica, Universitat de València, 46100-Burjassot, València, Spain*

<sup>3</sup>*Instituto de Física Corpuscular (IFIC), Universitat de València-CSIC, Apartado 22085, 46071-Paterna, València, Spain*

We formulate a unitary perturbation theory for quantum mechanics inspired by the Lie-Deprit formulation of canonical transformations. The original Hamiltonian is converted into a solvable one by a transformation obtained through a Magnus expansion. This ensures unitarity at every order in a small parameter. A comparison with the standard perturbation theory is provided. We work out the scheme up to order ten with some simple examples.

**Keywords:** perturbation theory; quantum mechanics; unitary transformation; Magnus expansion

## 1. Introduction

Historically, classical mechanics played a crucial role in the early formulation of quantum mechanics. Afterwards, the new ideas led to a quantum formalism that loosened contact with the classical language. But nevertheless, every now and then, mutual relationships between them have been profitably considered.

Fixing our attention on the treatment of the effect of a perturbation on a solvable physical system, matrix mechanics adapted, with daring modifications and reinterpretations, the classical lore of canonical transformations, generating functions, action and angle variables and the like. This is apparent, e.g. in the famous paper by Born *et al.* (1926). Nowadays, both time-independent and time-dependent quantum perturbation theories are presented in the Hilbert space formalism and make no mention of their classical counterparts.

Suppose one is interested in a system with Hamiltonian  $H(t, \epsilon) = H_0 + H'(t, \epsilon)$ , where  $H_0$  is time-independent and  $H'(t, \epsilon)$  is a time-dependent perturbation, depending on a small parameter  $\epsilon$  that controls the intensity of the perturbation in such a way that  $H'(t, \epsilon = 0) = 0$ . The aim is then to determine the quantum time-evolution operator  $U(t, t_0)$ , which takes the state of the system from time  $t_0$  to time  $t$ . One of its most salient features is its unitary

\*Author for correspondence ([rosj@uv.es](mailto:rosj@uv.es)).

character. The evolution operator satisfies the Schrödinger equation which, in the conventional approach, is solved by series expansion in  $\epsilon$ . After truncation at a given order, it allows approximation of the different observable magnitudes of physical interest. The main drawback of the method is that any finite order approximation of  $U(t, t_0)$  ceases to be unitary.

By contrast, the procedure followed in the founding years of quantum mechanics and directly inherited from the classical perturbation theory was more geometric in character. The power series expansions were also the basic tool, but the goal was to construct a geometric (unitary) transformation that renders solvable the Hamiltonian  $H(t, \epsilon)$  (e.g. Birtwistle 1928). The reader acquainted with the Hamilton–Jacobi equation will recognize the idea of finding a canonical transformation (namely, a change of phase space coordinates) that transforms a Hamiltonian  $H$  into a Hamiltonian  $K$  easily solvable in the new coordinates. The problem is then translated to the search of the corresponding generating function.

The original canonical transformations theory used in classical mechanics relies on the effective use of generating functions expressed in terms of mixed, old and new, variables (e.g. Arnold *et al.* 2006). It means that, once the equations of motion have been integrated, one has to solve implicit functional equations so as to express the results in terms of either the old or the new variables. In general, this is difficult to do.

This classical canonical theory, however, fitted in the late 1960s into a Lie-algebraic setting, mainly in connection with problems of celestial mechanics. In that respect, the paper (Deprit 1969) *Canonical transformations depending on a small parameter* played a seminal role by introducing what it is known nowadays as the *Lie-Deprit perturbation method*. Essentially, a canonical transformation is generated by a new ‘Hamiltonian’  $W$  in such a way that the shift by ‘time’  $\epsilon$  along the trajectories of  $W$  produces the required transformation. In this way, no mixed variables appear in the formalism. The algorithm can be readily implemented in a computer, thus allowing carrying out of perturbative analysis up to high order.

Since then, techniques and procedures of classical perturbation theory have been used in a quantum mechanical context from time to time (e.g. Kummer 1971; Eckhardt 1986; Scherer 1994; Daems *et al.* 2004). In fact, this is quite natural, given the existing parallelism between perturbation theory in classical and quantum systems when both are formulated in terms of geometric transformations (canonical maps and unitary transformations, respectively). The common formal setting is indeed provided by the language of Lie algebras. From this perspective, the perturbation problem defined by  $H_0 + H'(t, \epsilon)$  can be solved by changing through a (canonical or unitary) transformation to a Hamiltonian  $K$  whose dynamics is easier to solve, at least up to a certain order in the  $\epsilon$  parameter. When the perturbation expansion converges, the original problem is completely solved. Even when the convergence problem is not settled, the procedure very often provides very accurate results.

The purpose of the present paper consists in adapting and developing an implementation of the Lie-Deprit method that gives results amenable to analyse time-dependent problems in quantum mechanics. For this reason, we have adapted the title of the 1969 Deprit paper. The procedure leads to operator linear differential equations, which we solve by use of the Magnus expansion. This will ensure unitarity at any order of truncation in  $\epsilon$ . This is an important feature of the algorithm proposed here. Thus, it is guaranteed

that the constructed approximation shares several important qualitative features with the exact solution. In addition, in this formalism, several approximation schemes are possible, better suited to the mathematical nature of the problem at hand. We will explain how to take advantage of this. The connection with the standard time-dependent perturbation theory will also be provided. Here, we will proceed at a purely formal level, keeping the technical requisites to a minimum.

The plan of the paper is the following: in §2, we briefly summarize the standard time-dependent perturbation theory in quantum mechanics in the interaction picture defined by  $H_0$ . In §3, we develop the new unitary formalism, paying special attention to the construction of the transformation and the derivation of the basic equation. The procedure is highly flexible in the sense that several possible choices of the new Hamiltonian  $K$  are possible. Also the time-independent case fits naturally into the algorithm. These issues are analysed in §4, whereas §5 is devoted to a pair of examples to illustrate the main features of the scheme. Finally, §6 contains an additional discussion and further comparisons with the so-called quantum averaging approach.

