Quantum averaging II: Kolmogorov's algorithm



1997 J. Phys. A: Math. Gen. 30 2825

(http://iopscience.iop.org/0305-4470/30/8/026)

J. Phys. A: Math. Gen. **30** (1997) 2825–2851. Printed in the UK PII: S0305-4470(97)77781-3

# Quantum averaging II: Kolmogorov's algorithm

Wolfgang Scherer

Institut für Theoretische Physik A, TU Clausthal, Leibnizstr. 10, D-38678 Clausthal-Zellerfeld, Germany

Received 11 September 1996, in final form 9 December 1996

**Abstract.** A complete analogue of Kolmogorov's perturbation algorithm in classical mechanics is presented for perturbations of self-adjoint operators. The resulting perturbation theory is different from the usual Rayleigh–Schrödinger (or Kato–Rellich) perturbation theory.

## 1. Introduction

Despite the fact that quantum mechanics became necessary because classical mechanics could no longer explain the observed phenomena and despite the apparent difference between the two theories, the methods used and the results obtained in classical mechanics have always been an inspiration for the search of similar methods and results in quantum mechanics (e.g. Ehrenfest theorems, adiabatic theorems, etc).

Nevertheless, it seemed that the realm of the perturbation theory was somehow an exception in that the standard Rayleigh–Schrödinger perturbation theory, developed for linear operators (see e.g. [1]), did not seem to bear any resemblance to the usual perturbation theory as first developed by Poincaré [2] and commonly used in classical mechanics. For the first time Kummer [3] (see also [4, 5]) constructed a quantum mechanical perturbation theory in strict analogy with the normal form approach to perturbations in classical mechanics. A similar construction was given by Ali [6] and for the specific example of the anharmonic oscillator by Eckhardt [7]. In a recent paper [8] we have shown that this resemblance between the classical theory and the quantum perturbation theory can be extended to include the method of averaging. In fact, by describing quantum mechanics as a Hamiltonian system, where the Hilbert space of states is viewed as an infinite-dimensional real symplectic manifold, the usual quantum mechanical perturbation theory can be formulated as the 'classical perturbation theory' with canonical (i.e. unitary) transformations which are constructed with the help of the averaging method.

Once a quantum mechanical version of the averaging method is available, however, it may be used to construct a quantum analogue of Kolmogorov's perturbation method [9] which was designed to overcome the problem of small divisors and which is used in the proof of the famous KAM theorem [10, 11]. That is what is done in this paper: the construction of a quantum mechanical version of Kolmogorov's perturbation algorithm.

Kolmogorov's algorithm in classical mechanics consists of an iterative procedure which at each iteration requires the solution of certain differential equations (see equation (19)) which is constructed with the averaging method. Formulating the averaging method in a purely geometric way enables us to write down its quantum analogue, i.e. its analogue for self-adjoint operators. The iterative steps can also be formulated for operators in strict

0305-4470/97/082825+27$19.50  

analogy with the classical situation. In this way one obtains a perturbation algorithm suitable for self-adjoint operators. It should be emphasized that the analogy rests solely on the structure of the equations and its geometrical content and does not make any use of any classical↔quantum relation such as, for example quantization or semi-classical limit.

The paper is organized as follows. Since the quantum analogue is modelled in strict accordance with the classical method we give a review of the latter in section 2. This is a more cursory exhibition of the classical theory and is akin to a proof of the KAM theorem given by Benettin *et al* [12]. The main purpose of section 2 is to motivate the origin of certain similar constructions needed in the quantum version. Hence, most aspects of the classical theory which are not needed in the quantum version will be treated only marginally or not at all. This applies in particular to all conditions necessary for convergence. For a complete and readable account of the classical theory we refer the reader to [12].

In section 3 we develop the quantum version by first exhibiting the quantum analogue of the method of averaging in section 3.1 and then using this to build the perturbation algorithm in section 3.2. In section 3.3 it is then shown how this algorithm serves to approximate eigenvalues and eigenvectors of the perturbed operator. If the unperturbed operator has a purely discrete spectrum these results may be formulated in terms of the unperturbed eigenvalues and eigenvectors which is done in section 3.4. Section 3.5 contains a summary of the algorithm. In the special case of an unperturbed operator with purely discrete and non-degenerate spectrum the expressions simplify considerably and are given in section 3.6. In section 4 we apply the new theory to several examples which show that the quantum perturbation method when constructed is different from, and presumably better than, the usual Rayleigh–Schrödinger theory. We hasten to add that the examples themselves are not sufficient evidence for a considerable improvement over the standard theory. They are meant to illustrate how the new method works and that it does indeed differ from the standard theory. Section 5 contains a discussion of the method presented here, its relation with pre-existing perturbation methods, and several open questions which arise from it.

This paper is an extended version of the results which where announced in [13]. For symbolic and numerical calculations the *Mathematica* package together with its NCAlgebra routines was used.

## 2. Kolmogorov's algorithm in classical mechanics

### *2.1. Generating functions, canonical transformations and averaging*

We begin with some generalities about canonical transformations, their vector fields, and their generating functions. Let $M$ be a $2m$-dimensional symplectic manifold whose symplectic structure has an associated Poisson bracket

$$\{\cdot,\cdot\} : C^\infty(M) \times C^\infty(M) \to C^\infty(M) \tag{1}$$

which allows us to define the following map for each $f \in C^\infty(M)$

$$\begin{aligned} &ad_f : C^\infty(M) \to C^\infty(M) \\ &g \mapsto ad_f(g) := \{f, g\}. \end{aligned} \tag{2}$$

Hence, we shall also use the language

$$f \text{ and } g \text{ commute} :\Leftrightarrow ad_f(g) = \{f, g\} = 0 \tag{3}$$

which is conceptionally useful in view of the quantum mechanical version later. For later use we also define

$$(ad_f)^p(g) := \underbrace{ad_f \circ ad_f \circ \ldots \circ ad_f}_{p \text{ times}}(g) = \underbrace{\{f, \{f, \ldots, \{f}_{p \text{ times}}, g\}\ldots\}. \tag{4}$$

For any smooth map $\varphi : M \to M$ (transformation) we define its pull-back action on functions on $M$ by:

$$\begin{aligned} &\varphi^* : C^\infty(M) \to C^\infty(M) \\ &f \mapsto \varphi^* f := f \circ \varphi. \end{aligned} \tag{5}$$

Each function $f \in C^\infty(M)$ generates the one parameter group of canonical transformations $\xi_f(t) : M \to M$ which satisfy

$$\frac{\mathrm{d}}{\mathrm{d}t}\xi_f(t)^* = \xi_f(t)^* \circ ad_f \tag{6}$$

$$\xi_f(0) = \mathrm{i}d_M. \tag{7}$$

Consequently, the reverse (inverse) flow

$$\varphi_f(t) := \big(\xi_f(t)\big)^{-1} = \xi_f(-t) \tag{8}$$

satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_f(t)^* = -ad_f \circ \varphi_f(t)^* \qquad \varphi_f(0) = \mathrm{i}d_M. \tag{9}$$

In the next section we shall apply this to the case where the generating function is 'time' dependent and where the 'time' parametrizing the flow is just the perturbation parameter $\epsilon$.

Before we do this, however, we shall present the averaging method in this geometric language. Let $f, g \in C^\infty(M)$, $\xi_f(t) : M \to M$ be the flow generated by $f$, and let $f$ be completely integrable with functionally independent integrals $b_i : M \to \mathbb{R}, i = 1, \ldots, m$ such that the map

$$\begin{aligned} &\boldsymbol{b} : M \to \mathbb{R}^m \\ &x \mapsto \{b_1(x), \ldots, b_m(x)\} \end{aligned} \tag{10}$$

has compact inverse images:

$$\boldsymbol{b}^{-1}(\boldsymbol{c}) \subset M \text{ is compact.} \tag{11}$$

This implies [14] that these inverse images are $m$-dimensional tori and that the motion generated by the Hamiltonian $f$ (which by assumption is in involution with all $b_i$) is confined to the torus on which the initial point is to be found. Suppose $g$ is such that

$$\overline{g}^{(f)} := \lim_{T\to\infty} \frac{1}{T} \int_0^T \mathrm{d}s\, \xi_f(-s)^* g \qquad \in C^\infty(M) \tag{12}$$

$$s^{(f)}(g) := \lim_{T\to\infty} \int_0^T \mathrm{d}t \int_0^t \mathrm{d}s\, \xi_f(-s)^*(g - \overline{g}^{(f)}) \qquad \in C^\infty(M) \tag{13}$$

exist. Then it follows that

$$ad_f(\overline{g}^{(f)}) = \quad 0 \tag{14}$$

$$ad_{s^{(f)}(g)}(f) = \overline{g}^{(f)} - g. \tag{15}$$

We first prove (14):

$$\begin{aligned} ad_f(\overline{g}^{(f)}) &= \lim_{T\to\infty}\frac{1}{T}\int_0^T \mathrm{d}s\, ad_f \circ \xi_f(-s)^*(g) \\ &= \lim_{T\to\infty}\frac{1}{T}\int_0^T \mathrm{d}s\, \left(-\frac{\mathrm{d}}{\mathrm{d}s}\xi_f(-s)^*(g)\right) \\ &= -\lim_{T\to\infty}\frac{1}{T}\left(\xi_f(-T)^*(g) - g\right) \\ &= 0 \end{aligned} \tag{16}$$

where the first equality results from the definition of $\overline{g}^{(f)}$ in (12), the second from the properties (6), (8), and (9) of $\xi_f(-s)$, the third is obvious, and the last from the fact that $g \in C^\infty(M)$ and that the tori are compact. Noting that thus

$$\xi_f(-s)^*(\overline{g}^{(f)}) = \overline{g}^{(f)} \tag{17}$$

one shows in a similar fashion that

$$\begin{aligned} ad_{s^{(f)}(g)}(f) &= -ad_f\left(s^{(f)}(g)\right) \\ &= \lim_{T\to\infty}\frac{1}{T}\int_0^T \mathrm{d}t\, \left(\xi_f(-t)^*(g-\overline{g}^{(f)}) - (g-\overline{g}^{(f)})\right) \\ &= \overline{g}^{(f)} - g. \end{aligned} \tag{18}$$

In the construction of Kolmogorov's algorithm it will be repeatedly necessary that for given functions $f$ and $g$ we can find a function $w$ such that

$$ad_f(\,ad_w(f) + g) = 0. \tag{19}$$

Results (14) and (15) thus assure us that we can solve this problem with the help of the averaging constructions (12) and (13) by choosing

$$w = s^{(f)}(g). \tag{20}$$

It should be noted that this is not the only choice since $w = s^{(f)}(g) + v$ would also solve (19) as long as $ad_v(f) = 0$. However, solution (20) allows us to evaluate $ad_w(f)+g = \overline{g}^{(f)}$ directly.

