Quantum averaging. I. Poincare-von Zeipel is Rayleigh Schrodinger

1994 J. Phys. A: Math. Gen. 27 8231

(http://iopscience.iop.org/0305-4470/27/24/028)

J. Phys. A: Math. Gen. 27 (1994) 8231–8246. Printed in the UK

Quantum averaging: I. Poincaré–von Zeipel is Rayleigh–Schrödinger

Wolfgang Scherer

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

Received 9 August 1994, in final form 21 October 1994

Abstract. An exact analogue of the method of averaging in classical mechanics is constructed for self-adjoint operators. It is shown to be completely equivalent to the usual Rayleigh–Schrödinger perturbation theory but gives the sums over intermediate states in closed-form expressions. The anharmonic oscillator and the Henon–Heiles system are treated as examples to illustrate the quantum averaging method.

1. Introduction and motivation

The failure to obtain exact solutions for most mechanical systems of interest (e.g. planetary motion) has prompted the search for perturbation techniques almost immediately after the conception of Newtonian mechanics (see [1] for some history on the subject). At about the turn of the last century Lindstedt [2], Poincaré [3], and later von Zeipel [4] developed a perturbation method for classical Hamiltonian systems using an averaging procedure in phase space. Despite its lack of convergence in many cases this method, which we shall henceforth refer to as the Poincaré–von Zeipel method, has been widely used, since it yields at least asymptotic expansions.

Concerning the ability to find exact solutions nothing much changed with the advent of quantum mechanics. There it turned out to be equally important to develop perturbation methods and this was done simultaneously with the beginning of quantum mechanics by Schrödinger [5]. Due to previous contributions to similar perturbation techniques in other wave equations by Lord Rayleigh this theory has been named Rayleigh–Schrödinger perturbation theory and was later given a rigorous mathematical basis in the work of Kato [6] and Rellich [7].

In this paper it is shown that the two methods are identical. More precisely, it will be shown that in quantum mechanics an exact analogue of the classical Poincaré–von Zeipel method can be formulated with the help of an averaging techique for self-adjoint operators analogous to the classical method and that the resulting quantum Poincaré–von Zeipel perturbation theory is identical to the Rayleigh–Schrödinger theory.

The analogy between the classical and quantum cases is based entirely on the structure of the equations appearing in the algorithm and the structure of the method used to solve them (averaging). The starting point in the classical case is a Hamiltonian function on phase space with a perturbing Hamiltonian, whereas in the quantum case we start from a self-adjoint Hamiltonian operator with a perturbing operator. If and how the two Hamiltonians are related is of no interest here. No quantization or other quantum-classical map (e.g.

0305-4470/94/248231+16$19.50

semi-classical correspondence) is needed to construct the quantum analogue of the classical Poincaré–von Zeipel theory.

Viewing Rayleigh–Schrödinger as a quantum version of the classical Poincaré–von Zeipel method yields (apart from a purely conceptual viewpoint) one possible advantage: it gives closed expressions for the sums over intermediate states which appear in the corrective terms for the eigenvalues and eigenfunctions.

The method of quantum averaging has also been used to construct a quantum analogue of Kolmogorov's superconvergent perturbation theory [8]. This new quantum 'superconvergent' perturbation theory is substantially different and from the standard Rayleigh–Schrödinger method and initial numerical studies in some examples indicate much better convergence properties [9]. While using quantum averaging to construct analogues of the classical Poincaré–von Zeipel and superconvergent methods they needed to be compared with existing perturbation methods in quantum mechanics. In [9] we have shown that the quantum superconvergent method yields a new kind of perturbation theory and in this paper we show that the quantum Poincaré–von Zeipel method is identical to the standard Rayleigh–Schrödinger theory.

In classical mechanics the Poincaré–von Zeipel series is in most cases divergent and yields only an asymptotic series. This is similar in its quantum equivalent the Rayleigh–Schrödinger series. Since we prove the equivalence of the two methods we shall not state all technical details necessary to make all steps rigorous but refer the reader to the vast mathematical literature dealing with the Rayleigh–Schrödinger (or Kato–Rellich) perturbation theory (see, e.g., [10]). Just as certain quantities diverge in standard Rayleigh–Schrödinger theory (in cases where convergence conditions fail) the power expansions in ϵ formally written down here may not converge in which case the perturbation algorithm gives only asymptotic information, all sums have to be replaced by finite ones up to N, and equations have to be read modulo $O(\epsilon^{N+1})$ for any finite N.

The paper is organized as follows.

In section 2 we present the classical Poincaré–von Zeipel perturbation theory and method of averaging in such a way that it can easily be generalized to quantum mechnics which is done in section 3.

In section 4 we apply the quantum Poincaré–von Zeipel and averaging method to a Hamiltonian with pure point spectrum, we show that up to second order all results from the Rayleigh–Schrödinger theory are reproduced and discuss two examples which illustrate the method and show possible advantages of this new way of constructing the Rayleigh–Schrödinger series.

In section 5 the full equivalence of the two perturbation expansions in all orders is proven.

Finally, in section 6 we discuss previous constructions mimicking classical perturbation expansions in quantum mechanics by Kummer [12, 13], Ali [14], Eckhardt [15], and Ben Lemlih and Ellison [16] and their relation to the present method and conclude with some remarks about future investigations.