## 2. Time-dependent perturbation theory in quantum mechanics

Let us consider the dynamics governed by the Hamiltonian

$$H(t, \epsilon) = H_0 + H'(t, \epsilon) \equiv H_0 + \sum_{n=1}^{\infty} \epsilon^n H_n(t), \quad (2.1)$$

where we have assumed that  $H'(t, \epsilon)$  is analytic in  $\epsilon$ . In most situations, the power series expansion of  $H'(t, \epsilon)$  collapses to the first term:  $H'(t, \epsilon) = \epsilon H_1(t)$ .

The corresponding time-evolution operator  $U(t, t_0)$  is unitary, i.e.  $U(t, t_0) U^\dagger(t, t_0) = U^\dagger(t, t_0) U(t, t_0) = I$ , where the dagger means adjoint. This property guarantees, in physical terms, probability conservation. Although for simplicity we have not shown it explicitly,  $U(t, t_0)$  depends also on  $\epsilon$ . In turn, it obeys the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} U(t, t_0) = H(t) U(t, t_0), \quad U(t_0, t_0) = I. \quad (2.2)$$

When the Hamiltonian is independent of time, i.e.  $H(t) = H_0$ , the solution of (2.2) reads  $U(t, t_0) = \exp(-i(t - t_0)H_0/\hbar)$ . Otherwise, the solution is by no means so simple, unless  $\int H(t)dt$  and  $H(t)$  do commute.

For the Hamiltonian (2.1), the problem can be approximately solved by the conventional time-dependent perturbation theory, yielding the solution of (2.2) in power series of  $\epsilon$ . In this setting, it is very important to isolate the piece of the problem to be treated approximately and allow the solvable part of the Hamiltonian to be exactly integrated. A way to deal with this question consists in transforming the problem into the interaction picture, i.e. we factorize

$$U(t, t_0) = \exp\left(\frac{-i(t - t_0)H_0}{\hbar}\right) U_I(t, t_0), \quad (2.3)$$

where the unknown operator  $U_I$  obeys the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} U_I(t, t_0) = H_I(t, \epsilon) U_I(t, t_0), \quad U_I(t_0, t_0) = I, \quad (2.4)$$

with the new Hamiltonian given by

$$H_I(t, \epsilon) = \exp\left(\frac{-i(t-t_0)H_0}{\hbar}\right) H'(t, \epsilon) \exp\left(\frac{i(t-t_0)H_0}{\hbar}\right) \equiv \sum_{n=1}^{\infty} \epsilon^n H_n^I(t). \quad (2.5)$$

In this way, the interaction picture defined in (2.3) provides a unitary transformation that allows the mathematical integration of the solvable piece  $H_0$  in  $H(t, \epsilon)$  and focuses the approximations just on  $H'(t, \epsilon)$ . Explicitly, one deals with the expansion

$$U_I(t, t_0) = \sum_{n=0}^{\infty} \epsilon^n U_n(t, t_0), \quad (2.6)$$

where  $U_n(t, t_0)$  stands for the contribution of order  $n$ .

Substitution of (2.6) in (2.4) and then equating terms of the same power  $\epsilon^n$  yields the successive terms  $U_n$ . In particular, the first two read

$$U_0(t, t_0) = I, \quad U_1(t, t_0) = \int_{t_0}^t ds H_1^I(s). \quad (2.7)$$

Unfortunately, neither the approximant  $U_I \simeq I + U_1$  nor any of higher finite order are unitary. This, of course, has undesirable consequences in practice. For instance, the computed transition probability between different quantum states may exceed unity. Only when the infinite series is summed up, the unitary character is restored.

### 3. Unitary time-dependent perturbation theory

#### (a) The formalism

We will start by rephrasing, in modern quantum mechanics language, the main idea of perturbation theory such as it was first formulated in matrix mechanics. It will help us to introduce the central issue of the present article.

Suppose we have a *hard* time-independent quantum problem defined by the Hamiltonian  $H(\epsilon) = H_0 + \epsilon H_1 + \epsilon^2 H_2 + \dots$ , and assume the mathematical problem defined by  $H_0$  has been solved. This is equivalent to finding a basis of the Hilbert space such that  $H_0$  is diagonal. Then the idea is to look for a unitary transformation  $T(\epsilon)$  (or ‘canonical transformation’ in the language of matrix mechanics) such that the new Hamiltonian  $K = T^\dagger H T$  is diagonal. The transformation  $T$  can be obtained recursively as a power series in the perturbation parameter  $\epsilon$ . Notice that the starting point of the method is rather geometric in nature, in clear contrast with time-dependent perturbation theory that is algebraic from the very beginning. Of course, an appropriate choice of  $K$  is essential.

In the following, we consider the more general, time-dependent problem defined by (2.1). Assume that the dynamics corresponding to  $H_0$  has been solved, i.e. we have determined

$$U_{H_0}(t, t_0) = \exp\left(\frac{-i(t - t_0)H_0}{\hbar}\right).$$

Let us now define

$$U_K(t, t_0) = T^\dagger(t, \epsilon) U(t, t_0) T(t_0, \epsilon), \quad (3.1)$$

where  $T(t, \epsilon)$  is a unitary transformation we want to construct in order to solve our problem.

The Schrödinger equation for  $U_K(t, t_0)$  reads

$$i\hbar \frac{\partial}{\partial t} U_K(t, t_0) = K(t, \epsilon) U_K(t, t_0), \quad U_K(t_0, t_0) = I, \quad (3.2)$$

where the new Hamiltonian is

$$K(t, \epsilon) = T^\dagger(t, \epsilon) H(t, \epsilon) T(t, \epsilon) + i\hbar \frac{\partial T^\dagger(t, \epsilon)}{\partial t} T(t, \epsilon). \quad (3.3)$$

Our goal is to determine  $T(t, \epsilon)$  in such a way that (3.2) is easier to solve than the original equation (2.2). Notice that the factorization (3.1) may be considered as a generalization of (2.3). Very often we will take  $T(t_0, \epsilon) = I$  in (3.1) and  $t_0 = 0$ .