### *2.2. Transforming the Hamiltonian with parameter-dependent transformations*

Now we shall apply the general results of the previous section to the case where the flow parameter is the perturbation parameter, $\epsilon$, and generating functions depend on the flow parameter as well. At the $n$th iteration of Kolmogorov's algorithm we have to construct a canonical transformation $\varphi^n(\epsilon)$ which we shall do by choosing a corresponding generating function $w^n(\epsilon)$ depending on the perturbation parameter in the appropriate way. Let

$$w^n(\epsilon) := \sum_{p=0}^{\infty}\frac{\epsilon^p}{p!}w^n_{p+1} \in C^\infty(M) \tag{21}$$

be a family of $\epsilon$-dependent functions on $M$ such that each $w^n_{p+1} \in C^\infty(M)$ is independent of $\epsilon$ and let

$$\xi^n(\epsilon) := \xi_{-w^n(\epsilon)}(\epsilon) : M \to M \tag{22}$$

be the one-parameter group of canonical transformations generated by $-w^n(\epsilon)$ and let

$$\varphi^n(\epsilon) := \left(\xi^n(\epsilon)\right)^{-1} \tag{23}$$

be the corresponding inverse flow. Consequently, according to (9)

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\varphi^n(\epsilon)^* = ad_{w^n(\epsilon)} \circ \varphi^n(\epsilon)^* \qquad \varphi^n(0) = \mathrm{i}d_M. \tag{24}$$

Assuming that $\varphi^n(\epsilon)^* : C^\infty(M) \to C^\infty(M)$ is analytic in $\epsilon$ we can write:

$$\varphi^n(\epsilon)^* = \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} t_p^n \tag{25}$$

and use (21) and (24) in order to obtain a recursive relation for the $t_p^n$ in terms of $ad_{w_p^n}$:

$$t_0^n = \mathrm{i}d_{C^\infty(M)} \tag{26}$$

$$t_{p+1}^n = \sum_{l=0}^{p} \binom{p}{l} ad_{w_{l+1}^n} \circ t_{p-l}^n \qquad \forall p \geqslant 0. \tag{27}$$

The first few terms in expansion (25) are then

$$\begin{aligned} t_1^n &= ad_{w_1^n} \\ t_2^n &= ad_{w_2^n} + (ad_{w_1^n})^2 \\ t_3^n &= ad_{w_3^n} + (ad_{w_1^n})^3 + (ad_{w_1^n})^2 \circ ad_{w_2^n} + 2ad_{w_2^n} \circ ad_{w_1^n}. \end{aligned} \tag{28}$$

Note that $\varphi^n(\epsilon)^*$ and thus all $t_l^n$ are maps acting on $C^\infty(M)$ whereas the $\varphi^n(\epsilon)$ act on $M$ itself.

Suppose now

$$h(\epsilon) := h_0 + \sum_{p=1}^{\infty} \frac{\epsilon^p}{p!} h_p \tag{29}$$

is the perturbed Hamiltonian $h(\epsilon) : M \to \mathbb{R}$ whose unperturbed part is $h_0 = h(0)$ which is assumed integrable.

In the classical KAM theory we now have to restrict further considerations to a torus $\boldsymbol{b}^{-1}(\boldsymbol{c})$ which has *'sufficiently irrational'* frequencies $\nu_j := \frac{\partial h_0}{\partial b_j}$, i.e. it satisfies a certain Diophantine condition (see, e.g. [10, 12]) which we shall henceforth assume to hold but shall not go into detail about, because the purpose of the discussion of the classical theory is not to give a complete account of it, but to exhibit the structure of the underlying procedure such that it can be *translated to quantum mechanics*.

Moreover, the original proof of the KAM theorem required the use of the implicit function theorem in order to relate the frequencies of the higher approximants to the original ones. Benettin *et al* [12], however, gave a proof using the Lie method which eliminates the use of the implicit function theorem.

For the sake of notational completeness we introduce the following notation

$$h_p^0 := h_p \quad \forall p \in \mathbb{N} \qquad h^0(\epsilon) := \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} h_p^0 \tag{30}$$

and thus

$$h^0(\epsilon) = h(\epsilon). \tag{31}$$

Let $\varphi^1(\epsilon)$ be a canonical transformation as in (24) and let

$$k^1(\epsilon) : = \varphi^1(\epsilon)^* h^0(\epsilon) \tag{32}$$

$$= \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} k_p^1 \tag{33}$$

be the transformed Hamiltonian. Rather than computing a formula for $k_p^1$ in terms of the $t_p^1$ and $h_l^0$ directly from (33) it is more useful for our purposes to obtain a recursive formula in the following manner. Equation (32) implies

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}k^1(\epsilon) = \left(\frac{\mathrm{d}}{\mathrm{d}\epsilon}\varphi^1(\epsilon)^*\right)h^0(\epsilon) + \varphi^1(\epsilon)^*\frac{\mathrm{d}h^0}{\mathrm{d}\epsilon}(\epsilon) \tag{34}$$

$$= ad_{w^1(\epsilon)} \circ \varphi^1(\epsilon)^*(h^0(\epsilon)) + \varphi^1(\epsilon)^*\frac{\mathrm{d}h^0}{\mathrm{d}\epsilon}(\epsilon) \qquad \text{(from (24))} \tag{35}$$

$$= ad_{w^1(\epsilon)} \circ k^1(\epsilon) + \varphi^1(\epsilon)^*\frac{\mathrm{d}h^0}{\mathrm{d}\epsilon}(\epsilon) \qquad \text{(from (32))} \tag{36}$$

and by now using (33) gives the desired recursive formula

$$k_{p+1}^1 = \sum_{l=0}^{p}\binom{p}{l}\left(ad_{w_{l+1}^1}(k_{p-l}^1) + t_{p-l}^1(h_{l+1}^0)\right). \tag{37}$$

where one has to use

$$k_0^1 = h_0^0 \tag{38}$$

which follows from (30) and (32). In particular we have

$$k_1^1 = ad_{w_1^1}(k_0^1) + h_1^0 = ad_{w_1^1}(h_0^0) + h_1^0. \tag{39}$$

It is desired that $k^1(\epsilon)$ commute with $h_0$ up to $\mathrm{O}(\epsilon)$, i.e. that

$$ad_{h_0^0}(k_1^1) = 0. \tag{40}$$

Requirement (40) together with (39) comprises exactly the situation alluded to at the end of section 2.1 and which is solved according to (14) and (15) by the choice

$$w_1^1 = s^{(h_0^0)}(h_1^0) \tag{41}$$

which implies

$$ad_{w_1^1}(h_0^0) = \overline{h_1^0}^{(h_0^0)} - h_1^0 \tag{42}$$

and thus

$$k_1^1 = \overline{h_1^0}^{(h_0^0)}. \tag{43}$$

So far only $w_1^1$ has been determined. For the remaining terms in the expansion of the generating function $w^1(\epsilon)$ we make the choice

$$w_p^1 = 0 \qquad \forall p \geqslant 2 \tag{44}$$

which implies

$$t_p^1 = (ad_{w_1^1})^p \tag{45}$$

and

$$k_{p+1}^1 = ad_{w_1^1}(k_p^1) + \sum_{l=0}^{p}\binom{p}{l}(ad_{w_1^1})^{p-l}(h_{l+1}^0) \qquad p \geqslant 1. \tag{46}$$

Hence, after the first iteration we have the following situation

$$k^1(\epsilon) = h_0^0 + \epsilon\overline{h_1^0}^{(h_0^0)} + \sum_{p=2}^{\infty}\frac{\epsilon^p}{p!}k_p^1 \tag{47}$$

and $\overline{h_1^0}^{(h_0^0)}$ commutes with $h_0$. For the second step we treat the 'commuting' part of $k^1(\epsilon)$ as the unperturbed Hamiltonian and the higher-order terms as perturbations, i.e. we set

$$h^1(\epsilon) := k^1(\epsilon) = \underbrace{h_0^0 + \epsilon\overline{h_1^0}^{(h_0^0)}} + \sum_{p=2}^{\infty} \frac{\epsilon^p}{p!} k_p^1 \tag{48}$$

$$= h_0^1 + \sum_{p=2}^{\infty} \frac{\epsilon^p}{p!} h_p^1 \tag{49}$$

such that

$$h_0^1 = h_0^0 + \epsilon\overline{h_1^0}^{(h_0^0)} \qquad h_1^1 = 0 \qquad h_p^1 = k_p^1 \qquad \forall p \geqslant 2 \tag{50}$$

and $h^1(\epsilon)$ has no perturbation of first order in $\epsilon$. Now we do a second transformation $\varphi^2(\epsilon)$ on $h^1(\epsilon)$

$$k^2(\epsilon) : = \varphi^2(\epsilon)^* h^1(\epsilon) \tag{51}$$

$$= \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} k_p^2 \tag{52}$$

and in the same manner as before one finds that

$$k_0^2 = h_0^1 \tag{53}$$

$$k_{p+1}^2 = \sum_{l=0}^{p} \binom{p}{l} \left( ad_{w_{l+1}^2}(k_{p-l}^2) + t_{p-l}^2(h_{l+1}^1) \right) \qquad \forall p \geqslant 1. \tag{54}$$

Here we choose the generating function $w^2(\epsilon)$ such that

$$w_1^2 = 0 = w_p^2 \qquad \forall p \geqslant 4 \tag{55}$$

which implies

$$k_1^2 = 0 \tag{56}$$

$$k_2^2 = ad_{w_2^2}(k_0^2) + h_2^1 \tag{57}$$

$$k_3^2 = ad_{w_3^2}(k_0^2) + h_3^1 \tag{58}$$

$$k_{p+1}^2 = \binom{p}{1} ad_{w_2^2}(k_{p-1}^2) + \binom{p}{2} ad_{w_3^2}(k_{p-2}^2) + \sum_{l=1}^{p} \binom{p}{l} t_{p-l}^2(h_{l+1}^1). \tag{59}$$

Note that $k_0^2 = h_0^1 = h_0^0 + \epsilon\overline{h_1^1}^{(h_0^0)}$ is the unperturbed Hamiltonian after the first step which by our assumption is completely integrable, hence, in particular its flow can be computed.

Keeping in mind that we are working on a selected torus meeting certain conditions will allow us to assert that the motion of the higher-order approximants remains confined to deformations of the original torus and to apply the averaging construction. Thus (57) and (58) together with the averaging results (14) and (15) allow us to choose $w_2^2$ and $w_3^2$ such that $k_2^2$ and $k_3^2$ commute with $h_0^1$: the choice

$$w_2^2 = s^{(h_0^1)}(h_2^1) \qquad w_3^2 = s^{(h_0^1)}(h_3^1) \tag{60}$$

implies

$$k_2^2 = \overline{h_2^1}^{(h_0^1)} \qquad k_3^2 = \overline{h_3^1}^{(h_0^1)} \tag{61}$$

and $ad_{h_0^1}(k_2^2) = 0 = ad_{h_0^1}(k_3^2)$. Hence, after the second transformation we have

$$k^2(\epsilon) = \underbrace{h_0^0 + \epsilon\overline{h_1^0}^{(h_0^0)}}_{=h_0^1} + \underbrace{\frac{\epsilon^2}{2}\overline{h_2^1}^{(h_0^1)} + \frac{\epsilon^3}{3!}\overline{h_3^1}^{(h_0^1)}}_{\text{commutes with } h_0^1} + \sum_{p=4}^{\infty}\frac{\epsilon^p}{p!}k_p^2. \tag{62}$$

For the third step we treat the 'commuting' part of $k^2(\epsilon)$ as the unperturbed Hamiltonian and the higher-order terms as perturbations, i.e. we set

$$h^2(\epsilon) := k^2(\epsilon) = \underbrace{h_0^0 + \epsilon\overline{h_1^0}^{(h_0^0)} + \frac{\epsilon^2}{2}\overline{h_2^1}^{(h_0^1)} + \frac{\epsilon^3}{3!}\overline{h_3^1}^{(h_0^1)}}_{} + \sum_{p=4}^{\infty}\frac{\epsilon^p}{p!}k_p^2 \tag{63}$$

$$= h_0^2 + \sum_{p=4}^{\infty}\frac{\epsilon^p}{p!}h_p^2 \tag{64}$$

such that

$$h_0^2 = h_0^0 + \epsilon\overline{h_1^0}^{(h_0^0)} + \frac{\epsilon^2}{2}\overline{h_2^1}^{(h_0^1)} + \frac{\epsilon^3}{3!}\overline{h_3^1}^{(h_0^1)} \tag{65}$$

$$h_p^2 = 0 \qquad 1 \leqslant p \leqslant 3 \tag{66}$$

$$h_p^2 = k_p^2 \qquad \forall p \geqslant 4 \tag{67}$$

and $h^2(\epsilon)$ has no perturbation of second and third order in $\epsilon$. Now one does a third transformation $\varphi^3(\epsilon)$ on $h^2(\epsilon)$ etc.