2. Classical Poincaré–von Zeipel perturbation theory

In this section we will describe the classical Poincaré–von Zeipel perturbation theory along with the method of averaging in a geometric manner such that its generalization to quantum mechanics is almost self-evident. To avoid later confusion we will use lower-case Latin letters for the classical situation. The unperturbed Hamiltonian h_0 is a function on phase space γ which is equipped with a Poisson bracket structure $\{\cdot,\cdot\}$. h_0 is assumed to be

sufficiently smooth, integrable and non-degenerate in the sense of Liouville–Arnold, i.e. it has $n := \frac{1}{2}\dim\gamma$ functionally independent constants of motion $b_1, \ldots, b_n$ which are in involution, define the invariant tori and have the property that

$$\{h_0, g\} = 0 \Rightarrow g = g(b_1, \ldots, b_n).$$

Furthermore, let

$$h(\epsilon) := \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} h_p \tag{1}$$

be the perturbed Hamiltonian where the perturbations h_p, $p \geqslant 1$ are assumed to be sufficiently smooth functions on γ. The idea of Poincaré–von Zeipel perturbation theory is to look for an ϵ-dependent generating function

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

(with ϵ-independent smooth functions w_l) such that $-w(\epsilon)$ generates a canonical flow $\xi(\epsilon)$ with 'time' ϵ. Then $\varphi(\epsilon) := \xi(\epsilon)^{-1}$ is a transformation on γ determined uniquely by

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\varphi(\epsilon)^* = \operatorname{ad} w(\epsilon) \circ \varphi(\epsilon)^* \tag{3}$$

$$\varphi(0) = \mathrm{id}_\gamma \tag{4}$$

and gives rise to the following action on phase-space functions a:

$$\varphi(\epsilon)^* a := a \circ \varphi(\epsilon). \tag{5}$$

Here $\operatorname{ad} f(g)$ is defined for any two phase-space functions f, g as

$$\operatorname{ad} f(g) := \{f, g\} \tag{6}$$

and for future use we remark that $(\operatorname{ad} w(\epsilon))^p := \operatorname{ad} w(\epsilon) \circ \cdots \circ \operatorname{ad} w(\epsilon)$ (p times). For later use we shall need an expansion of φ^* in terms of some differential operators t_p independent of ϵ:

$$\varphi(\epsilon)^* = \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} t_p. \tag{7}$$

The t_p are then recursively defined through $t_0 = \mathrm{id}_\gamma$ and

$$t_{p+1} = \sum_{l=0}^{p} \binom{p}{l} \operatorname{ad} w_{l+1} \circ t_{p-l}. \tag{8}$$

Acting with $\varphi(\epsilon)^*$ on the perturbed Hamiltonian $h(\epsilon)$ gives a new Hamiltonian

$$k(\epsilon) := \varphi(\epsilon)^* h(\epsilon) \tag{9}$$

which is assumed to be analytic in ϵ:

$$k(\epsilon) = \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} k_p \tag{10}$$

and for which one finds

$$k_0 = h_0 \tag{11}$$

$$k_p = \operatorname{ad} w_p(h_0) + f_p \qquad p \geqslant 1 \tag{12}$$

where $f_1 = h_1$ and for $p \geqslant 2$

$$f_p := h_p + \sum_{l=0}^{p-2} \binom{p-1}{l} (\mathrm{ad}\, w_{l+1}(k_{p-l-1}) + t_{p-l-1} h_{l+1}) . \tag{13}$$

If we had a solution $\psi_{k(\epsilon)}(t) : \gamma \to \gamma$ of the motion with Hamiltonian $k(\epsilon)$ then

$$\psi_{h(\epsilon)}(t) := \varphi(\epsilon) \circ \psi_{k(\epsilon)}(t) \circ \varphi(\epsilon)^{-1} \tag{14}$$

would give us the desired solution for the perturbed Hamiltonian $h(\epsilon)$. In general it is not possible to find $\varphi(\epsilon)$ such that a solution for $k(\epsilon)$ may be found. However, we may choose the w_l successively in such a manner that each k_p is integrable, i.e. (since h_0 is integrable and non-degenerate) such that

$$\mathrm{ad}\, h_0(k_p) = 0 \qquad \forall p . \tag{15}$$

In this way one can trivially solve the equations of motion for $k(\epsilon)$ up to any finite order in ϵ and thus obtain via (14) solutions of $h(\epsilon)$ up to the same order in ϵ.

Thus we have to find the w_p successively such that

$$k_p = \mathrm{ad}\, w_p(h_0) + f_p \tag{16}$$

$$\mathrm{ad}\, h_0(k_p) = 0 . \tag{17}$$

Equation (16) and (17) are solved by the method of averaging [17, 18]. Let $\beta = (\beta_1, \ldots, \beta_n)$ be the coordinates canonically conjugate to the integrals $b = (b_1, \ldots, b_n)$. Then the flow φ_{h_0} generated by the unperturbed Hamiltonian h_0 expressed in the coordinates (b, β) is

$$\varphi_{h_0}(t) \begin{pmatrix} b_0 \\ \beta_0 \end{pmatrix} = \begin{pmatrix} b(t) \\ \beta(t) \end{pmatrix} = \begin{pmatrix} b_0 \\ \beta_0 + \omega(b_0)t \end{pmatrix} \tag{18}$$

where

$$\omega(b_0) := \left(\frac{\partial h_0}{\partial b_1}(b_0), \ldots, \frac{\partial h_0}{\partial b_n}(b_0) \right) \tag{19}$$

gives the frequencies of the unperturbed motion which we assume to be independent over the rationals (non-resonant) for the given b_0, i.e. $c \cdot \omega(b_0) = 0 \Rightarrow c = 0$ for any $c = (c_1, \ldots, c_n)$ with integer c_j. Let g be any function on phase space which has the Fourier decomposition

$$g(b, \beta) = \sum_{c \in \mathbb{Z}^n} g(b)_c e^{ic \cdot \beta} \tag{20}$$

and define the phase-space functions

$$\overline{g} := \lim_{T \to \infty} \frac{1}{T} \int_0^T \mathrm{d}t\, \varphi_{h_0}(-t)^* g \tag{21}$$

$$s(g) := \lim_{T \to \infty} \frac{1}{T} \int_0^T \mathrm{d}t \int_0^t \mathrm{d}s\, \left(\varphi_{h_0}(-s)^* g - \overline{g} \right) \tag{22}$$