In perturbation theory, the perturbed and the non-perturbed Hamiltonians are always very close to each other in some mathematical sense. It is right that what makes meaningful perturbative computations. It is then natural to determine  $T(t, \epsilon)$  as a unitary near-identity transformation

$$T^\dagger(t, \epsilon) = T^{-1}(t, \epsilon), \quad T(t, \epsilon) = I + \mathcal{O}(\epsilon).$$

Two different ways of guaranteeing these features are the following:

- Introduce a skew-Hermitian operator  $S(t, \epsilon) = \mathcal{O}(\epsilon)$  such that  $T(t, \epsilon) = \exp(S(t, \epsilon))$ .
- Alternatively, introduce a skew-Hermitian operator  $L(t, \epsilon)$  such that  $T(t, \epsilon)$  is the solution of the operator differential equation

$$\frac{\partial}{\partial \epsilon} T(t, \epsilon) = -T(t, \epsilon) L(t, \epsilon), \quad T(t, 0) = I. \quad (3.4)$$

This second prescription is precisely the approach proposed by Deprit in classical mechanics, and the one we follow here. Somehow, the perturbed Hamiltonian may be regarded as the *evolution*, with respect to the parameter  $\epsilon$ , of the non-perturbed Hamiltonian (which may be thought of as an *initial condition*). Deprit's idea consists in materializing such an abstract  $\epsilon$ -evolution in a manner similar to the usual  $t$ -evolution, since this way the unitary character of  $T$  is ensured. To this end, an evolution law similar to the one in (2.2) is introduced.

In (3.4),  $L(t, \epsilon)$  is the (unknown) generator of the transformation  $T$ , namely, of the  $\epsilon$ -evolution. Equivalently, one has

$$\frac{\partial}{\partial \epsilon} T^\dagger(t, \epsilon) = L(t, \epsilon) T^\dagger(t, \epsilon), \quad T^\dagger(t, 0) = I, \quad (3.5)$$

The assessment of the situation is the following. Initially, we are given a time-dependent Hamiltonian  $H(t, \epsilon) = H_0 + H'(t, \epsilon)$ , and the problem is to solve the Schrödinger equation. However, we are only able to solve the piece  $H_0$ . Then we introduce the idea of an  $\epsilon$ -evolution in such a way that a (unknown) continuous transformation  $T(t, \epsilon)$  takes  $H(t, \epsilon)$  to a new Hamiltonian  $K(t, \epsilon)$  to be determined in such a way that the corresponding Schrödinger equation (3.2) can be solved. The inverse transformation  $T^\dagger(t, \epsilon)$  may be used then to transform the solution of  $K(t, \epsilon)$  into the one for  $H(t, \epsilon)$ . However, by introducing the law of the  $\epsilon$ -evolution, we have added a further unknown: the generator  $L(t, \epsilon)$ . Thus, the situation may now appear worse than at the beginning: neither  $T$  nor  $L$  are known!

This apparently adverse scheme will provide, however, an excellent starting point to deal with power series expansions and preserve the unitary character even if the series are truncated.

We first notice that  $L$  and  $T^\dagger$  are related by the linear differential equation (3.5). Thus, once  $L$  has been determined, we can obtain  $T$  and  $T^\dagger$  by formally applying the Magnus expansion (Magnus 1954) to (3.5). As is well known, his proposal is to solve a generic operator linear equation  $dY(t)/dt = A(t)Y(t)$  with  $Y(0) = I$  in the form

$$Y(t) = \exp \Omega(t), \quad \Omega(0) = O.$$

The exponent  $\Omega(t)$  is given by an infinite series  $\Omega(t) = \sum_{m=1}^{\infty} \Omega_m(t)$  of terms that are linear combinations of integrals and nested commutators involving  $m$ -fold products of the operator  $A$  at different times (Blanes *et al.* 2009).

Although explicit formulae for  $\Omega_m$  exist, it is sometimes more convenient to generate them by a recursive procedure. The one proposed in Klarsfeld & Oteo (1989) is particularly well suited from a computational point of view (Blanes *et al.* 2009):

$$S_m^{(1)} = [\Omega_{m-1}, A], \quad S_m^{(j)} = \sum_{n=1}^{m-j} [\Omega_n, S_{m-n}^{(j-1)}], \quad 2 \leq j \leq m-1$$

and

$$\Omega_1 = \int_0^t A(t_1) dt_1, \quad \Omega_m = \sum_{j=1}^{m-1} \frac{B_j}{j!} \int_0^t S_m^{(j)}(t_1) dt_1, \quad m \geq 2, \quad (3.6)$$

where  $B_j$  stand for Bernoulli numbers and  $[\cdot, \cdot]$  represents the commutator,  $[A, B] \equiv AB - BA$ . Notice that this approach can be directly applied to obtain approximate formal solutions of equation (3.5), so that we can write

$$T^\dagger(t, \epsilon) = T^{-1}(t, \epsilon) = \exp \Omega(t, \epsilon), \quad (3.7)$$

where  $\Omega$  is a skew-Hermitian operator. Thus, the two above enumerated alternatives to render  $T$  unitary are, in fact, equivalent.

Next, we derive equation (3.3) with respect to  $\epsilon$ ,

$$\frac{\partial K}{\partial \epsilon} = [L, K] + T^\dagger \frac{\partial H}{\partial \epsilon} T + i\hbar \frac{\partial L}{\partial t}. \quad (3.8)$$

Since this expression involves both the operators  $T$  and  $L$ , it is convenient to insert (3.7) here. In fact, it is straightforward to show that

$$T^\dagger \frac{\partial H}{\partial \epsilon} T = e^{\mathcal{Q}} \frac{\partial H}{\partial \epsilon} e^{-\mathcal{Q}} = e^{\text{ad}_{\mathcal{Q}}} \frac{\partial H}{\partial \epsilon} = \sum_{n=0}^{\infty} \frac{1}{n!} \text{ad}_{\mathcal{Q}}^n \frac{\partial H}{\partial \epsilon}. \quad (3.9)$$

Here we have defined the operator  $\text{ad}_{\mathcal{Q}} B \equiv [\mathcal{Q}, B]$ , with  $\text{ad}_{\mathcal{Q}}^n B \equiv [\mathcal{Q}, \text{ad}_{\mathcal{Q}}^{n-1} B]$ . In consequence, we have

$$\frac{\partial K}{\partial \epsilon} = [L, K] + e^{\text{ad}_{\mathcal{Q}}} \frac{\partial H}{\partial \epsilon} + i\hbar \frac{\partial L}{\partial t}, \quad (3.10)$$

an equation formulated only in terms of  $L, H$  and  $K$  that is easier to handle than (3.8).