Note that $w_2^2, w_3^2, k_2^2$, and $k_3^2$ are constructed with averaging along the flow of $h_0^1$ and thus acquire an $\epsilon$-dependence. In fact, what we have described above is the final result of an expansion scheme in which one uses a new $\epsilon_n$ in each iteration and at the end sets $\epsilon_n = \ldots = \epsilon_0 = \epsilon$ as follows. Begin with

$$h^0(\epsilon_0) = h_0^0 + \sum_{p=1}^{\infty}\frac{(\epsilon_0)^p}{p!}h_p^0$$

$$w^1(\epsilon_0) = \sum_{p=0}^{\infty}\frac{(\epsilon_0)^p}{p!}w_{p+1}^1$$

$$k^1(\epsilon_0) = \varphi^1(\epsilon_0)^* h^0(\epsilon_0) = \sum_{p=0}^{\infty}\frac{(\epsilon_0)^p}{p!}k_p^1$$

where the $h_p^0, w_{p+1}^1, k_p^1$ do not depend on $\epsilon_0$ and $\varphi^1(\epsilon_0)$ is a solution of (24) for $n = 1$ and $\epsilon = \epsilon_0$. Now expand as above in $\epsilon_0$ to arrive at

$$k^1(\epsilon_0) = h_0^0 + \epsilon_0\overline{h_1^0}^{(h_0^0)} + \sum_{p=2}^{\infty}\frac{(\epsilon_0)^p}{p!}k_p^1$$

and define

$$h^1(\epsilon_1, \epsilon_0) := \underbrace{h_0^0 + \epsilon_0\overline{h_1^0}^{(h_0^0)}}_{=:h_0^1(\epsilon_0)} + \sum_{p=2}^{\infty}\frac{(\epsilon_1)^p}{p!}\underbrace{k_p^1}_{=:h_p^1}$$

$$w^2(\epsilon_1, \epsilon_0) := \sum_{p=0}^{\infty}\frac{(\epsilon_1)^p}{p!}w_{p+1}^2(\epsilon_0)$$

$$k^2(\epsilon_1, \epsilon_0) = \varphi^2(\epsilon_1, \epsilon_0)^* h^1(\epsilon_1, \epsilon_0) = \sum_{p=0}^{\infty}\frac{(\epsilon_1)^p}{p!}k_p^2(\epsilon_0)$$

where $\varphi^2(\epsilon_1, \epsilon_0)$ is defined as the solution of (24) for $n = 2$ and $\epsilon = \epsilon_1$. Now expand in $\epsilon_1$ as above, choose $w_2^2(\epsilon_0) = s^{(h_0^1)}(h_2^1)$, etc. Thus, in the $n$th iteration we expand in $\epsilon_{n-1}$ as was done above and at the end we set $\epsilon_{n-1} = \ldots = \epsilon_0 = \epsilon$. The final result is the expansion described in equations (30)–(67) for $n = 1, 2$. For arbitrary $n$ it is summarized in the next section (without the intermediate step of using a different $\epsilon$ for each iteration). Benettin *et al* [12] consider a perturbation not depending on a parameter $\epsilon$ which corresponds to the case $\epsilon = 1$ and utilizes $\epsilon = 1$-canonical transformation of the type (24). In view of the quantum mechanical procedure we have, however, chosen to treat the case of an $\epsilon$-dependent perturbation.

### *2.3. Summary of the classical algorithm*

Suppose now that after $n - 1$ transformations we have the following Hamiltonian

$$h^{n-1}(\epsilon) = h_0^{n-1} + \sum_{p=2^{n-1}}^{\infty} \frac{\epsilon^p}{p!} h_p^{n-1} \tag{68}$$

such that

$$h_p^{n-1} = 0 \qquad \forall 1 \leqslant p < 2^{n-1}. \tag{69}$$

To this Hamiltonian we apply the transformation $\varphi^n(\epsilon)$ defined by (24) to obtain

$$k^n(\epsilon) = \varphi^n(\epsilon)^* h^{n-1}(\epsilon) \tag{70}$$

$$= \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} k_p^n. \tag{71}$$

The $w_p^n$ are now chosen as

$$w_p^n = \begin{cases} 0 & \text{if } p < 2^{n-1} \\ s^{(h_0^{n-1})}(h_p^{n-1}) & \text{if } 2^{n-1} \leqslant p < 2^n \\ 0 & \text{if } 2^n \leqslant p. \end{cases} \tag{72}$$

This choice leads to the following expansion for $k^n(\epsilon)$

$$k_p^n = \begin{cases} h_0^{n-1} & \text{if } p = 0 \\ 0 & \text{if } 1 \leqslant p < 2^{n-1} \\ \overline{h_p^{n-1}}^{(h_0^{n-1})} & \text{if } 2^{n-1} \leqslant p < 2^n \\ \displaystyle\sum_{l=2^{n-1}-1}^{p-1} \binom{p-1}{l} (ad_{w_{l+1}^n}(k_{p-l-1}^n) + t_{p-l-1}^n(h_{l+1}^{n-1})) & \text{if } 2^n \leqslant p. \end{cases} \tag{73}$$

Thus we have

$$h^n(\epsilon) = h_0^n + \sum_{p=2^n}^{\infty} \frac{\epsilon^p}{p!} h_p^n \tag{74}$$

where

$$h_0^n = h_0^{n-1} + \underbrace{\sum_{p=2^{n-1}}^{2^n-1} \frac{\epsilon^p}{p!} \overline{h_p^{n-1}}^{(h_0^{n-1})}}_{\text{commutes with } h_0^{n-1}} \tag{75}$$

$$h_p^n = k_p^n \qquad \forall p \geqslant 2^n. \tag{76}$$

Equation (75) allows us to write down $h_0^n$ explicitly:

$$h_0^n = h_0^0 + \sum_{l=1}^{n} \left( \sum_{p=2^{l-1}}^{2^l-1} \frac{\epsilon^p}{p!} \overline{h_p^{l-1}(h_0^{l-1})} \right). \tag{77}$$

Note that at each step (i.e. after each transformation) the averaging procedure is executed with a new Hamiltonian which generates the flow along which the average is computed. This is one of the crucial improvements of Kolmogorov's method over the original Poincaré–von Zeipel perturbation theory in classical mechanics and in the quantum mechanical version this will be responsible for the difference to, and (presumably) improvement over, the usual Rayleigh–Schrödinger perturbation theory.

## 3. Kolmogorov's algorithm in quantum mechanics

### *3.1. Quantum mechanical transformations and averaging*

The perturbation theories in classical mechanics (Poincaré–von Zeipel and Kolmogorov) are built with the tools of the symplectic (or Poissonian) geometry of the classical phase space (e.g. generating functions, canonical flows). Using the well known fact (see e.g. [15]) that quantum mechanics may be formulated as an $\infty$-dimensional Hamiltonian system it is possible to construct the quantum perturbation theories using the symplectic language. The resulting algorithm, however, can be described using purely 'quantum mechanical language' (e.g. unitary groups and their generators, which, of course, are disguised objects of the $\infty$-dimensional symplectic geometry of quantum mechanics) and it is this more immediate route to the quantum mechanical Kolmogorov algorithm that we shall adopt.

Let $\mathcal{H}$ be a Hilbert space and $\Phi, F, G : \mathcal{H} \to \mathcal{H}$ linear operators. Then we define the pull-back action $\Phi^*$ of $\Phi$ and the adjoint action $\mathcal{AD}_F$ of $F$ on operators $G$ on $\mathcal{H}$ as

$$\Phi^* A := \Phi^\dagger A \Phi \tag{78}$$

$$\mathcal{AD}_F(G) := \frac{\mathrm{i}}{\hbar}[F, G] \tag{79}$$

where $\Phi^\dagger$ is the adjoint of $\Phi$ and $[\cdot, \cdot]$ denotes the commutator. Let

$$W^n(\epsilon) := \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} W_{p+1}^n \tag{80}$$

be a family of self-adjoint operators such that the $W_p^n$ do not depend on $\epsilon$. Then each $-W^n(\epsilon)$ generates a one parameter group of unitary transformations $\Xi\epsilon : \mathcal{H} \to \mathcal{H}$ which satisfies

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon} \Xi^n(\epsilon) = \frac{\mathrm{i}}{\hbar} W^n(\epsilon) \Xi^n(\epsilon) \qquad \Xi^n(0) = \mathbb{1}. \tag{81}$$

For the inverse transformation

$$\Phi^n(\epsilon) := \Xi^n(\epsilon)^{-1} \tag{82}$$

it follows that

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon} \Phi^n(\epsilon) = -\frac{\mathrm{i}}{\hbar} \Phi^n(\epsilon) W^n(\epsilon) \qquad \Phi^n(0) = \mathbb{1} \tag{83}$$

and for its pull-back we have just as in (24)

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon} \Phi^n(\epsilon)^* = \mathcal{AD}_{W^n(\epsilon)} \circ \Phi^n(\epsilon)^* \qquad \Phi^n(0) = \mathrm{i}d_{\mathcal{H}}. \tag{84}$$

Following the classical construction we also expand $\Phi^n(\epsilon)^*$ in terms of $\epsilon$-independent operators $T_p^n$ (which should be viewed as acting on $L(\mathcal{H})$ (=linear operators on $\mathcal{H}$)):

$$\Phi^n(\epsilon)^* = \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} T_p^n \tag{85}$$

and for those $T_p^n$ we obtain the same recursive formula as in (26) and (27)

$$T_0^n = \mathrm{i}d_{L(\mathcal{H})} \tag{86}$$

$$T_{p+1}^n = \sum_{l=0}^{p} \binom{p}{l} \mathcal{AD}_{W_{l+1}^n} \circ T_{p-l}^n \qquad \forall p \geqslant 0 \tag{87}$$

and the first terms in this expansion are identical to those in (28) with capital letters and the appropriate replacement $ad \to \mathcal{AD}$.

Because we are dealing with linear flows, unlike in the general classical case, we can use (83) to construct recursively the $\epsilon$-expansion

$$\Phi^n(\epsilon) = \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} \Phi_p^n \tag{88}$$

of $\Phi^n(\epsilon)$ itself in terms of the $W_p^n$:

$$\Phi_{p+1}^n = -\frac{\mathrm{i}}{\hbar} \sum_{l=0}^{p} \binom{p}{l} \Phi_{p-l}^n W_{l+1}^n. \tag{89}$$

The first few terms in this expansion are

$$\Phi_0^n = \mathbb{1} \tag{90}$$

$$\Phi_1^n = -\frac{\mathrm{i}}{\hbar} W_1^n \tag{91}$$

$$\Phi_2^n = -\frac{\mathrm{i}}{\hbar} W_2^n - \frac{1}{\hbar^2} (W_1^n)^2 \tag{92}$$

$$\Phi_3^n = -\frac{\mathrm{i}}{\hbar} W_3^n + \frac{\mathrm{i}}{\hbar^3} (W_1^n)^3 - \frac{2}{\hbar^2} W_1^n W_2^n - \frac{1}{\hbar^2} W_2^n W_1^n. \tag{93}$$

Before we turn to Kolmogorov's algorithm itself in section 3.2 we first show that averaging constructions identical to those in (12) and (13) can also be made in quantum mechanics and that they lead to the same results as in (14) and (15). Let $F, G : \mathcal{H} \to \mathcal{H}$ be linear self-adjoint operators and let

$$\Xi_F(t) = \exp\left(-\frac{\mathrm{i}}{\hbar} t F\right) : \mathcal{H} \to \mathcal{H} \tag{94}$$

be the unitary flow generated by $F$. Its pull-back satisfies an equation similar to the classical case (6)

$$\frac{\mathrm{d}}{\mathrm{d}t} \Xi_F(t)^* = \Xi_F(t)^* \circ \mathcal{AD}_F. \tag{95}$$

Similar to the classical case (12) and (13) we now define

$$\overline{G}^{(F)} := \lim_{T\to\infty} \frac{1}{T} \int_0^T \mathrm{d}s\, \Xi_F(-s)^* G \qquad \in L(\mathcal{H}) \tag{96}$$

$$S^{(F)}(G) := \lim_{T\to\infty} \int_0^T \mathrm{d}t \int_0^t \mathrm{d}s\, \Xi_F(-s)^* (G - \overline{G}^{(F)}) \qquad \in L(\mathcal{H}). \tag{97}$$

Again as in the classical case this implies

$$\mathcal{AD}_F(\overline{G}^{(F)}) = 0 \tag{98}$$

$$\mathcal{AD}_{S^{(F)}(G)}(F) = \overline{G}^{(F)} - G \tag{99}$$

provided that one has

$$\lim_{T\to\infty}\left(\frac{\Xi_F(-T)^* G - G}{T}\right) = 0 \tag{100}$$

this being the quantum equivalent of the classical equation (16). In the classical situation the vanishing of the term corresponding to the left-hand side of (100) is a result of the (assumed) compactness of the tori so that (100) can be viewed as the 'quantum analogue of tori'.