then one finds

$$\overline{g} = g(b)_0 \tag{23}$$

$$s(g) = \sum_{c \in \mathbb{Z}^n - \{0\}} \frac{g(b)_c}{ic \cdot \omega(b)} e^{ic \cdot \beta} \tag{24}$$

and it is easy to see that

$$\overline{g} = \mathrm{ad}\,(s(g))(h_0) + g \tag{25}$$

$$\mathrm{ad}\,h_0(\overline{g}) = 0\,. \tag{26}$$

With this construction we now make the following choice for w_p in (16):

$$w_p = s(f_p) \tag{27}$$

then $k_p = \overline{f_p}$ commutes with h_0 as desired. The important point to note here is that we have formulated the basic (averaging) constructions (21) and (22) necessary to solve (16) and (17) in a coordinate-free way. This geometric description using the time average is immediately suitable for generalization to self-adjoint operators, i.e. to quantum mechanics. It should be noted, however, that (27) is not the only possible solution of (16) and (17) since

$$w_p' := w_p + v_p \tag{28}$$

is also a solution of our problem as long as

$$\mathrm{ad}\,v_p(h_0) = 0\,. \tag{29}$$

This non-uniqueness will also emerge in the quantum mechanical setting since it is also present in Kato's rigorous exposition of Rayleigh–Schrödinger perturbation theory [6].

3. Quantum Poincaré–von Zeipel perturbation theory and averaging

Now we shall develop the quantum mechanical analogue of the classical theory presented in the previous section. For this purpose we use upper-case (capital) Latin letters to denote operators on some Hilbert space Γ. Let H_0 be the unperturbed Hamiltonian operator which is assumed to be diagonalized in some basis and let

$$H(\epsilon) := \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} H_p \tag{30}$$

be the perturbed Hamiltonian. Here we do not impose any conditions (like, e.g., boundedness) on the perturbations H_p, $p \geqslant 1$, but proceed purely on a formal level. A mathematically rigorous justification of each step is notoriously intricate and will not be attempted here because it would completely obscure the basic ideas of the method. Since ultimately we shall prove the equivalence of this method to the usual Rayleigh–Schrödinger perturbation theory the conditions of the latter theory needed to guarantee convergence (see, e.g., [10]) will be sufficient to make the quantum Poincaré–von Zeipel perturbation theory convergent as well. In analogy with the classical situation we seek a self-adjoint generator (operator)

$$W(\epsilon) := \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} W_{p+1} \tag{31}$$

such that $-W(\epsilon)$ induces the unitary flow $\Xi(\epsilon)$ with 'time' ϵ, i.e.

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\Xi(\epsilon) = \frac{\mathrm{i}}{\hbar} W(\epsilon)\Xi(\epsilon) \qquad \Xi(0) = \mathbb{1}\,. \tag{32}$$

Then $\Phi(\epsilon) := \Xi(\epsilon)^{-1}$ is the unique solution of the initial value problem

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\Phi(\epsilon)^* = \mathrm{AD}\, W(\epsilon) \circ \Phi(\epsilon)^* \tag{33}$$

$$\Phi(0) = \mathbb{1} \tag{34}$$

where $\Phi(\epsilon)^*$ acts on any operator A via

$$\Phi(\epsilon)^* A := \Phi(\epsilon)^{-1} A\, \Phi(\epsilon) \tag{35}$$

and where $\mathrm{AD}\, F(G)$ is now defined as

$$\mathrm{AD}\, F(G) := \tfrac{\mathrm{i}}{\hbar}[F, G] \tag{36}$$

for any two operators F, G (again we omit the technical details necessary to make (36) well defined for unbounded operators) and as in the classical case $(\mathrm{AD}\, W(\epsilon))^p := \mathrm{AD}\, W(\epsilon) \circ \cdots \circ \mathrm{AD}\, W(\epsilon)$ (p times).

As in the classical case, it will be useful to expand Φ^* in terms of ϵ-independent operators T_p:

$$\Phi(\epsilon)^* = \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} T_p \tag{37}$$

where the T_p are then recursively defined through $T_0 = \mathbb{1}$ and

$$T_{p+1} = \sum_{l=0}^{p} \binom{p}{l} \mathrm{AD}\, W_{l+1} \circ T_{p-l}\,. \tag{38}$$

Note that $\Phi(\epsilon)^*$ and thus the T_p act on operators, whereas $\Phi(\epsilon)$ itself is a transformation on Hilbert space which can also be expanded as

$$\Phi(\epsilon) = \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} \Phi_p \tag{39}$$

and where the following recursive relation for the Φ_p can be derived from (32):

$$\Phi_{p+1} = -\frac{\mathrm{i}}{\hbar} \sum_{l=0}^{p} \binom{p}{l} \Phi_{p-l} W_{l+1} \tag{40}$$

and $\Phi_0 = \mathbb{1}$.

Transforming the perturbed Hamiltonian $H(\epsilon)$ with $\Phi(\epsilon)$ gives a new Hamiltonian

$$K(\epsilon) := \Phi(\epsilon)^* H(\epsilon) = \Phi(\epsilon)^{-1} H(\epsilon) \Phi(\epsilon) \tag{41}$$

which is assumed to be analytic in ϵ:

$$K(\epsilon) = \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} K_p \tag{42}$$

and for which one finds

$$K_0 = H_0 \tag{43}$$

$$K_p = \mathrm{AD}\, W_p(H_0) + F_p \qquad p \geqslant 1 \tag{44}$$

with $F_1 = H_1$ and for $p \geqslant 2$

$$F_p := H_p + \sum_{l=0}^{p-2} \binom{p-1}{l} (\mathrm{AD}\, W_{l+1}(K_{p-l-1}) + T_{p-l-1} H_{l+1})\,. \tag{45}$$