In the present formalism, then, three different problems have to be studied:

1. Choose the new Hamiltonian  $K$  such that equation (3.2) is easy to solve.
2. Compute the skew-Hermitian generator  $L$  of the required transformation to the new picture.
3. Construct the unitary transformation  $T$  from the generator  $L$ , or equivalently, the operator  $\mathcal{Q}$  in  $T = \exp(-\mathcal{Q})$ .

The first two problems enumerated earlier can be solved perturbatively with equation (3.10), whereas the third one can be tackled independently. We treat these problems in reverse order.

### (b) Determining $T$

We next proceed to compute  $\mathcal{Q}$ , once  $L$  has been fixed. To do that, we introduce, in addition to the expression (2.1) for the Hamiltonian  $H$ , the following series expansions:

$$K(t, \epsilon) = \sum_{n=0}^{\infty} \epsilon^n K_n(t), \quad (3.11)$$

and

$$L(t, \epsilon) = \sum_{n=0}^{\infty} \epsilon^n L_{n+1}(t). \quad (3.12)$$

It turns out that  $\mathcal{Q}(t, \epsilon)$  in (3.7) can also be determined as a power series in  $\epsilon$ :

$$\mathcal{Q}(t, \epsilon) = \sum_{n=1}^{\infty} \epsilon^n v_n(t), \quad (3.13)$$

with  $v_n(t)$  expressed in terms of  $L_j(t)$ . This can be done algorithmically with the Magnus expansion up to any given order. In fact, the procedure already used to construct numerical integrators based on Magnus expansion

(Blanes *et al.* 2009) is of use here also. Specifically, if the series (3.12) is inserted into the recurrence (3.6), we get up to order  $\epsilon^5$

$$\begin{aligned}\Omega_1 &= \epsilon L_1 + \frac{1}{2}\epsilon^2 L_2 + \frac{1}{3}\epsilon^3 L_3 + \frac{1}{4}\epsilon^4 L_4 + \frac{1}{5}\epsilon^5 L_5, \\ \Omega_2 &= -\frac{1}{12}\epsilon^3 [L_1, L_2] - \frac{1}{12}\epsilon^4 [L_1, L_3] - \epsilon^5 \left( \frac{3}{40}[L_1, L_4] + \frac{1}{60}[L_2, L_3] \right), \\ \Omega_3 &= \epsilon^5 \left( \frac{1}{360}[L_1, [L_1, L_3]] - \frac{1}{240}[L_2, [L_1, L_2]] \right) \\ \text{and} \quad \Omega_4 &= \epsilon^5 \frac{1}{720}[L_1, [L_1, [L_1, L_2]]].\end{aligned}$$

Reordering according with the power of  $\epsilon$ , we have the following expressions for the first terms in the series (3.13):

$$\begin{aligned}v_1 &= L_1, \\ v_2 &= \frac{1}{2}L_2, \\ v_3 &= \frac{1}{3}L_3 - \frac{1}{12}[L_1, L_2], \\ v_4 &= \frac{1}{4}L_4 - \frac{1}{12}[L_1, L_3] \\ \text{and} \quad v_5 &= \frac{1}{5}L_5 - \frac{3}{40}[L_1, L_4] - \frac{1}{60}[L_2, L_3] + \frac{1}{360}[L_1, [L_1, L_3]] \\ &\quad - \frac{1}{240}[L_2, [L_1, L_2]] + \frac{1}{720}[L_1, [L_1, [L_1, L_2]]].\end{aligned}\tag{3.14}$$

This procedure can be carried out for higher orders with a symbolic algebra package. On the other hand, if the series (3.13) is inserted into (3.9), then

$$e^{\text{ad}_\Omega} \frac{\partial H}{\partial \epsilon} = \sum_{n=0}^{\infty} \epsilon^n w_n(t),\tag{3.15}$$

and, again, this series can be determined algorithmically. The first terms read

$$\begin{aligned}w_0 &= H_1, \\ w_1 &= 2H_2 + [L_1, H_1], \\ w_2 &= 3H_3 + 2[L_1, H_2] + \frac{1}{2}[L_2, H_1] + \frac{1}{2}[L_1, [L_1, H_1]] \\ \text{and} \quad w_3 &= 4H_4 + \frac{1}{12}([H_1, [L_1, L_2]] + 2[L_1, [L_1, [L_1, H_1]]]) \\ &\quad + 12[L_1, [L_1, H_2]] + 3[L_1, [L_2, H_1]] + 36[L_1, H_3] \\ &\quad + 3[L_2, [L_1, H_1]] + 12[L_2, H_2] + 4[L_3, H_1].\end{aligned}\tag{3.16}$$

It is worth stressing here that the series (3.13) and (3.15) have to be determined only once. In fact, computing these series up to  $n = 10$  takes only a few minutes on a personal computer. Terms up to  $n = 10$  and  $n = 9$ , respectively, can be found at the website (<http://www.gicas.uji.es/Research/Perturbation.html>) as *Mathematica* files.

### (c) Determining $L$

When the series expansions for  $K$ ,  $L$  and (3.15) are substituted into (3.10), collecting terms of the same power in  $\epsilon$  yields a system of recurrence equations.

Specifically, we arrive at  $K_0 = H_0$  and

$$i\hbar \frac{\partial L_n}{\partial t} + [L_n, H_0] = nK_n - \tilde{F}_n, \quad n = 1, 2, \dots, \quad (3.17)$$

with

$$\tilde{F}_1 = w_0, \quad \tilde{F}_n = \sum_{j=1}^{n-1} [L_{n-j}, K_j] + w_{n-1}, \quad n > 1. \quad (3.18)$$

The next step in the procedure is to propose a suitable  $K_n$  and solve (3.17) to determine  $L_n$  and finally the transformation  $T$ , which is unitary at any order by construction.