The proof of the equalities (98) and (99) is then similar to the classical case since it is based entirely on the geometrical properties of the flow represented in (95). As long as $G$ is bounded (100) may be understood in the uniform topology, i.e. convergence in the operator norm and the limits appearing in (96), (97) exist under (certain conditions) in the weak operator topology. In the case of unbounded $G$ these limits might be understood in the norm resolvent sense [16] but this latter case still needs to be worked out in more detail.

Having established that the method of averaging can also be defined for quantum mechanics we now turn to the perturbation algorithm in the quantum case.

### *3.2. General algorithm*

Let $H_0^0 \in L(\mathcal{H})$ be the unperturbed Hamiltonian and let

$$H^0(\epsilon) := H_0^0 + \sum_{p=1}^{\infty}\frac{\epsilon^p}{p!}H_p^0 \tag{101}$$

be the perturbed Hamiltonian which we transform with $\Phi^1(\epsilon)$ to $K^1(\epsilon)$:

$$K^1(\epsilon) : = \Phi^1(\epsilon)^* H^0(\epsilon) \tag{102}$$

$$= \sum_{p=0}^{\infty}\frac{\epsilon^p}{p!}K_p^1 \tag{103}$$

where one has, as in the classical case,

$$K_0^1 = H_0^0 \tag{104}$$

$$K_{p+1}^1 = \sum_{l=0}^{p}\binom{p}{l}(\mathcal{AD}_{W_{l+1}^1}(K_{p-l}^1) + T_{p-l}^1(H_{l+1}^0)). \tag{105}$$

In particular, one has for $K_1^1$

$$K_1^1 = \mathcal{AD}_{W_1^1}(H_0^0) + H_1^0 \tag{106}$$

such that the choice

$$W_1^1 = S^{(H_0^0)}(H_1^0) \tag{107}$$

$$W_p^1 = 0 \qquad \forall p \geqslant 2 \tag{108}$$

leads to

$$K_1^1 = \overline{H_1^0}^{(H_0^0)} \tag{109}$$

$$K_{p+1}^1 = \mathcal{AD}_{W_1^1}(K_p^1) + \sum_{l=0}^{p}\binom{p}{l}(\mathcal{AD}_{W_1^1})^{p-l}(H_{l+1}^0) \qquad \forall p \geqslant 1 \tag{110}$$

and thus

$$K^1(\epsilon) = H_0^0 + \epsilon\overline{H_1^0}^{(H_0^0)} + \sum_{p=2}^{\infty} \frac{\epsilon^p}{p!} K_p^1 \tag{111}$$

where $\overline{H_1^0}^{(H_0^0)}$ commutes with $H_0^0$. For the second step we treat the 'commuting' part of $K^1(\epsilon)$ as the unperturbed Hamiltonian and the higher-order terms as perturbations, i.e. we set

$$H^1(\epsilon) := K^1(\epsilon) = \underbrace{H_0^0 + \epsilon\overline{H_1^0}^{(H_0^0)}} + \sum_{p=2}^{\infty} \frac{\epsilon^p}{p!} K_p^1 \tag{112}$$

$$= \qquad H_0^1 \qquad + \sum_{p=2}^{\infty} \frac{\epsilon^p}{p!} H_p^1 \tag{113}$$

such that

$$H_0^1 = H_0^0 + \epsilon\overline{H_1^0}^{(H_0^0)} \qquad H_1^1 = 0 \qquad H_p^1 = K_p^1 \qquad \forall p \geqslant 2 \tag{114}$$

and $H^1(\epsilon)$ has no perturbation of first order in $\epsilon$. Now one proceeds exactly as in the classical case and the algorithm may be summarized likewise: suppose now that after $n-1$ transformations we have the following Hamiltonian

$$H^{n-1}(\epsilon) = H_0^{n-1} + \sum_{p=2^{n-1}}^{\infty} \frac{\epsilon^p}{p!} H_p^{n-1} \tag{115}$$

such that

$$H_p^{n-1} = 0 \qquad \forall 1 \leqslant p < 2^{n-1}. \tag{116}$$

To this Hamiltonian we apply the transformation $\Phi^n(\epsilon)$ defined by (84) to obtain

$$K^n(\epsilon) = \Phi^n(\epsilon)^* H^{n-1}(\epsilon) \tag{117}$$

$$= \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} K_p^n. \tag{118}$$

The $W_p^n$ are now chosen as

$$W_p^n = \begin{cases} 0 & \text{if } p < 2^{n-1} \\ S^{(H_0^{n-1})}(H_p^{n-1}) & \text{if } 2^{n-1} \leqslant p < 2^n \\ 0 & \text{if } 2^n \leqslant p. \end{cases} \tag{119}$$

This choice leads to the following expansion for $K^n(\epsilon)$

$$K_p^n = \begin{cases} H_0^{n-1} & \text{if } p = 0 \\ 0 & \text{if } 1 \leqslant p < 2^{n-1} \\ \overline{H_p^{n-1}}^{(H_0^{n-1})} & \text{if } 2^{n-1} \leqslant p < 2^n \\ \displaystyle\sum_{l=2^{n-1}-1}^{p-1} \binom{p-1}{l} (\mathcal{AD}_{W_{l+1}^n}(K_{p-l-1}^n) + T_{p-l-1}^n(H_{l+1}^{n-1})) & \text{if } 2^n \leqslant p. \end{cases} \tag{120}$$

Thus we have

$$H^n(\epsilon) = H_0^n + \sum_{p=2^n}^{\infty} \frac{\epsilon^p}{p!} H_p^n \tag{121}$$

where

$$H_0^n = H_0^{n-1} + \underbrace{\sum_{p=2^{n-1}}^{2^n-1} \frac{\epsilon^p}{p!} \overline{H_p^{n-1}}^{(H_0^{n-1})}}_{\text{commutes with } \mathrm{H}_0^{n-1}} \tag{122}$$

$$H_p^n = K_p^n \qquad \forall p \geqslant 2^n. \tag{123}$$

Equation (122) allows us to write down $H_0^n$ explicitly:

$$H_0^n = H_0^0 + \sum_{l=1}^{n} \left( \sum_{p=2^{l-1}}^{2^l-1} \frac{\epsilon^p}{p!} \overline{H_p^{l-1}}^{(H_0^{l-1})} \right). \tag{124}$$

The first few terms in this expansion are

$$\begin{aligned} H_0^1 &= H_0^0 + \epsilon \overline{H_1^0}^{(H_0^0)} \\ H_0^2 &= H_0^1 + \frac{\epsilon^2}{2!} \overline{H_2^1}^{(H_0^1)} + \frac{\epsilon^2}{3!} \overline{H_3^1}^{(H_0^1)} \\ H_0^3 &= H_0^2 + \frac{\epsilon^4}{4!} \overline{H_4^2}^{(H_0^2)} + \cdots + \frac{\epsilon^7}{7!} \overline{H_7^2}^{(H_0^2)} \end{aligned} \tag{125}$$

where the appearing operators are calculated from (87), (120), and (123) as

$$\begin{aligned} H_2^1 &= H_2^0 + \frac{\mathrm{i}}{\hbar}[W_1^1, \overline{H_1^0}^{(H_0^0)} + H_1^0] \\ H_3^1 &= H_3^0 + 3\frac{\mathrm{i}}{\hbar}[W_1^1, H_2^0] + \left(\frac{\mathrm{i}}{\hbar}\right)^2 [W_1^1, [W_1^1, \overline{H_1^0}^{(H_0^0)} + 2H_1^0]] \\ H_4^1 &= H_4^0 + \frac{\mathrm{i}}{\hbar}[W_1^1, H_3^1 + 3H_3^0] \\ &\quad + 3\left(\frac{\mathrm{i}}{\hbar}\right)^2 [W_1^1, [W_1^1, H_2^0]] + \left(\frac{\mathrm{i}}{\hbar}\right)^3 [W_1^1, [W_1^1, [W_1^1, H_1^0]]] \\ H_4^2 &= H_4^1 + 3\frac{\mathrm{i}}{\hbar}[W_2^2, \overline{H_2^1}^{(H_0^1)} + H_2^1] \\ &\vdots \end{aligned} \tag{126}$$

This shows that we can execute Kolmogorov's algorithm in quantum mechanics in exactly the same way as in classical mechanics if we only make the suitable substitutions $ad \to \mathcal{AD}, h \to H, w \to W, \ldots$. So far it is not clear how this can be used to obtain eigenvalues and eigenvectors of the perturbed Hamiltonian. For this we turn to the following section.

### *3.3. Diagonalization, eigenvalues and eigenvectors*

Let $H_0^0$ be the unperturbed Hamiltonian which we assume to be diagonal in some basis

$$\mathcal{B} := \{|j, \alpha\rangle^0 | \alpha \in D_j\}_{j=1}^{\infty} \tag{127}$$

where $j$ denotes the level, $\alpha$ is the degeneracy-index, $D_j := \{1, \ldots, d_j\}$, and $d_j$ is the degeneracy of the $j$th eigenvalue $E_j^0$ of $H_0^0$:

$$H_0^0 = \sum_{j=1}^{\infty} \sum_{\alpha \in D_j} |j, \alpha\rangle^0 E_j^{00} \langle j, \alpha|. \tag{128}$$

In view of the intended application to quantum mechanics we have started with an $\infty$-dimensional space but the method is, of course, also applicable for operators in finite-dimensional spaces in which case $j$ is confined to a finite index set.

By construction (see (122)) $\overline{H_1^0}^{(H_0^0)}$ commutes with $H_0^0$, hence

$$H_0^1 = H_0^0 + \epsilon \overline{H_1^0}^{(H_0^0)} \tag{129}$$

can be diagonalized in the same basis $\mathcal{B}$. Having done this we note that again by construction $\overline{H_2^1}^{(H_0^1)}$ and $\overline{H_3^1}^{(H_0^1)}$ commute with $H_0^1$, hence

$$H_0^2 = H_0^1 + \frac{\epsilon^2}{2} \overline{H_2^1}^{(H_0^1)} + \frac{\epsilon^3}{3!} \overline{H_3^1}^{(H_0^1)} \tag{130}$$

can be diagonalized in the same basis $\mathcal{B}$. Suppose now that $H_0^{n-1}$ is diagonal in $\mathcal{B}$. By construction (see again (122)) all $\overline{H_p^{n-1}}^{(H_0^{n-1})}$ with $2^{n-1} \leqslant p \leqslant 2^n - 1$ commute with $H_0^{n-1}$ and thus

$$H_0^n = H_0^{n-1} + \sum_{p=2^{n-1}}^{2^n-1} \frac{\epsilon^p}{p!} \overline{H_p^{n-1}}^{(H_0^{n-1})} \tag{131}$$

can be diagonalized in $\mathcal{B}$.