All these equations are exactly analogous to the classical case, but it should be emphasized that they are perfectly well defined operator equations. But how are we to choose W now? If we could diagonalize $K(\epsilon)$ up to a given finite order in ϵ, we could read off its eigenvalues and eigenvectors to that order as well, but this means that we have found the eigenvalues and the eigenvectors of the perturbed Hamiltonian $H(\epsilon)$, since by (41) H and K are unitarily equivalent. Before we write this out in formulae let us first see how we can diagonalize K order by order using the method of quantum averaging. It is obvious that the necessary and sufficient condition for diagonalization is the equivalent of (17), i.e. requiring

$$\mathrm{AD}\, H_0(K_p) = 0 \qquad p \geqslant 1 \tag{46}$$

means that all K_p commute with H_0 and thus H_0 and K_p can be diagonalized simultaneously, hence K can be made diagonal to any finite order N in ϵ. Consequently, in order to diagonalize K to any finite order we need to solve the quantum analogue of (16) and (17), i.e.

$$K_p = \mathrm{AD}\, W_p(h_0) + F_p \tag{47}$$

$$\mathrm{AD}\, H_0(K_p) = 0\,. \tag{48}$$

This, too, is done in a manner analogous to that in the classical case. Let Φ_{H_0} be the unitary flow generated by the unperturbed Hamiltonian H_0 such that for any operator G

$$\Phi_{H_0}(-t)^* G = \exp(-\tfrac{\mathrm{i}}{\hbar} t H_0)\, G\, \exp(\tfrac{\mathrm{i}}{\hbar} t H_0) =: G(t) \tag{49}$$

where the last equation introduces a simplified notation. Suppose now that G is such that

$$\overline{G} := \lim_{T\to\infty} \frac{1}{T} \int_0^T \mathrm{d}t\, G(t) \tag{50}$$

$$S(G) := \lim_{T\to\infty} \frac{1}{T} \int_0^T \mathrm{d}t \int_0^t \mathrm{d}s\, \left(G(s) - \overline{G}\right) \tag{51}$$

exist, and such that

$$\lim_{T\to\infty} \frac{G(T) - G}{T} = 0 \tag{52}$$

then it follows that:

$$\overline{G} = \mathrm{AD}\,(S(G))(H_0) + G \tag{53}$$

$$\mathrm{AD}\, H_0(\overline{G}) = 0\,. \tag{54}$$

We first prove (54):

$$\begin{aligned}
\mathrm{AD}\, H_0(\overline{G}) &= \frac{\mathrm{i}}{\hbar}[H_0, \overline{G}] \\
&= \lim_{T\to\infty} \frac{1}{T} \int_0^T \frac{\mathrm{i}}{\hbar}[H_0, G(t)]\, \mathrm{d}t \\
&= -\lim_{T\to\infty} \frac{1}{T} \int_0^T \frac{\mathrm{d}}{\mathrm{d}t} G(t)\, \mathrm{d}t \\
&= -\lim_{T\to\infty} \frac{G(T) - G}{T} = 0
\end{aligned}$$

by assumption. Moreover

$$\begin{aligned}
\mathrm{AD}\,(S(G))(H_0) &= \lim_{T\to\infty}\frac{1}{T}\int_0^T \mathrm{d}t \int_0^t \mathrm{d}s\,\frac{\mathrm{i}}{\hbar}[G(s)-\overline{G}, H_0] \\
&= \lim_{T\to\infty}\frac{1}{T}\int_0^T \mathrm{d}t \int_0^t \mathrm{d}s\,\frac{\mathrm{i}}{\hbar}[G(s), H_0] \\
&= \lim_{T\to\infty}\frac{1}{T}\int_0^T \mathrm{d}t \int_0^t \mathrm{d}s\,\frac{\mathrm{d}}{\mathrm{d}s}G(s) \\
&= \lim_{T\to\infty}\frac{1}{T}\int_0^T \mathrm{d}t\ (G(t)-G) \\
&= \overline{G}-G
\end{aligned}$$

which proves (53). Equations (50)–(54) are the exact quantum analogue of the classical averaging technique with the noteworthy absence of any non-resonance condition. Thus for any $p \geqslant 1$ equation (47) and (48) are successively solved by

$$W_p = S(F_p) \tag{55}$$

$$K_p = \overline{F_p}\,. \tag{56}$$

Using (45) we may simplify $\overline{F_p}$ by noting that

$$\overline{\mathrm{AD}\,W_{l+1}(\overline{F_{p-l-1}})} = \mathrm{AD}\,\overline{W_{l+1}}(\overline{F_{p-l-1}}) \tag{57}$$

and with (55) we have $\overline{W_{l+1}} = \overline{S(F_{l+1})}$. Assuming continuity of the maps S and $\bar{\ }$ one can formally show that $\overline{S(B)} = S(\overline{B})$ for any operator B for which $S(B)$ and $\overline{B}$ exist. On the other hand it is evident that $S(\overline{B}) = 0$. Putting these together shows that

$$\overline{\mathrm{AD}\,W_{l+1}(\overline{F_{p-l-1}})} = 0 \tag{58}$$

such that the expression (56) for K_p does not contain contributons arising from averaging the terms $\mathrm{AD}\,W_{l+1}(K_{p-l-1})$ in the expression (45) for the F_p.