Equation (3.17) has the following formal solution:

$$\begin{aligned} L_n(t) = & \exp\left(\frac{-itH_0}{\hbar}\right) L_n(0) \exp\left(\frac{itH_0}{\hbar}\right) - \frac{i}{\hbar} \int_0^t du \exp\left(\frac{-i(t-u)H_0}{\hbar}\right) \\ & \times (nK_n(u) - \tilde{F}_n(u)) \exp\left(\frac{i(t-u)H_0}{\hbar}\right), \end{aligned} \quad (3.19)$$

where  $n > 0$ , and we have taken  $t_0 = 0$  for simplicity. Very often we will fix  $L_n(0) = 0$ , so that

$$L_n(t) = -\frac{i}{\hbar} \int_0^t du e^{-i(t-u)H_0/\hbar} (nK_n(u) - \tilde{F}_n(u)) e^{i(t-u)H_0/\hbar}. \quad (3.20)$$

#### 4. On the choice of $K$

The choice of a particular  $K$  is a degree of freedom of the method but it may also constitute a nuisance. It might be the case that two particular  $H$  and  $K$  are associated in a natural way. Here we will comment about a number of possible choices of  $K$  in the general case.

$$(a) K = H_0$$

Perhaps, this is the most immediate option. It means that systematically

$$K_n = 0 \quad \text{for } n \geq 1.$$

Equivalently, one tries to construct a unitary transformation in such a way that in the new picture there is no perturbation at all. In that case,  $U_K(t) = \exp(-iH_0 t/\hbar)$  and

$$U(t) = T(t, \epsilon) U_K(t) = e^{-\Omega(t, \epsilon)} e^{-(i/\hbar)H_0 t}. \quad (4.1)$$

Notice that with this choice, the whole  $\epsilon$  dependency is contained in  $T$ .

$$(b) K \text{ diagonal}$$

Suppose that  $H_0$  has a pure non-degenerate point spectrum. In that case, we can choose  $K_n$  to be diagonal. In other words, we perturbatively construct a

unitary transformation to a new picture where the Hamiltonian  $K$  is diagonal and therefore equation (3.2) is easy to solve. Specifically,

$$U_K(t) = \exp\left(\frac{-i}{\hbar} \sum_{n \geq 0} \epsilon^n \int_0^t K_n(u) du\right)$$

and finally

$$U(t) = e^{-\Omega(t, \epsilon)} U_K(t).$$

(c) *Time-independent case*

The formalism developed earlier works also when  $H'$  is time-independent. In that case, it makes sense to construct a time-independent unitary transformation  $T(\epsilon)$  so that

$$K(\epsilon) = T^\dagger(\epsilon) H(\epsilon) T(\epsilon) = e^{\Omega(\epsilon)} H(\epsilon) e^{-\Omega(\epsilon)}.$$

A natural choice is then to take the new Hamiltonian so that  $[H_0, K] = 0$ . This is analogous to the standard procedure of constructing the Birkhoff–Gustavson normal form, both in the classical and quantum setting (Ali 1985; Eckhardt 1986). From (3.17), the equations to solve read then

$$[L_n, H_0] = nK_n - \tilde{F}_n, \quad [K_n, H_0] = 0. \quad (4.2)$$

It has been shown in Jauslin *et al.* (2000) that if these equations are defined in the space of bounded skew-Hermitian multiplication operators, then the formal solution can be written in terms of averages as

$$K_n = \frac{1}{n} \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau du e^{-(i/\hbar)H_0 u} \tilde{F}_n e^{(i/\hbar)H_0 u},$$

and

$$L_n = \frac{i}{\hbar} \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau dt \int_0^t ds (e^{-(i/\hbar)H_0 s} \tilde{F}_n e^{(i/\hbar)H_0 s} - nK_n),$$

even if  $H_0$  has degenerate eigenvalues. As a matter of fact, these expressions reproduce the standard results (Birtwistle 1928; Wu 1986). If  $|\nu, j\rangle$  denotes an eigenbasis of  $H_0$ , where  $\nu$  labels the distinct eigenvalues  $E_\nu^0$  and  $j$  distinguishes different basis vectors of the degeneracy subspace, then

$$K_n = \frac{1}{n} \sum_{\nu, j, j'} |\nu, j\rangle \langle \nu, j| \tilde{F}_n |\nu, j'\rangle \langle \nu, j'|,$$

whereas

$$L_n = \frac{i}{\hbar} \sum_{\nu, j, j', \nu' \neq \nu} \frac{|\nu, j\rangle \langle \nu, j| \tilde{F}_n |\nu', j'\rangle \langle \nu', j'|}{E_\nu^0 - E_{\nu'}^0}.$$

The operator evolution in this setting reads

$$U(t) = \exp(-\Omega(\epsilon)) \exp\left(\frac{-i}{\hbar} tK(\epsilon)\right),$$

being unitary even upon truncation of the series. This coincides with the formal solution proposed by Kummer (1971) to avoid secular terms in perturbative

approximations. In this sense, the formalism developed here could be considered as a generalization to time-dependent problems of Kummer's approach.

## 5. Examples

In this section, we apply the previous formalism to two simple examples with the aim of illustrating its main features.

**Example 5.1.** First, we consider the Hamiltonian

$$H(t) = \frac{1}{2}\hbar\omega_0\sigma_3 + \epsilon(\sigma_1 \cos \omega t + \sigma_2 \sin \omega t), \quad (5.1)$$

where  $\sigma_j$  denote the Pauli matrices and  $\epsilon$ ,  $\omega_0$ ,  $\omega \neq \omega_0$  are real parameters. The exact solution of the Schrödinger equation (2.2) with  $U(0) = I$  is then

$$U(t) = \exp\left(\frac{-i}{2}\omega t\sigma_3\right) \exp\left[-it\left(\frac{1}{2}(\omega_0 - \omega)\sigma_3 + \frac{\epsilon}{\hbar}\sigma_1\right)\right].$$

Let 1 and 2 denote the spin up and down states, respectively. Then the exact transition probability from state 1 to state 2 is given by

$$|U_{21}(t)|^2 = \left(\frac{2\epsilon}{\omega'} \sin \frac{\omega' t}{2}\right)^2 \quad \text{with} \quad \omega' = \left((\omega_0 - \omega)^2 + \frac{4\epsilon^2}{\hbar^2}\right)^{1/2},$$

and may serve as a test to analyse the various results in the following.