Note that

$$\begin{aligned} H^n(\epsilon) &= \Phi^n(\epsilon)^* \circ \cdots \circ \Phi^1(\epsilon)^* H^0(\epsilon) \\ &= \Phi^n(\epsilon)^\dagger \cdots \Phi^1(\epsilon)^\dagger H^0(\epsilon) \Phi^1(\epsilon) \cdots \Phi^n(\epsilon) \end{aligned} \tag{132}$$

implies that $H^n(\epsilon)$ and our original perturbed Hamiltonian $H(\epsilon) = H^0(\epsilon)$ are unitarily equivalent:

$$H^n(\epsilon) = U^n(\epsilon)^\dagger H^0(\epsilon) U^n(\epsilon) \tag{133}$$

where $U^n(\epsilon)$ is the unitary transformation

$$U^n(\epsilon) := \Phi^1(\epsilon) \cdots \Phi^n(\epsilon). \tag{134}$$

Moreover, one has

$$H^n(\epsilon) = H_0^n + \mathrm{O}(\epsilon^{2^n}) \tag{135}$$

which implies that $H_0^n$ unitarily approximates the original perturbed Hamiltonian:

$$H(\epsilon) = H^0(\epsilon) = U^n(\epsilon) H_0^n U^n(\epsilon)^\dagger + \mathrm{O}(\epsilon^{2^n}). \tag{136}$$

Consequently, since $H_0^n$ can be diagonalized in $\mathcal{B}$ as shown above, we can read off its eigenvalues which coincide with those of our original perturbed Hamiltonian $H(\epsilon)$ up to $\mathrm{O}(\epsilon^{2^n})$. In the formulae:

- Let $E_{j,\alpha}^n(\epsilon)$ be an eigenvalue of $H_0^n$ with eigenvector $|j, \alpha\rangle^n(\epsilon)$:

$$H_0^n \epsilon |j, \alpha\rangle^n(\epsilon) = E_{j,\alpha}^n \epsilon |j, \alpha\rangle^n(\epsilon) \tag{137}$$

- and let $E_{j,\alpha}(\epsilon)$ be an eigenvalue of $H^0(\epsilon) = H(\epsilon)$ with eigenvector $|j, \alpha\rangle(\epsilon)$:

$$H(\epsilon)|j, \alpha\rangle(\epsilon) = E_{j,\alpha}(\epsilon)|j, \alpha\rangle(\epsilon) \tag{138}$$

then

$$E_{j,\alpha}(\epsilon) = E^n_{j,\alpha}(\epsilon) + \mathrm{O}(\epsilon^{2^n}) \tag{139}$$

$$|j,\alpha\rangle(\epsilon) = U^n(\epsilon)|j,\alpha\rangle^n(\epsilon) + \mathrm{O}(\epsilon^{2^n}). \tag{140}$$

Equation (139) gives the desired approximation of the eigenvalues of $H(\epsilon)$ and for the approximation of its eigenvectors one can determine $U^n(\epsilon)$ up to $\mathrm{O}(\epsilon^{2^n})$ in terms of the $W^n_p$ from (88), (89), and (134). For example, one finds

$$U^2(\epsilon) = \mathbb{1} - \left(\frac{\mathrm{i}}{\hbar} W^1_1\right)\epsilon - \left(\frac{\mathrm{i}}{2\hbar} W^2_2 + \frac{1}{2\hbar^2}(W^1_1)^2\right)\epsilon^2 + \left(\frac{\mathrm{i}}{6\hbar^3}(W^1_1)^3 - \frac{\mathrm{i}}{6\hbar} W^2_3 - \frac{1}{2\hbar^2} W^1_1 W^2_2\right)\epsilon^3 + \mathrm{O}(\epsilon^4) \tag{141}$$

In this way one has constructed approximations of the eigenvalues and the eigenvectors of $H(\epsilon)$ to any desired order in $\epsilon$. It should be noted, however, that the expansions obtained this way are not simple power series expansions since the $\overline{H^{n-1}_p}^{(H^{n-1}_0)}$ themselves depend on $\epsilon$. This dependence arises because of the averaging along the flow of $H^{n-1}_0$ which by construction depends on $\epsilon$ and it is this dependence which makes the current perturbation theory distinct from the usual Rayleigh–Schrödinger and which seems to improve convergence because it is akin to a partial re-summation at each step.

### *3.4. Expressions in terms of an eigenbasis*

In this section we shall exhibit the previously developed perturbation algorithm in terms of a basis consisting of eigenvectors of the unperturbed Hamiltonian, $H^0_0$, which is now assumed to have a completely discrete but possibly degenerate spectrum.

To begin with we derive expressions for the averaging constructions (96) and (97) in terms of a basis of eigenvectors. Let $F$ be a self-adjoint operator with a discrete spectrum $\{E^F_j\}^\infty_{j=1}$ with $D_j := \{1, \ldots, d_j = \text{degeneracy of } j\text{th level}\}$ whose eigenvectors $\{|j,\alpha\rangle^F | \alpha \in D_j\}^\infty_{j=0}$ form an orthonormalized basis

$$F = \sum_{j;\alpha\in D_j} |j,\alpha\rangle^F E^F_j {}^F\langle j,\alpha|. \tag{142}$$

The unitary flow generated by $F$ (94) may thus be written as

$$\Xi_F(t) = \sum_{j;\alpha\in D_j} |j,\alpha\rangle^F \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}E^F_j t\, F}\langle j,\alpha| \tag{143}$$

and a simple integration yields

$$\overline{G}^{(F)} = \sum_j \sum_{\alpha,\beta\in D_j} |j,\alpha\rangle^F {}^F\langle j,\alpha|G|j,\beta\rangle^F {}^F\langle j,\beta| \tag{144}$$

for an operator $G$. Similarly one finds that

$$S^{(F)}(G) = \frac{\hbar}{\mathrm{i}} \sum_{j\neq k} \sum_{\alpha\in D_j;\beta\in D_k} |j,\alpha\rangle^F \frac{{}^F\langle j,\alpha|G|k,\beta\rangle^F}{E^F_j - E^F_k} {}^F\langle k,\beta|. \tag{145}$$

Let us now use the same notation as in (137) and (138) with the additional simplification that

$$|j,\alpha\rangle := |j,\alpha\rangle^0 \tag{146}$$

denotes the ($\epsilon$-independent) eigenvector of the unperturbed Hamiltonian $H_0^0$. We assume that for each $n$ and $\epsilon$ the $\{|j,\alpha\rangle^n(\epsilon)|\alpha \in D_j\}_{j=0}^{\infty}$ form an orthonormalized basis. Then it follows that

$$|j,\alpha\rangle^n(\epsilon) = \sum_{\beta\in D_j} c_j^n(\epsilon)_{\alpha\beta}|j,\beta\rangle \tag{147}$$

where the $\epsilon$-dependent $d_j \times d_j$ matrix

$$c_j^n(\epsilon)_{\alpha\beta} = \langle j,\beta|j,\alpha\rangle^n(\epsilon) \tag{148}$$

is unitary. This last statement (unitarity of $c_j^n$) follows trivially from the assumption of orthonormality of the $|j,\alpha\rangle^n(\epsilon)$. The statement that all $|j,\alpha\rangle^n(\epsilon)$ can be written in this form is proven by induction. It is obviously true for $n = 0$. Suppose it is true for $n-1$:

$$|j,\alpha\rangle^{n-1}(\epsilon) = \sum_{\beta\in D_j} c_j^{n-1}(\epsilon)_{\alpha\beta}|j,\beta\rangle \tag{149}$$

and this implies

$$H_0^{n-1} = \sum_j \sum_{\alpha\in D_j} |j,\alpha\rangle^{n-1}(\epsilon)E_{j,\alpha}^{n-1}(\epsilon)(\epsilon)^{n-1}\langle j,\alpha| \tag{150}$$

$$= \sum_j \sum_{\alpha,\beta\in D_j} |j,\alpha\rangle\left(\sum_{\gamma\in D_j} c_j^{n-1}(\epsilon)^{\dagger}_{\beta\gamma}E_{j,\gamma}^{n-1}(\epsilon)c_j^{n-1}(\epsilon)_{\gamma\alpha}\right)\langle j,\beta|. \tag{151}$$

From (144) it follows that

$$\overline{H_p^{n-1}}^{(H_0^{n-1})} = \sum_j \sum_{\alpha,\beta\in D_j} |j,\alpha\rangle^{n-1}(\epsilon)(\epsilon)^{n-1}\langle j,\alpha|H_p^{n-1}|j,\beta\rangle^{n-1}(\epsilon)(\epsilon)^{n-1}\langle j,\beta| \tag{152}$$

but it is easily verified that

$$\sum_{\alpha\in D_j} |j,\alpha\rangle^{n-1}(\epsilon)(\epsilon)^{n-1}\langle j,\alpha| = \sum_{\alpha\in D_j} |j,\alpha\rangle\langle j,\alpha| \tag{153}$$

and thus

$$\overline{H_p^{n-1}}^{(H_0^{n-1})} = \sum_j \sum_{\alpha,\beta\in D_j} |j,\alpha\rangle\langle j,\alpha|H_p^{n-1}|j,\beta\rangle\langle j,\beta|. \tag{154}$$

Inserting this into (131) together with (151) yields

$$H_0^n = H_0^{n-1} + \sum_{p=2^{n-1}}^{2^n-1} \frac{\epsilon^p}{p!}\overline{H_p^{n-1}}^{(H_0^{n-1})} \tag{155}$$

$$= \sum_j \sum_{\alpha,\beta\in D_j} |j,\alpha\rangle\langle j,\alpha|H_p^n|j,\beta\rangle\langle j,\beta| \tag{156}$$

where

$$\langle j,\alpha|H_p^n|j,\beta\rangle = \sum_{\gamma\in D_j} c_j^{n-1}(\epsilon)^{\dagger}_{\beta\gamma}E_{j,\gamma}^{n-1}(\epsilon)c_j^{n-1}(\epsilon)_{\gamma\alpha} + \sum_{p=2^{n-1}}^{2^n-1}\frac{\epsilon^p}{p!}\langle j,\alpha|H_p^{n-1}|j,\beta\rangle \tag{157}$$

is a $d_j \times d_j$ matrix which has to be diagonalized in order to determine $E_{j,\alpha}^n(\epsilon)$ and it follows that the eigenvectors $|j,\alpha\rangle^n(\epsilon)$ of $H_0^n$ are of the form of (147).

With this in mind we then have for $2^{n-1} \leqslant p < 2^n$

$$W_p^n \underbrace{=}_{(119)} S^{(H_0^{n-1})}(H_p^{n-1})$$

$$\underbrace{=}_{(145)} \frac{\overline{h}}{\mathrm{i}} \sum_{j\neq k} \sum_{\substack{\alpha\in D_j \\ \beta\in D_k}} |j,\alpha\rangle^{n-1}(\epsilon) \frac{(\epsilon)^{n-1}\langle j,\alpha|H_p^{n-1}|k,\beta\rangle^{n-1}(\epsilon)}{E_{j,\alpha}^{n-1}(\epsilon) - E_{k,\beta}^{n-1}(\epsilon)} (\epsilon)^{n-1}\langle k,\beta|$$

$$= \frac{\overline{h}}{\mathrm{i}} \sum_{j\neq k} \sum_{\substack{\alpha,\lambda,\mu\in D_j \\ \beta,\gamma,\delta\in D_k}} |j,\lambda\rangle$$

$$\times \frac{c_j^{n-1}(\epsilon)_{\alpha\lambda} c_j^{n-1}(\epsilon)^\dagger_{\mu\alpha} \langle j,\mu|H_p^{n-1}|k,\delta\rangle c_k^{n-1}(\epsilon)_{\beta\delta} c_k^{n-1}(\epsilon)^\dagger_{\gamma\beta}}{E_{j,\alpha}^{n-1}(\epsilon) - E_{k,\beta}^{n-1}(\epsilon)} \langle k,\gamma|. \tag{158}$$

Expression (158) for $W_p^n$ is reminiscent of similar sums over intermediate states appearing in the standard quantum perturbation theory with the notable exception that the denominator now contains the corrected and $\epsilon$-dependent eigenvalues. The rigorous existence of the operators $W_p^n$—which via (119) is equivalent to the existence of operator-valued integrals of the form of (97)—has to be established with functional analytic tools and cannot be ascertained with the present method (see also section 5.2). For example, we cannot *a priori* and without further knowledge of the spectrum of $H_0^0$ guarantee the existence of a non-empty $\epsilon$-interval in which all denominators in (158) are non-zero. Robnik [17], however, has presented heuristic arguments leading to the conjecture that the series for the $W^n$ converge for 'certain admissible perturbations'.

### *3.5. Summary of the quantum mechanical algorithm*

Once the expressions for the $H_p^n$ are available we can start to compute, with the algorithm, eigenvalues and eigenvectors of $H(\epsilon) = H^0(\epsilon)$ by computing those of $H_0^n$. The recursive procedure runs as follows. Starting with $n = 1$ do the following loop.