Let us summarize what we have done so far: Given a perturbed Hamiltonian $H(\epsilon) = \sum_{p=0}^{\infty}\frac{\epsilon^p}{p!}H_p$ we have shown that choosing $W_p = S(F_p)$ in $W(\epsilon) = \sum_{p=0}^{\infty}\frac{\epsilon^p}{p!}W_{p+1}$ leads to

$$K(\epsilon) = H_0 + \sum_{p=1}^{\infty}\frac{\epsilon^p}{p!}\overline{F_p} = \Phi(\epsilon)^{-1}\,H(\epsilon)\,\Phi(\epsilon) \tag{59}$$

and the $\overline{F_p}$ all commute with the unperturbed operator H_0 and are given by

$$\overline{F_1} = \overline{H_1} \tag{60}$$

$$\overline{F_p} = \overline{H_p} + \sum_{l=0}^{p-2}\binom{p-1}{l}\overline{T_{p-l-1}H_{l+1}} \qquad p \geqslant 2\,. \tag{61}$$

The first few terms in the expansion (59) are

$$\begin{aligned}
F_0 &= 0 & K_0 &= H_0 \\
F_1 &= H_1 & K_1 &= \overline{H_1} \\
F_2 &= H_2 + \tfrac{\mathrm{i}}{\hbar}[W_1, K_1 + H_1] & K_2 &= \overline{H_2} + \tfrac{\mathrm{i}}{\hbar}\overline{[W_1, H_1]}
\end{aligned} \tag{62}$$

where $W_1 = S(H_1)$. Hence, $\sum_{p=1}^{N} \frac{\epsilon^p}{p!}\overline{F_p}$ and H_0 can be simultaneously diagonalized for any finite N. Let us introduce the following notation:

$$K^N(\epsilon) = \sum_{p=0}^{N} \frac{\epsilon^p}{p!}\overline{F_p} \tag{63}$$

$$\Phi^N(\epsilon) = \sum_{p=0} \frac{\epsilon^p}{p!}\Phi_p \tag{64}$$

which implies

$$K^N(\epsilon) = (\Phi^N(\epsilon))^{-1} H(\epsilon)\, \Phi^N(\epsilon) + O(\epsilon^{N+1}). \tag{65}$$

Let

$$K^N(\epsilon)|j\rangle^N(\epsilon) = E_j^N(\epsilon)|j\rangle^N(\epsilon) \tag{66}$$

and

$$H(\epsilon)|j\rangle(\epsilon) = E_j(\epsilon)|j\rangle(\epsilon). \tag{67}$$

Evidently

$$H(\epsilon)\Phi^N(\epsilon)|j\rangle^N(\epsilon) = E_j^N(\epsilon)\Phi^N(\epsilon)|j\rangle^N(\epsilon) + O(\epsilon^{N+1}) \tag{68}$$

i.e. the eigenvalues $E_j(\epsilon)$ and eigenvectors $|j\rangle(\epsilon)$ of the perturbed Hamiltonian $H(\epsilon)$ are approximated as follows:

$$E_j(\epsilon) = E_j^N(\epsilon) + O(\epsilon^{N+1}) \tag{69}$$

$$|j\rangle(\epsilon) = \Phi^N(\epsilon)|j\rangle^N(\epsilon) + O(\epsilon^{N+1}). \tag{70}$$

Hence, we have used the quantum analogue of the averaging method to construct a quantum mechanical perturbation theory. Just as in the classical case, however, the solutions for W_p constructed here are not the only ones. One encounters the same non-uniqueness as given by (28) and (29) in the classical case.

4. Examples: discrete spectra

4.1. General second-order terms

In this section we will apply the theory developed in section 3 to the case of a Hamiltonian H_0 which is assumed to have a purely discrete spectrum with finite degeneracy

$$H_0 = \sum_{j;\,\alpha\in D_j} |j,\alpha\rangle E_j^0 \langle\alpha, j| \tag{71}$$

where the sum over α runs over $D_j := \{1, \ldots, d_j = \dim(\mathrm{Eig}(H_0, E_j^0))\}$. For any self-adjoint G one then obtains

$$\overline{G} = \sum_{j;\,\alpha,\beta\in D_j} |j,\alpha\rangle\langle\alpha, j|G|j,\beta\rangle\langle\beta, j| \tag{72}$$

$$S(G) = \frac{\hbar}{\mathrm{i}} \sum_{j\neq k;\,\alpha\in D_j,\beta\in D_k} |j,\alpha\rangle \frac{\langle\alpha, j|G|k,\beta\rangle}{E_j^0 - E_k^0}\langle\beta, k|. \tag{73}$$

Using (62), (72), and (73) one finds after straightforward calculations

$$K^2(\epsilon) = \sum_{j;\,\alpha,\beta\in D_j} |j,\alpha\rangle \left(E_j^0 + \epsilon\langle\alpha, j|H_1|j,\beta\rangle\right)\langle\beta, j|$$

$$+ \sum_{j;\,\alpha,\beta\in D_j} |j,\alpha\rangle\left(\epsilon^2\left\{\frac{\langle\alpha, j|H_2|j,\beta\rangle}{2} + \sum_{j\neq k;\,\gamma\in D_k} \frac{\langle\alpha, j|H_1|k,\gamma\rangle\langle\gamma, k|H_1|j,\beta\rangle}{E_j^0 - E_k^0}\right\}\right)\langle\beta, j|\,. \tag{74}$$

Consequently the eigenvalues $E^2_{j,\alpha}(\epsilon)$ of $K^2(\epsilon)$ are determined to be solutions of the secular equation of d_j-dimensional matrices:

$$\det\left(E^2\delta_{\alpha\beta} - \langle\alpha, j|K^2(\epsilon)|j,\beta\rangle\right) = 0 \tag{75}$$

which coincides with the usual Rayleigh–Schrödinger result. The corresponding eigenvectors of $K^2(\epsilon)$ are then

$$|j,\alpha\rangle^2(\epsilon) = \sum_{\beta\in D_j} c^{j(2)}_{\alpha\beta}(\epsilon)|j,\beta\rangle \tag{76}$$

and one has for the eigenvector $|j,\alpha\rangle(\epsilon)$ of $H(\epsilon)$:

$$|j,\alpha\rangle(\epsilon) = \left(\mathbb{1} - \epsilon\frac{\mathrm{i}}{\hbar}W_1 + \frac{\epsilon^2}{2}\left(\left(\frac{\mathrm{i}}{\hbar}W_1\right)^2 - \frac{\mathrm{i}}{\hbar}W_2\right)\right)|j,\alpha\rangle^2(\epsilon) + \mathrm{O}(\epsilon^3) \tag{77}$$