Our perturbative scheme starts by writing  $H(t) = H_0 + \epsilon H_1(t)$ , with

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

We have applied the procedure up to  $n = 10$  twice. First by choosing  $K_n = 0$  for  $n \geq 1$  and second by taking  $K_n$  as a diagonal matrix. In either case, we get

$$K(t, \epsilon) \simeq \sum_{n=0}^{10} \epsilon^n K_n(t),$$

and the corresponding evolution operator  $U_K(t)$ , whereas

$$L(t, \epsilon) \simeq \sum_{n=0}^{10} \epsilon^n L_{n+1}(t)$$

leads to  $\mathcal{Q}(t, \epsilon)$  up to this order. Finally,

$$U(t) = T(t, \epsilon) U_K(t) \simeq \exp(-\mathcal{Q}(t, \epsilon)) \exp\left(\frac{-i}{\hbar} \int_0^t K(u) du\right).$$

In figure 1, we depict the difference between the exact and approximate transition probabilities as a function of time computed with this procedure up to  $n = 4$  and  $n = 10$ . Solid lines correspond to choosing  $K = H_0$ , whereas dot-dashed curves are generated by taking  $K_n$  as a diagonal matrix. Notice that increasing the number of terms in  $L(t, \epsilon)$  in both cases provides a more accurate approximation.

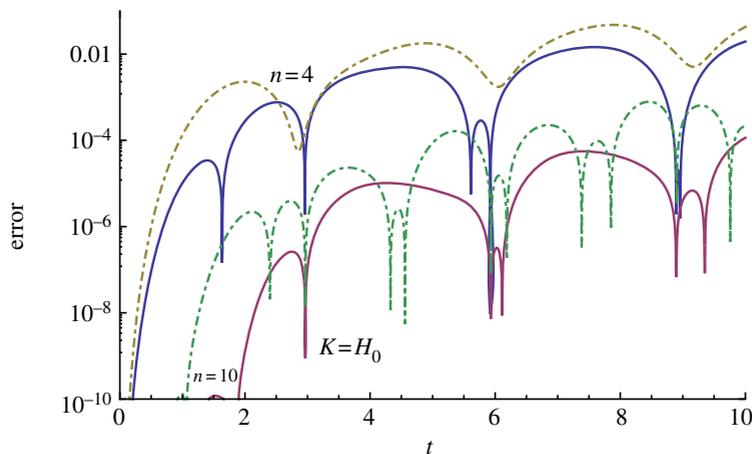


Figure 1. Error in the transition probability for the Hamiltonian (5.1) computed with the unitary perturbation theory developed in this paper, with  $n = 4$  and  $n = 10$ . The parameter values are  $\beta = 0.35$ ,  $\omega = 3$  and  $\omega_0 = 1$ . Solid lines correspond to taking  $K = H_0$ , whereas dotted–dashed curves are obtained by choosing  $K_n$  diagonal,  $n \geq 1$ . (Online version in colour.)

**Example 5.2.** The second example illustrates the time-independent case. We take the two-level spin Hamiltonian

$$H = H_0 + \epsilon H_1 \equiv \frac{1}{2} \hbar \omega_0 \sigma_3 + \epsilon \sigma_1.$$

Then the corresponding equations (4.2) provide for the first orders the following results:

$$\begin{aligned} \text{order } \epsilon^0 : K_0 = H_0, \quad \text{order } \epsilon^1 : K_1 = 0, L_1 = \frac{i}{\omega_0} \sigma_2, \\ \text{order } \epsilon^2 : K_2 = \frac{1}{\omega_0} \sigma_3, L_2 = 0 \quad \text{and} \quad \text{order } \epsilon^3 : K_3 = 0, L_3 = -\frac{4i}{\omega_0^3} \sigma_2. \end{aligned}$$

Carrying out the computation up to order  $\epsilon^{10}$ , we get

$$K(\epsilon) = \left( \frac{\omega_0}{2} + \frac{\epsilon^2}{\omega_0} - \frac{\epsilon^4}{\omega_0^3} + \frac{2\epsilon^6}{\omega_0^5} - \frac{5\epsilon^8}{\omega_0^7} + \frac{14\epsilon^{10}}{\omega_0^9} \right) \sigma_3. \quad (5.2)$$

On the other hand, the exact eigenvalues of the Hamiltonian  $H$  are in this case  $\pm \frac{1}{2} \sqrt{\omega_0^2 + 4\epsilon^2}$  and its expansion in power series of  $\epsilon$  agrees with (5.2). Finally,  $\Omega(\epsilon) = i\omega_a \sigma_2$ , with

$$\omega_a \equiv \frac{\epsilon}{\omega_0} - \frac{4\epsilon^3}{3\omega_0^3} + \frac{16\epsilon^5}{5\omega_0^5} - \frac{64\epsilon^7}{7\omega_0^7} + \frac{256\epsilon^9}{9\omega_0^9} + \mathcal{O}(\epsilon^{11}),$$

and thus the approximate solution is given by  $U(t) \simeq e^{-i\omega_a \sigma_2} e^{-(i/\hbar)K(\epsilon)t}$ .

## 6. Discussion

### (a) Connection with the standard time-dependent perturbation theory

The formalism we have developed here contains as a limiting case the usual time-dependent perturbation theory summarized in §2. This indeed corresponds to choosing  $K = H_0$  as a new Hamiltonian and expanding the operator  $T(t, \epsilon) = e^{-\mathcal{Q}(t, \epsilon)}$  in power series of  $\epsilon$ . Obviously, the resulting scheme is no longer unitary when the series is truncated.

Let us illustrate this point by computing explicitly the first order in  $\epsilon$ . Since  $\tilde{F}_1 = H_1$ , it is clear from (3.20) that

$$L_1(t) = \frac{i}{\hbar} \int_0^t du e^{-i(t-u)H_0/\hbar} H_1(u) e^{i(t-u)H_0/\hbar},$$

whereas

$$U(t) = T(t, \epsilon) e^{-(i/\hbar)H_0 t} = (I - \epsilon L_1(t)) e^{-(i/\hbar)H_0 t} + \mathcal{O}(\epsilon^2).$$

But then

$$\begin{aligned} U(t) &= e^{-(i/\hbar)H_0 t} - \epsilon \frac{i}{\hbar} \int_0^t du e^{-i(t-u)H_0/\hbar} H_1(u) e^{-iuH_0/\hbar} \\ &= e^{-(i/\hbar)H_0 t} \left( I - \epsilon \frac{i}{\hbar} \int_0^t H_I(u) du \right), \end{aligned}$$

with

$$H_I(u) = e^{iuH_0/\hbar} H_1(u) e^{-iuH_0/\hbar}.$$

This, as we know, is precisely the result achieved by standard perturbation theory in the interaction picture defined by  $H_0$ .