Step 1: Compute

$$H_0^n = H_0^{n-1} + \sum_{p=2^{n-1}}^{2^n-1} \frac{\epsilon^p}{p!} \overline{H_p^{n-1}}^{(H_0^{n-1})}. \tag{159}$$

It is block-diagonal, hence its diagonalization can be carried out as follows. For each $j$ diagonalize the $d_j \times d_j$ matrix

$$\langle j,\alpha|H_0^n|j,\beta\rangle = \sum_{\gamma\in D_j} c_j^{n-1}(\epsilon)^\dagger_{\beta\gamma} E_{j,\gamma}^{n-1}(\epsilon) c_j^{n-1}(\epsilon)_{\gamma\alpha} + \sum_{p=2^{n-1}}^{2^n-1} \frac{\epsilon^p}{p!} \langle j,\alpha|H_p^{n-1}|j,\beta\rangle \tag{160}$$

i.e. determine its eigenvalues $E_{j,\alpha}^n(\epsilon), \alpha = 1,\ldots,d_j$ and the diagonalizing $d_j \times d_j$ matrix $c_j^n(\epsilon)_{\alpha\beta}$ such that

$$|j,\alpha\rangle^n(\epsilon) = \sum_{\beta\in D_j} c_j^n(\epsilon)_{\alpha\beta} |j,\beta\rangle \tag{161}$$

is the diagonal basis.

If only approximations of eigenvalues up to a given order $\leqslant 2^n - 1$ are desired stop here. The $E_{j,\alpha}^n(\epsilon)$ give an approximation up to $\mathrm{O}(\epsilon^{2^n})$. Otherwise continue.

Step 2: For $2^{n-1} \leqslant p < 2^n$ calculate

$$W_p^n = S^{(H_0^{n-1})}(H_p^{n-1}) = \frac{\overline{h}}{\mathrm{i}} \sum_{j \neq k} \sum_{\substack{\alpha,\lambda,\mu \in D_j \\ \beta,\gamma,\delta \in D_k}} |j,\lambda\rangle \times \frac{c_j^{n-1}(\epsilon)_{\alpha\lambda} c_j^{n-1}(\epsilon)^\dagger_{\mu\alpha} \langle j,\mu|H_p^{n-1}|k,\delta\rangle c_k^{n-1}(\epsilon)_{\beta\delta} c_k^{n-1}(\epsilon)^\dagger_{\gamma\beta}}{E_{j,\alpha}^{n-1}(\epsilon) - E_{k,\beta}^{n-1}(\epsilon)} \langle k,\gamma|. \tag{162}$$

For all other $p$'s set $W_p^n = 0$.

If approximations of eigenvectors up to a given order $\leqslant 2^n - 1$ are desired then use the $W_p^1, \ldots, W_p^n$ in (88), (89), and (134), in order to compute $U^n(\epsilon)$ up to the desired order and calculate the desired approximation to the eigenvector from (140).
Otherwise continue.

Step 3: For $p \geqslant 2^n$ use the $W_p^n$ of step 2 and

$$K_p^n = \begin{cases} H_0^{n-1} & \text{if } p = 0 \\ 0 & \text{if } 1 \leqslant p < 2^{n-1} \\ \displaystyle\sum_j \sum_{\alpha,\beta \in D_j} |j,\alpha\rangle\langle j,\alpha|H_p^{n-1}|j,\beta\rangle\langle j,\beta| & \text{if } 2^{n-1} \leqslant p < 2^n \end{cases} \tag{163}$$

and for $2^n \leqslant p$

$$H_p^n = K_p^n = \sum_{l=2^{n-1}-1}^{p-1} \binom{p-1}{l} (\mathcal{AD}_{W^{l+1}}(K_{p-l-1}^n) + T_{p-l-1}^n(H_{l+1}^{n-1})) \tag{164}$$

to calculate $H_p^n$ in terms of the basis $\{|j,\alpha\rangle | \alpha \in D_j\}_{j=0}^\infty$.

Step 4: Replace $n$ by $n+1$ and go to step 1.

### *3.6. Non-degenerate levels*

In the case of a completely non-degenerate spectrum ($d_j = 1, \forall j$) the algorithm simplifies considerably [13] since we do not have to diagonalize finite-dimensional matrices at each step. First (147) becomes

$$|j\rangle^n(\epsilon) = |j\rangle^0 = |j\rangle \tag{165}$$

and (152) is in this case

$$\overline{H_p^{n-1}}^{(H_0^{n-1})} = \sum_j |j\rangle\langle j|H_p^{n-1}|j\rangle\langle j|. \tag{166}$$

Consequently (155) turns out to be

$$H_0^n = \sum_j |j\rangle \left( E_j^{n-1}(\epsilon) + \sum_{p=2^{n-1}}^{2^n-1} \frac{\epsilon^p}{p!} \langle j|H_p^{n-1}|j\rangle \right) \langle j| \tag{167}$$

which implies a simple formula for the eigenvalues:

$$E_j^n(\epsilon) = E_j^{n-1}(\epsilon) + \sum_{p=2^{n-1}}^{2^n-1} \frac{\epsilon^p}{p!} \langle j|H_p^{n-1}|j\rangle \tag{168}$$

$$= E_j^0 + \sum_{l=1}^n \left( \sum_{p=2^{n-1}}^{2^n-1} \frac{\epsilon^p}{p!} E_j^{(p)}(\epsilon) \right) \tag{169}$$

where we have set

$$E_j^{(p)}(\epsilon) := \langle j|H_p^{k-1}|j\rangle \qquad \text{if } 2^{k-1} \leqslant p \leqslant 2^k - 1. \tag{170}$$

Moreover, (158) becomes

$$W_p^n = \frac{\hbar}{\mathrm{i}} \sum_{j\neq k} |j\rangle \frac{\langle j|H_p^{n-1}|k\rangle}{E_j^{n-1}(\epsilon) - E_k^{n-1}(\epsilon)} \langle k|. \tag{171}$$

In [13] we have worked out the approximations to the eigenvalues including terms of $\mathrm{O}(\epsilon^4)$ for the case of a perturbation which is only linear in $\epsilon$, i.e.

$$H(\epsilon) = H_0^0 + \epsilon H_1^0 = \sum_j |j\rangle^0 E_j^{00} \langle j| + \epsilon \sum_{j,k} |j\rangle^0 V_{jk}{}^0 \langle k| \tag{172}$$

and $H_p^0 = 0, \forall p \geqslant 2$. In this case one finds

$$E_j^{(1)} = V_{jj} \tag{173}$$

$$E_j^{(2)} = 2! \sum_{j\neq k} \frac{|V_{jk}|^2}{E_j^0 - E_k^0} \tag{174}$$

$$E_j^{(3)} = 3!\left( \sum_{m\neq j\neq k} \frac{V_{jk}V_{km}V_{mj}}{(E_j^0 - E_k^0)(E_j^0 - E_m^0)} - \sum_{j\neq k} \frac{|V_{jk}|^2 V_{jj}}{(E_j^0 - E_k^0)^2} \right) \tag{175}$$

which is identical to the standard Rayleigh–Schrödinger result. However, a different theory emerges when we look at the coefficient of $\epsilon^4$ (and higher ones):

$$\begin{aligned}
E_j^{(4)} = 24 \sum_{j\neq l} & \frac{|V_{lj}|^2 (V_{ll} - V_{jj})^2}{(E_j^0 - E_l^0)^2 (E^1(\epsilon)_j - E^1(\epsilon)_l)} \\
& + 6 \sum_{\substack{j\neq l\neq k\neq \\ k\neq j\neq m\neq k}} \frac{V_{jl}V_{lk}V_{km}V_{mj}}{E^1(\epsilon)_j - E^1(\epsilon)_k} \left( \frac{1}{E_j^0 - E_l^0} - \frac{1}{E_l^0 - E_k^0} \right) \\
& \times \left( \frac{1}{E_k^0 - E_m^0} - \frac{1}{E_m^0 - E_j^0} \right) \\
& + 12 \sum_{j\neq l\neq k\neq j} V_{jl}V_{lk}V_{kj} \left\{ \frac{V_{ll} - V_{kk}}{(E_j^0 - E_l^0)(E_l^0 - E_k^0)(E_k^0 - E_j^0)} \right. \\
& + \frac{V_{ll} - V_{jj}}{(E_j^0 - E_l^0)(E^1(\epsilon)_j - E^1(\epsilon)_l)} \left( \frac{1}{E_l^0 - E_k^0} - \frac{1}{E_k^0 - E_j^0} \right) \\
& \left. + \frac{V_{kk} - V_{jj}}{(E_j^0 - E_k^0)(E^1(\epsilon)_j - E^1(\epsilon)_k)} \left( \frac{1}{E_j^0 - E_l^0} - \frac{1}{E_l^0 - E_k^0} \right) \right\} \\
& + \sum_{k\neq l\neq j\neq m\neq k} V_{jl}V_{lk}V_{km}V_{mj} \left\{ \frac{9}{(E_m^0 - E_j^0)(E_j^0 - E_l^0)} \right. \\
& \times \left( \frac{1}{E_k^0 - E_m^0} - \frac{1}{E_l^0 - E_k^0} \right) \\
& \left. + \frac{3}{(E_l^0 - E_k^0)(E_k^0 - E_m^0)} \left( \frac{1}{E_j^0 - E_l^0} - \frac{1}{E_m^0 - E_j^0} \right) \right\}
\end{aligned} \tag{176}$$

which now depends on $\epsilon$ via

$$E_j^1(\epsilon) = E_j^0 + \epsilon V_{jj} \tag{177}$$

making the coefficient of $\epsilon^4$ a function of $\epsilon$. This continues to be the case in all higher orders. In the case where the summation runs over an infinite number of eigenstates the above formula implies that $E_j^{(4)}$ is of the form of

$$E_j^{(4)} = E_j^{(4)}(0) + \lim_{N\to\infty} \frac{Q_j^{N-1}(\epsilon)}{P_j^N(\epsilon)} \tag{178}$$

where $E_j^{(4)}(0)$ does not depend on $\epsilon$, $Q_j^{N-1}(\epsilon)$ is a polynomial of maximal order $N-1$ in $\epsilon$ and $P_j^N(\epsilon)$ is polynomial of maximal order $N$ in $\epsilon$.

It is this feature (emerging more explicitly in the following examples) which makes Kolmogorov's perturbation method distinct from the usual Rayleigh–Schrödinger perturbation theory.

## 4. Examples

### *4.1. A two-dimensional toy model*

To illustrate the method before we turn to more physical applications we consider the following simple two-dimensional Hamiltonian

$$H_0^0 = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} \tag{179}$$

with a perturbation which is only linear in $\epsilon$

$$H_1^0 = \begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix} \qquad H_p^0 = 0 \qquad \forall p \geqslant 2. \tag{180}$$

The exact eigenvalues of the perturbed Hamiltonian $H(\epsilon) = H_0^0 + \epsilon H_1^0$ are

$$E_{1,2}(\epsilon)_{EX} = \tfrac{1}{2}\left(5 + 3\epsilon \pm \sqrt{37\epsilon^2 + 6\epsilon + 9}\right) \tag{181}$$

whereas the usual Rayleigh–Schrödinger perturbation theory gives up to seventh order

$$E_1^3(\epsilon)_{RS} = 1 + \epsilon - 3\epsilon^2 + \epsilon^3 + \tfrac{8}{3}\epsilon^4 - \tfrac{26}{9}\epsilon^5 - \tfrac{109}{27}\epsilon^6 + \tfrac{721}{81}\epsilon^7 \tag{182}$$

$$E_2^3(\epsilon)_{RS} = 4 + 2\epsilon + 3\epsilon^2 - \epsilon^3 - \tfrac{8}{3}\epsilon^4 + \tfrac{26}{9}\epsilon^5 + \tfrac{109}{27}\epsilon^6 - \tfrac{721}{81}\epsilon^7 \tag{183}$$

where in accordance with previous notation for the Kolmogorov theory the superscript $3 = n$ is used here to denote that the expansion is up to $2^{n=3} - 1 =$ seventh order although the number, $n$, of transformations is meaningless in the Rayleigh–Schrödinger perturbation theory. With Kolmogorov's algorithm we find

$$H_0^1 = \begin{pmatrix} 1+\epsilon & 0 \\ 0 & 4+2\epsilon \end{pmatrix} \tag{184}$$

and thus

$$W_1^1 = S^{(H_0^0)}(H_1^0) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{185}$$

with which we compute

$$H_2^1 = \begin{pmatrix} -6 & -2 \\ -2 & 6 \end{pmatrix} \qquad H_3^1 = \begin{pmatrix} 6 & -24 \\ -24 & -6 \end{pmatrix} \qquad H_4^1 = \begin{pmatrix} 72 & 16 \\ 16 & -72 \end{pmatrix} \tag{186}$$