with

$$W_1 = \frac{\hbar}{\mathrm{i}}\sum_{j\neq k;\,\alpha\in D_j,\beta\in D_k} |j,\alpha\rangle\frac{\langle\alpha, j|H_1|k,\beta\rangle}{E_j^0 - E_k^0}\langle\beta, k| \tag{78}$$

and $W_2 = S(H_2 + \frac{\mathrm{i}}{\hbar}[W_1, \overline{H_1} + H_1])$ which we shall not write down here but which inserted into (78) yields the corrections to the eigenvctors to second order known from Rayleigh–Schrödinger perturbation theory in the case of a non-degenerate spectrum. In this case we see from (77) that the eigenvectors of $K^2(\epsilon)$ and H_0 coincide: $|j\rangle^2(\epsilon) = |j\rangle$. In fact, since

$$K^N(\epsilon) = K^{N-1}(\epsilon) + \frac{\epsilon^N}{N!}K_N \tag{79}$$

and $[H_0, K_N] = 0$, $\forall N$ we have by induction

$$|j\rangle^N(\epsilon) = |j\rangle \qquad \forall N \tag{80}$$

in the non-degenerate case.

The general equivalence of the quantum version of the Poincaré–von Zeipel perturbation theory to the standard Rayleigh–Schrödinger perturbation theory will be proven in section 5. The fact that the corrections to the eigenvalues in the Poincaré–von Zeipel theory are derived from an averaging procedure may, however, provide some computational advantage since it gives the sums over intermediate states so common to standard perturbation theory in closed form. For example, the second-order term in the non-degenerate case is given by

$$\tfrac{1}{2}\langle j|H_2|j\rangle + \sum_{j\neq k}\frac{|\langle j|H_1|k\rangle|^2}{E_j^0 - E_k^0} = \lim_{T\to\infty}\frac{1}{T}\int_0^T \mathrm{d}t\,\left(\tfrac{1}{2}H_2(t) + \tfrac{1}{\hbar}[W_1, H_1](t)\right)\,. \tag{81}$$

4.2. Example 1: anharmonic oscillator

For the non-degenerate case we shall illustrate the method in the example of the harmonic oscillator $H_0 = \frac{1}{2}(-\frac{d^2}{dx^2} + x^2)$ with cubic perturbation $H_1 = \frac{1}{4}x^4$, $H_p = 0$, $p \geqslant 2$ (anharmonic oscillator with $\hbar = 1$) where the quantum Poincaré–von Zeipel method will permit us to compute the sums over intermediate states and the corrections up to $O(\epsilon^2)$ without much effort. All calculations are straightforward if we use the operators

$$a := \frac{1}{\sqrt{2}}\left(\frac{d}{dx} + x\right) \qquad a^\dagger := \frac{1}{\sqrt{2}}\left(-\frac{d}{dx} + x\right) \tag{82}$$

for which one finds

$$a(t) = e^{it}a \qquad a^\dagger(t) = e^{-it}a^\dagger \tag{83}$$

such that

$$H_1(t) = \frac{1}{16}\left(e^{4it}a^4 + e^{-4it}(a^\dagger)^4 + 4e^{2it}aH_0a + 4e^{-2it}a^\dagger H_0 a^\dagger\right) + \frac{3}{8}(H_0)^2 + \frac{2}{32}$$

$$\overline{H_1} = \frac{3}{8}(H_0)^2 + \frac{2}{32}$$

$$W_1 = \frac{1}{64i}\left((a^\dagger)^4 - a^4 + 8a^\dagger H_0 a^\dagger - 8aH_0a\right)$$

$$W_1(t) = \frac{1}{64i}\left(e^{-4it}(a^\dagger)^4 - e^{4it}a^4 + 8e^{-2it}a^\dagger H_0 a^\dagger - 8e^{2it}aH_0a\right)$$

$$K_2 = \frac{1}{2^9}\left([(a^\dagger)^4, a^4] + 32[a^\dagger H_0 a^\dagger, aH_0a]\right)$$

$$E_j^2(\epsilon) = j + \frac{1}{2} + \epsilon\left(\frac{3}{8}(j^2 + j) + \frac{3}{16}\right) - \epsilon^2\left(\frac{17}{64}j^3 + \frac{51}{128}j^2 + \frac{59}{128}j + \frac{21}{128}\right).$$

This formula for the correction to the eigenvalues was first derived by Heisenberg [19] and is also reproduced by Kummer using his normal form approach [12].

4.3. Example 2: the Henon–Heiles system

To illustrate the method for the degenerate case we apply it to the two-dimensional Henon–Heiles system whose unperturbed Hamiltonian is (we again choose $\hbar = 1$)

$$H_0 = -\tfrac{1}{2}\left(\frac{d^2}{dx_1^2} + \frac{d^2}{dx_2^2}\right) + \frac{(x_1)^2 + (x_2)^2}{2} =: H_{01} + H_{02} =: N_1 + \tfrac{1}{2} + N_2 + \tfrac{1}{2} \tag{84}$$

where the N_j, $j = 1, 2$ are the number operators and the perturbation is

$$H_1(\alpha, \beta) := \alpha(x_1)^2 x_2 + \beta(x_2)^3 \qquad H_p = 0 \quad \text{if } p \geqslant 2. \tag{85}$$

It is only for convenience that we have chosen the 'degenerate' case (i.e. equal frequencies for the two one-dimensional oscillators in H_0). The method is completely oblivious to that distinction. In this example we treat α and β as one perturbation parameter in the sense that $\alpha = \epsilon\tilde{\alpha}$, $\beta = \epsilon\tilde{\beta}$ and ϵ is the single perturbation parameter which we set equal to one at the end of the calculation. As in the case of the anharmonic oscillator it is very convenient to use the operators

$$a_j := \frac{1}{\sqrt{2}}\left(\frac{d}{dx_j} + x_j\right) \qquad a_j^\dagger := \frac{1}{\sqrt{2}}\left(-\frac{d}{dx_j} + x_j\right) \tag{86}$$

which also evolve according to

$$a_j(t) = \mathrm{e}^{\mathrm{i}t} a_j \qquad a_j^\dagger(t) = \mathrm{e}^{-\mathrm{i}t} a_j^\dagger . \tag{87}$$