The previous treatment can be generalized to any order in  $\epsilon$ . In fact, by following this approach, we have explicitly reproduced up to order  $\epsilon^6$  the results achieved by the usual time-dependent perturbation theory for example 5.1.

### (b) Quantum averaging

In a series of papers published in the mid-1990s, Scherer used the averaging method to construct quantum mechanical analogues of the classical perturbation expansions by Poincaré and Kolmogorov, both in the autonomous case and for time-dependent operators (Scherer 1994, 1997, 1998). The idea was further elaborated by Daems *et al.* (2003a, 2004) and Sugny *et al.* (2004), and successfully applied in the perturbative treatment of pulse-driven quantum problems. This quantum averaging algorithm can also be reproduced with our formalism as follows.

Scherer's basic proposal also consists of transforming the original Hamiltonian (2.1) with the help of a unitary transformation  $T(t, \epsilon)$  such that the problem of finding the time evolution of the transformed Hamiltonian  $K(t, \epsilon)$  is easier than

the original one. In particular, he chooses the terms  $K_n$  in the series defining  $K$  so as to satisfy

$$i\hbar \frac{\partial K_n}{\partial t} + [K_n, H_0] = 0, \quad n \geq 1. \quad (6.1)$$

Then, it can be shown that the time evolution generated by  $K(t, \epsilon)$  is given by ( $t_0 = 0$ )

$$U_K(t) = U_{H_0}(t) \exp\left(-\frac{i}{\hbar} t(K(0, \epsilon) - H_0)\right).$$

In other words, the time evolution is reduced to a time-independent problem if one introduces an interaction picture defined by  $H_0$  for the new Hamiltonian  $K$ . Finally,

$$U(t) = T(t, \epsilon) U_{H_0}(t) \exp\left(\frac{-i}{\hbar} t \sum_{n=1}^{\infty} \epsilon^n K_n(0)\right), \quad (6.2)$$

whereas equation (6.1) can be solved by averaging, generalizing the autonomous case:

$$K_n(t) = \frac{1}{n} \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau du e^{-(i/\hbar)H_0 u} \tilde{F}_n(t-u) e^{(i/\hbar)H_0 u}. \quad (6.3)$$

The unitary transformation  $T(t, \epsilon)$  is then computed order by order from  $L_n(t)$ . In fact, a recursive procedure is proposed for determining the terms  $T_n(t)$  in the series

$$T(t, \epsilon) = I + \sum_{n=1}^{\infty} \epsilon^n T_n(t). \quad (6.4)$$

If not all terms  $L_n$  can be computed, the corresponding truncated series (6.4) provides approximations that are no longer unitary. It is argued, however, that these violations of unitarity do not grow with time, since they do not contain secular terms.

Notice that Scherer's algorithm fits quite naturally in the present scheme. Moreover, with the choice (6.3), we can determine  $L_n(t)$  by applying (3.20) and then compute the transformation as  $T(t, \epsilon) = \exp(-\mathcal{Q}(t, \epsilon))$  with the series (3.13). Now, contrarily to the original formulation, the resulting scheme is unitary upon truncation.

We have applied this modified scheme again to the Hamiltonian (5.1) and compared with the previous choices of  $K_n = 0$  and  $K_n$  diagonal ( $n > 0$ ). In figure 2, we represent the corresponding errors in the transition probability obtained with the three procedures when  $n = 10$  terms are considered in the series. The results illustrate the fairly different character of each approximation. It should be stressed again that the approximation (6.2) with (6.3) considered here is a unitary modification of that originally proposed by Scherer (1997, 1998), i.e. with  $T(t, \epsilon) = \exp(-\mathcal{Q}(t, \epsilon))$  instead of (6.4).

### (c) Open problems and further comments

There are several issues only touched upon in this work that deserve a much more detailed analysis. First, we have been only concerned here with the formal aspects of the transformations and expansions. It is clear, then, that a more

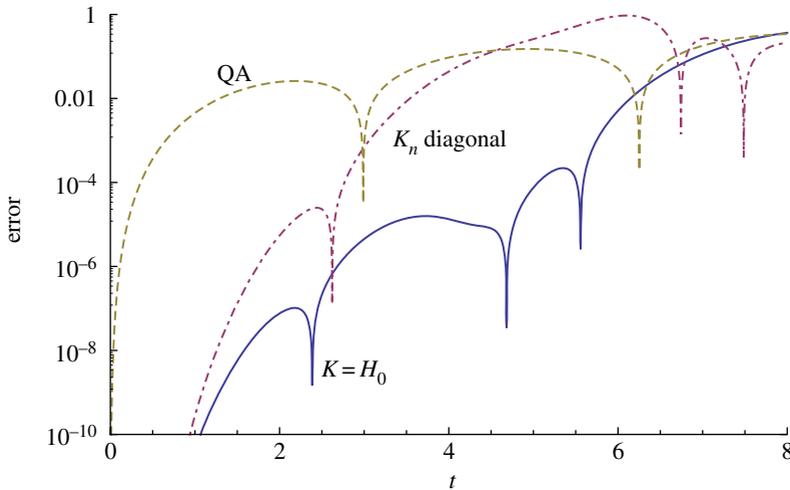


Figure 2. Error in the transition probability for the Hamiltonian (5.1) computed by the new unitary perturbation scheme proposed here, with  $K_n = 0$  ( $n > 0$ ) (solid line) and  $K_n$  diagonal (dotted-dashed line). The dashed curve labelled QA corresponds to the unitary modification of the quantum averaging procedure (6.2). In all cases, we have computed  $n = 10$  terms in the series. The parameter values are  $\beta = 0.4$ ,  $\omega = 1.8$  and  $\omega_0 = 1$ . (Online version in colour.)

rigorous treatment is necessary, with special attention to the convergence of the expansion. Second, the procedure developed here is fairly flexible since it enables different choices of the new Hamiltonian. We have seen on a simple example that different options lead indeed to different approximations; so the natural question is: for any particular problem, which is the optimal choice? In other words, which particular  $K$  allows one to get the best approximation once the expansion is truncated at a given order? In this sense, it is worth mentioning that there are several free parameters in the formalism that eventually may be used to tune the algorithm and improve the accuracy (e.g. Daems *et al.* 2004).