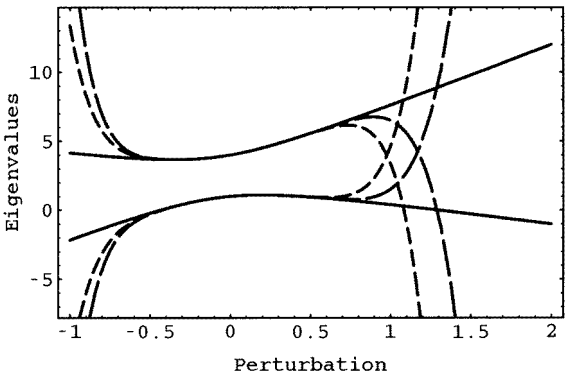


**Figure 1.** The two eigenvalues of the two-dimensional toy model as functions of the perturbation parameter, calculated by explicit diagonalization (full curves), standard perturbation theory (short broken curves), and Kolmogorov's algorithm (long broken curves).

and thus

$$W_2^2 = S^{(H_0^1)}(H_2^1) = \begin{pmatrix} 0 & \frac{2}{3+\epsilon} \\ -\frac{2}{3+\epsilon} & 0 \end{pmatrix} \tag{187}$$

and so on. Working until $n = 3$ (i.e. up to $2^n - 1 =$ seventh order) one finds for the eigenvalues of the perturbed matrix

$$\begin{aligned} E_1^3(\epsilon)_{SU} &= (405 + 675\epsilon - 900\epsilon^2 - 360\epsilon^3 + 1215\epsilon^4 - 405\epsilon^5 - 2295\epsilon^6 + 2385\epsilon^7 \\ &\quad +366\epsilon^8 + 2\epsilon^9)[45(3+\epsilon)^2]^{-1} \\ E_2^3(\epsilon)_{SU} &= (1620 + 1890\epsilon + 1935\epsilon^2 + 495\epsilon^3 - 1215\epsilon^4 + 405\epsilon^5 + 2295\epsilon^6 - 2385\epsilon^7 \\ &\quad -366\epsilon^8 - 2\epsilon^9)[45(3+\epsilon)^2]^{-1}. \end{aligned} \tag{188}$$

In figure 1 all three results $E_{1,2}^3(\epsilon)$ are plotted as a function of $\epsilon$ and it is evident that Kolmogorov's method performs better for both eigenvalues than the usual Rayleigh–Schrödinger perturbation theory.

### *4.2. Anharmonic oscillator*

Turning to more 'physical models' we now treat the anharmonic oscillator in one dimension, i.e. we now have as the Hilbert space $\mathcal{H} = L^2(\mathbb{R}, \mathrm{d}x)$ and

$$H_0^0 = -\frac{\mathrm{d}^2}{\mathrm{d}x^2} + x^2 \qquad H_1^0 = x^4 \qquad H_p^0 = 0 \qquad \forall p \geqslant 2. \tag{189}$$

Although Kolmogorov's algorithm can be formulated in closed form expressions of the form of (96) and (97) we have not been able to evaluate those in the case at hand and have used the expressions given in section 3.4 with the help of the eigenbasis of $H_0^0$. In doing so we can only deal with a finite number of eigenfunctions and operators represented by matrices of finite size. Computational bounds have limited us to the consideration of $16 \times 16$ matrices and calculations of again up to $2^3 - 1 =$ seventh order in $\epsilon$.

Under these restrictions one finds, for example, for the ground-state energy from the standard Rayleigh–Schrödinger theory up to the order of 7 [18]:

$$\begin{aligned} E_0^3(\epsilon)_{RS} &= 1 + \tfrac{3}{4}\epsilon - \tfrac{21}{16}\epsilon^2 + \tfrac{333}{64}\epsilon^3 - \tfrac{30\,885}{1024}\epsilon^4 + \tfrac{916\,731}{4096}\epsilon^5 - \tfrac{65\,518\,401}{32\,768}\epsilon^6 \\ &\quad +\tfrac{2\,723\,294\,673}{131\,072}\epsilon^7 \end{aligned} \tag{190}$$

$$= 1 + 0.75\epsilon - 1.3125\epsilon^2 + 5.203\,12\epsilon^3 - 30.1611\epsilon^4 + 223.811\epsilon^5 - 1999.46\epsilon^6 + 20\,777.0\epsilon^7 \tag{191}$$

whereas Kolmogorov's method gives

$$\begin{aligned} E_0^3(\epsilon)_{SU} &= 1 + \tfrac{3}{4}\epsilon - \tfrac{21}{16}\epsilon^2 + \tfrac{333}{64}\epsilon^3 \\ &\quad - \frac{3(1317\,760 + 12\,935\,472\epsilon + 36\,433\,368\epsilon^2 + 25\,183\,305\epsilon^3)}{2048(4+9\epsilon)(4+15\epsilon)(4+21\epsilon)}\epsilon^4 \\ &\quad + \frac{9(10\,518\,848 + 100\,086\,096\epsilon + 260\,912\,556\epsilon^2 + 140\,051\,835\epsilon^3)}{8192(4+9\epsilon)(4+15\epsilon)(4+21\epsilon)}\epsilon^5 \\ &\quad - \epsilon^6\,[1.144\,41 \times 10^7(2.248\,19 + 59.2375\epsilon + 636.794\epsilon^2 + 3585.19\epsilon^3 \\ &\quad + 11\,259.9\epsilon^4 + 19\,263.9\epsilon^5 + 15\,955\epsilon^6 + 4480.84\epsilon^7)] \\ &\quad \times[(4+9\epsilon)^2(4+15\epsilon)^2(4+21\epsilon)^2(4+27\epsilon)]^{-1} + \epsilon^7 \\ &\quad \times[1.029\,97 \times 10^8(2.401\,02 + 60.9076\epsilon + 623.15\epsilon^2 + 3290.11\epsilon^3 \\ &\quad + 9489.37\epsilon^4 + 14\,456.5\epsilon^5 + 10\,097.2\epsilon^6 + 2183.24\epsilon^7)] \\ &\quad \times[(4+9\epsilon)^2(4+15\epsilon)^2(4+21\epsilon)^2(4+27\epsilon)]^{-1} \end{aligned} \tag{192}$$

$$= 1 + 0.75\epsilon - 1.3125\epsilon^2 + 5.20312\epsilon^3 - 30.1611\epsilon^4 + 223.811\epsilon^5 - 1999.46\epsilon^6 + 20\,777.0\epsilon^7 + \mathcal{O}(\epsilon^8). \tag{193}$$

Equation (193) is a series expansion of the rational function $E_0^3(\epsilon)_{SU}$ and it reproduces the (unique) asymptotic result of the divergent series $E_0(\epsilon)_{RS}$. Of course, in this case no complete analytical solution of the eigenvalue problem is known and thus we have to resort to other numerical results for comparisons. One way to obtain such numerical results is to compute the eigenvalues by finding the zeros of the characteristic polynomial of a finite size matrix for each value of $\epsilon$ (see e.g. [19]). For this purpose we have chosen the diagonalization function in *Mathematica* to compute the lowest eigenvalue of the $16 \times 16$ matrix $H(\epsilon)$ for certain $\epsilon$ and compared its results with the results for the lowest eigenvalue of the anharmonic oscillator as shown in table 1. Figure 2 shows the performance of the three methods ('exact', i.e. numerical diagonalization by *Mathematica*, Rayleigh–Schrödinger, and Kolmogorov's method) and it is again evident that Kolmogorov's method outperforms the standard Rayleigh–Schrödinger theory. This improved performance of Kolmogorov's method over the standard theory as shown in table 1 and figure 2 for the two lowest eigenvalues is actually representative for the whole spectrum of the cut-off anharmonic oscillator Hamiltonian. For larger values of $\epsilon$ both Rayleigh–Schrödinger as well as Kolmogorov's method deviate considerably from the 'exact' results but the latter much less than the former. In any case higher-order computations are needed to exhibit the behaviour of the new method more clearly.

### *4.3. Two-electron atoms*

As an application of the degenerate case we compute the ground-state energies of two-electron atoms whose Hamiltonian is

$$H_{2e}(Z) = -\frac{\hbar^2}{2m_e}(\Delta_{\boldsymbol{x}_1} + \Delta_{\boldsymbol{x}_2}) - \frac{Ze^2}{|\boldsymbol{x}_1|} - \frac{Ze^2}{|\boldsymbol{x}_2|} + \frac{e^2}{|\boldsymbol{x}_1 - \boldsymbol{x}_2|} \tag{194}$$

where $m_e$ is the electron mass (infinite mass of nucleus is assumed), $e$ the electron charge, $Z$ the charge number of the nucleus and $\boldsymbol{x}_i$ denotes the position of the electrons $i = 1, 2$

**Table 1.** Distance of the result from Kolmogorov's algorithm (second column) and the Rayleigh–Schrödinger result (third column) from the numerical result as function of the perturbation (first column) for the ground-state energy of the anharmonic oscillator

| $\epsilon$ | $E_0(\epsilon)_{EX}, -E_0^3(\epsilon)_{SU}$ | $E_0(\epsilon)_{EX}, -E_0^3(\epsilon)_{RS}$ |
|---|---|---|
| 0.01 | $-1.429\,61 \times 10^{-11}$ | $-2.172\,57 \times 10^{-11}$ |
| 0.02 | $-3.158\,14 \times 10^{-9}$ | $-4.994\,68 \times 10^{-9}$ |
| 0.03 | $-7.081\,58 \times 10^{-8}$ | $-1.163\,12 \times 10^{-7}$ |
| 0.04 | $-6.257\,32 \times 10^{-7}$ | $-1.065\,46 \times 10^{-6}$ |
| 0.05 | $-3.329\,32 \times 10^{-6}$ | $-5.867\,8 \times 10^{-6}$ |
| 0.06 | $-1.287\,69 \times 10^{-5}$ | $-2.345\,82 \times 10^{-5}$ |
| 0.07 | $-4.001\,67 \times 10^{-5}$ | $-7.525\,44 \times 10^{-5}$ |
| 0.08 | $-1.060\,5 \times 10^{-4}$ | $-2.056\,36 \times 10^{-4}$ |
| 0.09 | $-2.490\,25 \times 10^{-4}$ | $-4.973\,49 \times 10^{-4}$ |
| 0.1 | $-5.318\,07 \times 10^{-4}$ | $-1.092\,86 \times 10^{-3}$ |

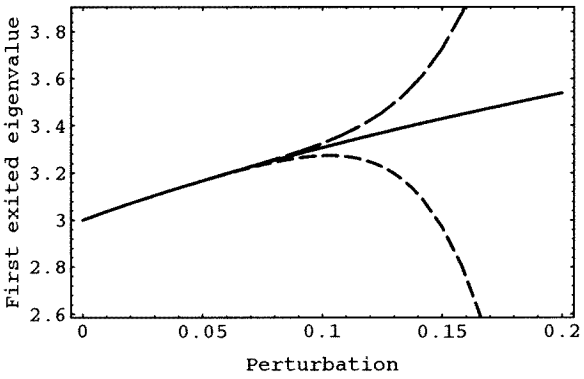


**Figure 2.** First excited energy of the anharmonic oscillator as function of the perturbation parameter, calculated by numerical diagonalization (full curve), standard perturbation theory (short broken curve), and Kolmogorov's algorithm (long broken curve).

relative to the nucleus. We wish to find the energy $\mathcal{E}_0(Z)$ of the spatially symmetric (para) ground state

$$H_{2e}(Z)\psi_0(\boldsymbol{x}_1, \boldsymbol{x}_2, Z) = \mathcal{E}_0(Z)\psi_0(\boldsymbol{x}_1, \boldsymbol{x}_2, Z). \tag{195}$$