This yields

$$\begin{aligned} H_1(t) = \frac{1}{2\sqrt{2}} \Big\{ & \mathrm{e}^{3\mathrm{i}t} \left(\alpha (a_1)^2 a_2 + \beta (a_2)^3\right) + \mathrm{e}^{-3\mathrm{i}t} \left(\alpha (a_1^\dagger)^2 a_2^\dagger + \beta (a_2^\dagger)^3\right) \\ & + \mathrm{e}^{\mathrm{i}t} \left(2\alpha H_{01} a_2 + \alpha (a_1)^2 a_2^\dagger + 2\beta H_{02} a_2 + \beta (a_2)^2 a_2^\dagger\right) \\ & + \mathrm{e}^{-\mathrm{i}t} \left(\alpha (a_1^\dagger)^2 + 2\alpha H_{01} a_2^\dagger + \beta (a_2^\dagger)^2 a_2 + 2\beta H_{02} a_2^\dagger\right) \Big\} \end{aligned} \tag{88}$$

$$\overline{H_1} = 0 \tag{89}$$

and

$$\begin{aligned} W_1 = \frac{1}{2\sqrt{2}\mathrm{i}} \Big\{ & \tfrac{1}{3} \left(\alpha (a_1^\dagger)^2 a_2^\dagger + \beta (a_2^\dagger)^3\right) - \tfrac{1}{3} \left(\alpha (a_1)^2 a_2 + \beta (a_2)^3\right) \\ & + \alpha (a_1^\dagger)^2 + 2\alpha H_{01} a_2^\dagger + \beta (a_2^\dagger)^2 a_2 + 2\beta H_{02} a_2^\dagger \\ & - \left(2\alpha H_{01} a_2 + \alpha (a_1)^2 a_2^\dagger + 2\beta H_{02} a_2 + \beta (a_2)^2 a_2^\dagger\right) \Big\} . \end{aligned} \tag{90}$$

From this one can read off $W_1(t)$ and obtain after some tedious but straightforward calculations (which we have executed with the help of the symbolic computation language MAPLE)

$$\begin{aligned} K_2(\alpha, \beta) = & -\left(\frac{4 N_1 N_2}{3} + \frac{5}{12} + \frac{5 N_1^2}{6} + (a_1)^2 (a_2^\dagger)^2 + (a_2)^2 (a_1^\dagger)^2 + \frac{2 N_2}{3} + \frac{3 N_1}{2}\right) \alpha^2 \\ & - \left(\frac{15 N_2}{2} + \frac{11}{4} + \frac{15 N_2^2}{2}\right) \beta^2 \\ & - \left(\frac{3}{2} + 6 N_1 N_2 - \frac{(a_1)^2 (a_2^\dagger)^2 + (a_2)^2 (a_1^\dagger)^2}{2} + 3 N_2 + 3 N_1\right) \alpha\beta . \end{aligned} \tag{91}$$

Keeping in mind that we set $\epsilon = 1$ and that $K_1 = \overline{H_1} = 0$, we now have to find the eigenvalues $E^2_{(k,\kappa)}(\alpha, \beta)$ of $K^2(\alpha, \beta) = H_0 + \frac{1}{2} K_2$ which will give us the correct eigenvalues of $H_0 + H_1$ up to second order. Let $k = 0$ denote the ground state (no degeneracy: $(0, 1)$), $k = 1$ the first excited state (double degeneracy: $(1, 1)$, $(1, 2)$), and $k = 3$ the second excited state (triple degeneracy: $(2, 1)$, $(2, \pm)$) then we find the following eigenvalues:

$$\begin{aligned} E^2_{(0,1)}(\alpha, \beta) &= 1 - \frac{11\beta^2}{8} - \frac{5\alpha^2}{24} - \frac{3\beta\alpha}{4} \\ E^2_{(1,1)}(\alpha, \beta) &= 2 - \frac{11\beta^2}{8} - \frac{11\alpha^2}{8} - \frac{9\beta\alpha}{4} \\ E^2_{(1,2)}(\alpha, \beta) &= 2 - \frac{71\beta^2}{8} - \frac{13\alpha^2}{24} - \frac{9\beta\alpha}{4} \\ E^2_{(2,1)}(\alpha, \beta) &= 3 - \frac{71\beta^2}{8} - \frac{19\alpha^2}{8} - \frac{27\beta\alpha}{4} \\ E^2_{(2,\pm)}(\alpha, \beta) &= 3 - \frac{101\beta^2}{8} - \frac{15\beta\alpha}{4} - \frac{17\alpha^2}{8} \pm \frac{\sqrt{2025\beta^4 - 446\beta^2\alpha^2 - 16\alpha^3\beta + 41\alpha^4}}{4} . \end{aligned} \tag{92}$$

The results for E^2 agree with those obtained by Kummer [13] and Ali [14] (except for the factor of α^2 in $E^2_{(2,1)}$ in [14], which is probably due to a typographical error).