In dealing with quantum time-dependent problems, one also finds in the literature the methods that are iterative in character. So, for example, when the Hamiltonian is slowly varying with time, sequences of unitary transformations intended to diagonalize instantaneously the Hamiltonian lead to non-perturbative approaches (Berry 1987, 1990). In turn, the so-called Fer expansion (Blanes *et al.* 2009) is based on the iterative use of the transformation that defines the interaction picture. A perturbation theory can also be used in an iterative way: the Hamiltonian obtained in each step is used as unperturbed Hamiltonian for the next step. This quantum mechanical version of the KAM scheme has been shown to produce very accurate results in treatments, different to the one proposed in this paper, both for time-independent and time-dependent quantum perturbations (Scherer 1997; Daems *et al.* 2003*a,b*, 2004). The same idea could be implemented with the procedure developed here. In this work, by following closely the Lie-Deprit method in classical mechanics, we have developed a unitary perturbative algorithm well adapted to quantum problems. The technique provides a, in practice approximate, single unitary transformation designed to simplify the dynamics as much as possible.

This work has been partially supported by MICINN (Spain), under grant no. AYA2010-22111-C03-02 (J.A.O. and J.R.) and MTM2010-18246-C03-02 (F.C.), co-financed by FEDER Funds of the European Union.

## References

- Ali, M. K. 1985 The quantum normal form and its equivalents. *J. Math. Phys.* **26**, 2565–2572. (doi:10.1063/1.526775)
- Arnold, V. I., Kozlov, V. & Neishtadt, A. I. 2006 *Mathematical aspects of classical and celestial mechanics*, 3rd edn. Berlin, Germany: Springer.
- Berry, M. V. 1990 Histories of adiabatic quantum transitions. *Proc. R. Soc. Lond. A* **429**, 61–72. (doi:10.1098/rspa.1990.0051)
- Berry, M. V. 1987 Quantum phase corrections from adiabatic iteration. *Proc. R. Soc. Lond. A* **414**, 31–46. (doi:10.1098/rspa.1987.0131)
- Birtwistle, G. 1928 *The new quantum mechanics*. Cambridge, UK: Cambridge University Press.
- Blanes, S., Casas, F., Oteo, J. A. & Ros, J. 2009 The Magnus expansion and some of its applications. *Phys. Rep.* **470**, 151–238. (doi:10.1016/j.physrep.2008.11.001)
- Born, M., Heisenberg, W. & Jordan, P. 1926 Zur Quantenmechanik II. *Z. Phys.* **35**, 557–615 (transl. *Sources of quantum mechanics* (ed. B. L. Van der Waerden), New York, NY: Dover, 1968).
- Daems, D., Keller, A., Guérin, S., Jauslin, H. R. & Atabek, O. 2003a Unitary time-dependent superconvergent technique for pulse-driven quantum dynamics. *Phys. Rev. A* **67**, 052505. (doi:10.1103/PhysRevA.67.052505)
- Daems, D., Guérin, S., Jauslin, H. R., Keller, A. & Atabek, O. 2003b Optimized time-dependent perturbation theory for pulse-driven quantum dynamics in atomic or molecular systems. *Phys. Rev. A* **68**, 051402(R). (doi:10.1103/PhysRevA.68.051402)
- Daems, D., Guérin, S., Jauslin, H. R. & Atabek, O. 2004 Pulse-driven quantum dynamics beyond the impulsive regime. *Phys. Rev. A* **69**, 033411. (doi:10.1103/PhysRevA.69.033411)
- Deprit, A. 1969 Canonical transformations depending on a small parameter. *Celes. Mech.* **1**, 12–30. (doi:10.1007/BF01230629)
- Eckhardt, B. 1986 Birkhoff–Gustavson normal form in classical and quantum mechanics. *J. Phys. A. Math. Gen.* **19**, 2961–2972. (doi:10.1088/0305-4470/19/15/020)
- Jauslin, H. R., Guérin, S. & Thomas, S. 2000 Quantum averaging for driven systems with resonances. *Phys. A* **279**, 432–442. (doi:10.1016/S0378-4371(99)00540-3)
- Klarsfeld, S. & Oteo, J. A. 1989 Recursive generation of higher-order terms in the Magnus expansion. *Phys. Rev. A* **39**, 3270–3273. (doi:10.1103/PhysRevA.39.3270)
- Kummer, M. 1971 How to avoid ‘secular’ terms in classical and quantum mechanics. *Il Nuovo Cimento B* **1**, 123–148.
- Magnus, W. 1954 On the exponential solution of differential equations for a linear operator. *Commun. Pure Appl. Math.* **7**, 649–673. (doi:10.1002/cpa.3160070404)
- Scherer, W. 1994 Quantum averaging. I. Poincaré–von Zeipel is Rayleigh–Schrödinger. *J. Phys. A Math. Gen.* **27**, 8231–8246. (doi:10.1088/0305-4470/27/24/028)
- Scherer, W. 1997 New perturbation algorithms for time-dependent quantum systems. *Phys. Lett. A* **233**, 1–6. (doi:10.1016/S0375-9601(97)00446-5)
- Scherer, W. 1998 Quantum averaging. III. Time-dependent systems. *J. Math. Phys.* **39**, 2597–2634. (doi:10.1063/1.532411)
- Sugny, D., Keller, A. & Atabek, O. 2004 Time-dependent unitary perturbation theory for intense laser-driven molecular orientation. *Phys. Rev. A* **69**, 043407. (doi:10.1103/PhysRevA.69.043407)
- Wu, T.-Y. 1986 *Quantum mechanics*. Singapore: World Scientific.