In order to treat this problem perturbatively we re-scale the position coordinates of the electrons as [20, 21]

$$\boldsymbol{r}_i := \frac{me^2Z}{\hbar^2}\boldsymbol{x}_i \tag{196}$$

define the perturbation parameter as

$$\epsilon := \frac{1}{Z} \tag{197}$$

and introduce the rescaled energy

$$E_0(\epsilon) := \frac{1}{2R_\infty Z^2}\mathcal{E}_0(Z) \tag{198}$$

(where $R_\infty = 13.605\,804$ eV is the Rydberg energy) such that the eigenvalue problem (195) is equivalent to

$$(H_0^0 + \epsilon H_1^0)\psi_0(\boldsymbol{r}_1, \boldsymbol{r}_2, \epsilon) = E_0(\epsilon)\psi_0(\boldsymbol{r}_1, \boldsymbol{r}_2, \epsilon) \tag{199}$$

**Table 2.** The third column shows the Rayleigh–Schrödinger results, the fourth column shows the results of Kolmogorov's method, and the fifth column shows the experimental results [22].

| $Z$ | Ion | $\mathcal{E}_0(Z)_{RS}$ | $\mathcal{E}_0(Z)_{SU}$ | $\mathcal{E}_0(Z)_{\exp}$ |
|---|---|---|---|---|
| 2 | He | −77.1253 | −77.0507 | −79.003 |
| 3 | Li II | −195.808 | −195.805 | −198.083 |
| 4 | Be III | −369.125 | −369.125 | −371.574 |
| 5 | B IV | −596.938 | −596.938 | −599.502 |
| 6 | C V | −879.205 | −879.205 | −881.885 |
| 7 | N VI | −1215.91 | −1215.91 | −1218.741 |
| 8 | O VII | −1607.05 | −1607.05 | −1610.068 |
| 9 | F VIII | −2052.62 | −2052.62 | −2055.909 |

where now

$$H_0^0 = -\frac{1}{2}(\Delta_{r_1} + \Delta_{r_2}) - \frac{1}{|r_1|} - \frac{1}{|r_2|} \tag{200}$$

$$H_1^0 = \frac{1}{|r_1 - r_2|} \tag{201}$$

and all higher-order perturbations are zero ($H_p^0 = 0, \forall p \geqslant 2$). The equivalent eigenvalue problem (199) is now treated in perturbation theory. Unfortunately the computational resources at hand did not permit the explicit evaluation of the eigenvalue expansion according to Kolmogorov's algorithm in this problem in the case of matrices of dimensions larger than six. The main reason for this is the growth of memory requirement for *Mathematica* if the $\epsilon$-dependent eigenvalues of matrices of dimensions larger than three are to be evaluated.

For the ground-state energy, $E_0(\epsilon)$, of the equivalent problem one finds in Rayleigh–Schrödinger perturbation theory

$$E_0^3(\epsilon)_{RS} = -1 + 0.625\epsilon - 0.051\,279\epsilon^2 - 0.045\,3062\epsilon^3 - 0.033\,3645\epsilon^4 - 0.017\,717\epsilon^5 - 0.001\,82935\epsilon^6 + 0.010\,5439\epsilon^7 \tag{202}$$

whereas Kolmogorov's perturbation theory yields

$$E_0^3(\epsilon)_{SU} = 0.000\,988\,718(-2.09906 + \epsilon)(-0.963\,952\ + \epsilon)(-0.785\,025 + \epsilon) \times (3.834\,48 - 3.198\,94\epsilon + \epsilon^2)(0.973\,851 - 1.940\,77\epsilon + \epsilon^2) \times \frac{(7.779\,92 + 1.672\,84\epsilon + \epsilon^2)(16.7351 + 8.079\,92\epsilon + \epsilon^2)}{(-1.010\,16 + \epsilon)^2(-0.865\,025 + \epsilon)^2} \tag{203}$$

which up to the order of 7 has exactly the same series expansion as $E_0(\epsilon)_{RS}$ in (202) and thus does not give results that differ too much for the ground-state energy, $\mathcal{E}_0(Z)$, of two-electron atoms as is evident from table 2 where the two perturbation theories are compared with the experimental results. So for this limited size of matrices and up to an order of 7 the new perturbation method performs slightly worse than the standard Rayleigh–Schrödinger theory which in turn is not as good as the variational method for such systems [20]. The weak results for $Z = 2$ and $Z = 3$ in table 2 with Kolmogorov's method are due to the double pole of $E_0^3(\epsilon)_{SU}$ at $\epsilon = 0.865\,025$. For the exited energies one finds a similar situation: whereas for $Z \geqslant 3$ Kolmogorov's method yields good approximations a double pole at $\epsilon = 0.489\,689$ leads to very bad results for $Z = 2$.

## 5. Conclusion

### *5.1. Discussion and comparison with other methods*

The question which has been answered affirmatively in this paper was: is it possible to construct a quantum mechanical version of the perturbation algorithm used in the classical KAM theorem? The perturbation method for self-adjoint operators presented here is constructed in complete analogy with Kolmogorov's perturbation theory for classical Hamiltonian systems. It yields approximations to the eigenvalues and to the eigenvectors of the perturbed operators which are no longer pure power series expansions in the perturbation parameter, $\epsilon$, but expansions in terms of functions (which are rational ones in the non-degenerate case) of $\epsilon$. On the purely formal level presented here it is applicable to any self-adjoint operator and free of any limitations such as space dimensions, form of potentials, eigenvalues to be considered, etc.

The method is clearly distinct from the standard Rayleigh–Schrödinger perturbation theory commonly used in such cases. This difference shows up in the $\epsilon$-dependent denominators in the following way. Approximations to the eigenvalues of the perturbed operator are given by the eigenvalues of $H_0^n$. This operator is a function of the $H_p^{n-1}$ (see (159)) which in turn is a function of the $W_p^{n-1}$ (see (120) and (123)) which in turn contain denominators of the type $E_{j,\alpha}^{n-2}(\epsilon) - E_{k,\beta}^{n-2}(\epsilon)$ (see (162)) such that in the $n$th iteration only the corrections up to $n-2$ are taken into account. This is the reason why up to third order the current method does not differ from Rayleigh–Schrödinger perturbation theory since the first-order corrections only start to be taken into account with iteration $n = 3$ yielding corrections from the order of $2^{n-1} = 4$ onwards. This is also the reason why we have not included any systems with zero first-order perturbations in the examples since the difference to the standard method for the second-order perturbation will be felt only from order of 8 onwards.

Having constructed this perturbation method the next question to be answered is: what, if any, is its relation to existing methods? Although in the non-degenerate case this method yields approximations by rational functions or limits thereof (e.g. as in (178)) the expansion method does not give the Padé approximation to the eigenvalue. This negative statement can be proven by computing all Padé approximants of the order of (numerator $\leqslant 9$, denominator $\leqslant 2$) of the exact eigenvalues $E_1^3(\epsilon)_{RS}$ in (181) of our two-dimensional toy model in section 4.1 which shows that none of them matches the rational function $E_1^3(\epsilon)_{SU}$ in (188) which is the result of the new perturbation algorithm.

A great deal of perturbation methods either deal with techniques for special systems or are re-summation techniques of the initially divergent Rayleigh–Schrödinger expansion [19, 23, 24]. In fact, none of the perturbation methods outlined in the books by Bender and Orszag [19], Morse and Feshbach [25], and Arteca *et al* [23] or the reviews by Killingbeck [24] and Simon [26] seem to be related to the method presented in this paper.

### *5.2. Problems and future work*

To put this theory on a rigorous footing one first has to establish criteria for the convergence of the integrals of the type $\overline{G}^{(F)}$ in (96) and $S^{(F)}(G)$ in (97) since they are needed to construct the algorithm. In the case of bounded perturbations, these integrals and limits can be shown to converge under certain assumptions [27]. The unbounded case needs work.

Of course, the most important question still to be answered is if and how Kolmogorov's method converges? Work on this is in progress but from the preliminary results on two-

electron systems in section 4.3 it seems that poles caused by perturbative level-crossing may in certain cases impede improvements on convergence over the standard theory. At this time the numerical results are too scarce and due to the small size of the matrices cannot be considered convincing evidence. Therefore, more extensive computations with a greater number of examples that are done to higher orders are needed.

It should be noted that the non-uniqueness appearing in the classical case (see the remarks after equation (20)) also appears in the quantum analogue. This means that it might be possible to improve the algorithm presented here by optimizing the choice for the operators $W_p^n$ and thus combining the method with some kind of variational technique.

The extension of the new method to time-dependent systems is easier and more straightforward (at least on the formal level). Here we only need to construct the theory in strict analogy with the classical theory which is well known. Work on this is in progress.

Kummer and Gompa [4] have used a perturbation expansion modelled on the classical normal form method to derive rigorous bounds on the quantum mechanical time evolution and recently Delshams and Gutiérrez [28] have exhibited the close relationship between the classical KAM theorem and Nekhoroshev-estimates on the classical time evolution. Using the method presented in this paper it is perhaps possible to get better Nekhoroshev-type estimates for quantum systems.

## References


[1] Kato T 1966 *Perturbation Theory for Linear Operators* (New York: Springer)
[2] Poincaré H 1957 *Les Méthodes Nouvelles de la Mécanique Célèste I, II, III Reprint* (New York: Dover)
[3] Kummer M 1971 *Nuovo Cimento* B **1** 123
[4] Kummer M and Gompa R 1988 *J. Math. Phys.* **29** 1405
[5] Kummer M 1996 *Fields Institute Communications* **8** 35
[6] Ali M K 1985 *J. Math. Phys.* **26** 2565
[7] Eckhardt B 1986 *J. Phys. A: Math. Gen.* **19** 2961
[8] Scherer W 1994 *J. Phys. A: Math. Gen.* **27** 8231
[9] Kolmogorov A N 1954 *Dokl. Akad. Nauk SSSR* **98** 527
[10] Arnold V I 1963 *Russian Math. Surveys* **18** 9
[11] Moser J 1966 *Annali della Scuola normale superiore di Pisa, Scienze fisiche e matematiche* **20** 265 and 499
[12] Benettin G, Galgani L, Giorgilli A and Strelcyn J-M 1984 *Nuovo Cimento* B **79** 201
[13] Scherer W 1995 *Phys. Rev. Lett.* **74** 1495
[14] Arnold V I 1978 *Mathematical Methods of Classical Mechanics* (Berlin: Springer)
[15] Chernoff P R and Marsden J E 1974 *Properties of Infinite Dimensional Hamiltonian Systems (Lecture Notes in Mathematics 425)* (Berlin: Springer)
[16] Reed M and Simon B 1978 *Methods of Modern Mathematical Physics I: Functional Analysis* (New York: Academic)
[17] Robnik M 1986 *J. Phys. A: Math. Gen.* **19** L841
[18] Bender C M and Wu T T 1969 *Phys. Rev.* **184** 1231
[19] Bender C M and Orszag S A 1978 *Advanced Mathematical Methods for Scientists and Engineers* (New York: McGraw-Hill)
[20] Midtdal J 1965 *Phys. Rev.* **138** A1010
[21] Stillinger F H 1966 *J. Chem. Phys.* **45** 3623
[22] Galindo A and Pascal P 1991 *Quantum Mechanics II* (Berlin: Springer)
[23] Arteca G A, Fernandez F M and Castro E A 1990 *Large Order Perturbation Theory and Summation Methods in Quantum Mechanics (Lecture Notes in Chemistry 53)* (Berlin: Springer)
[24] Killingbeck J 1977 *Rep. Prog. Phys.* **40** 963
[25] Morse P M and Feshbach H 1953 *Methods of Theoretical Physics II* (New York: McGraw-Hill)
[26] Simon B 1990 *Bull. Am. Math. Soc.* **24** 303
[27] Scherer W 1997 work in progress
[28] Delshams A and Gutiérrez P 1995 Effective stability and KAM theory *Preprint* (Universitat Politècnica de Catalunya)