5. Equivalence to Rayleigh–Schrödinger perturbation theory

5.1. Non-degenerate case

In section 4 we have already seen that at least up to second order the quantum analogue of the Poincaré–von Zeipel method and the standard Rayleigh–Schrödinger perturbation theory coincide. In this section we show that this is indeed true for the full perturbation expansions. To do this we recall briefly how the standard Rayleigh–Schrödinger expansion is constructed in the non-degenerate case. With the help of a suitably chosen contour integral in the complex E plane one can show that the projector

$$P_j(\epsilon) = \frac{1}{2\pi \mathrm{i}} \oint_{|E-E_j^0|=r} \frac{\mathrm{d}E}{E - H(\epsilon)} \tag{93}$$

on the jth eigenspace of $H(\epsilon)$ is analytic in ϵ and that for ϵ sufficiently small $\langle j|P_j(\epsilon)|j\rangle > 0$, [6, 10]. This projector then gives a normalized eigenvector $|j\rangle(\epsilon)$ of $H(\epsilon)$ to the eigenvalue $E_j(\epsilon)$ via

$$|j\rangle(\epsilon) = \frac{P_j(\epsilon)|j\rangle}{\sqrt{\langle j|P_j(\epsilon)|j\rangle}}. \tag{94}$$

From this one obtains

$$E_j(\epsilon) = (\epsilon)\langle j|H(\epsilon)|j\rangle(\epsilon) = \frac{\langle j|P_j(\epsilon)H(\epsilon)P_j(\epsilon)|j\rangle}{\langle j|P_j(\epsilon)|j\rangle}. \tag{95}$$

Using the expansion for $P_j(\epsilon)$ the right-hand side of (96) then yields an expansion for $E_j(\epsilon)$ in ϵ which is the usual Rayleigh–Schrödinger perturbation series.

On the other hand it follows from (81) that

$$K(\epsilon)|j\rangle = E_j(\epsilon)|j\rangle \tag{96}$$

where $K(\epsilon) = \Phi(\epsilon)^{-1} H(\epsilon)\, \Phi(\epsilon)$ which implies

$$|j\rangle(\epsilon) = \Phi(\epsilon)|j\rangle \tag{97}$$

such that

$$P_j(\epsilon) = |j\rangle(\epsilon)(\epsilon)\langle j| = \Phi(\epsilon))|j\rangle\langle j|\Phi(\epsilon)^{-1}. \tag{98}$$

Inserting (99) into (96) yields

$$E_j(\epsilon) = \frac{\langle j|P_j(\epsilon)H(\epsilon)P_j(\epsilon)|j\rangle}{\langle j|P_j(\epsilon)|j\rangle} = \langle j|\Phi(\epsilon)^{-1} H(\epsilon)\, \Phi(\epsilon) j\rangle$$

$$= \langle j|K(\epsilon)|j\rangle = \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} \langle j|K_p|j\rangle \tag{99}$$

which proves the equivalence in all orders for the non-degenerate case.

5.2. Degenerate case

Suppose now that $H_0|j,\alpha\rangle = E_j^0|j,\alpha\rangle$ with possible degeneracies $\alpha \in D_j = \{1,\ldots,d_j\}$, that E_j^0 is an isolated point of the spectrum $\sigma(H_0)$, that $P_j = \sum_{\alpha\in D_j} |j,\alpha\rangle\langle\alpha,j|$ is the projector on the jth eigenspace of H_0 and that $P_j(\epsilon)$ defined as in (94) exists and is analytic in ϵ. Then it has been shown that [6, 10]

$$\sigma\left(H(\epsilon)\,|_{\mathrm{Ran}\,P_j(\epsilon)}\right) = \sigma(H(\epsilon)) \cap \{E \mid |E - E_j^0| < r\} \tag{100}$$

(where $A\,|_{\mathrm{Ran}\,B}$ means restriction of the operator A to the range of B) and that there exists a unitary operator $U(\epsilon)$ such that

$$P_j(\epsilon) = U(\epsilon) P_j U(\epsilon)^{-1} \tag{101}$$

and

$$\tilde{H}(\epsilon) := U(\epsilon)^{-1} H(\epsilon) U(\epsilon) \tag{102}$$

satisfies

$$\tilde{H}(\epsilon) P_j = P_j \tilde{H}(\epsilon)\,. \tag{103}$$

Then $P_j(\tilde{H}(\epsilon) - E)P_j$ is a finite-dimensional operator analytic in ϵ and the eigenvalues $E_{j,\alpha}(\epsilon)$ of $H(\epsilon)$ are found as the d_j roots of the equation

$$\det\left(P_j(\tilde{H}(\epsilon) - E)P_j\right) = 0\,. \tag{104}$$

The operators $L_j(\epsilon) := P_j U(\epsilon)^{-1}$ and $R_j(\epsilon) := U(\epsilon)P_j$ sandwiching $H(\epsilon)$ in (105) satisfy certain differential equations involving $P_j(\epsilon)$ whose expansion in terms of ϵ is known from (94). With the help of these differential equations $L_j(\epsilon)$ and $R_j(\epsilon)$ can be expanded in ϵ as well. The Nth-order approximation in the Rayleigh–Schrödinger series for the degenerate case is then obtained by solving (105) where terms of order higher than N are neglected. We refer the reader to [6] for more details on Kato's rigorous exposition of the usual quantum mechanical Rayleigh–Schrödinger perturbation theory.

As was pointed out in [6] the unitary transformation $U(\epsilon)$ exists but is not necessarily unique. As we shall see below this non-uniqueness is equivalent to the non-uniqueness of the choice of W_p mentioned at the end of section 3.

In order to prove equivalence with the quantum Poincaré–von Zeipel method we first note that

$$\Phi(\epsilon)^{-1} U(\epsilon) P_j = P_j \Phi(\epsilon)^{-1} U(\epsilon)\,. \tag{105}$$

In fact, since $[K(\epsilon), H_0] = 0$ we may write the orthonormalized eigenvectors of $K(\epsilon)$ as $\sum_{\beta\in D_j} u^j(\epsilon)_{\alpha\beta}|j,\beta\rangle$ where $u^j(\epsilon)$ is a d_j-dimensional unitary matrix. Hence, the projector on the space $\bigoplus_{\alpha\in D_j} \mathrm{Eig}(K(\epsilon), E_{j,\alpha}(\epsilon))$ is

$$\sum_{\alpha,\beta,\gamma\in D_j} u^j(\epsilon)_{\alpha\beta}|j,\beta\rangle\langle\gamma,j|\left(u^j(\epsilon)\right)^\dagger_{\gamma\alpha} = \sum_{\beta\in D_j} |j,\beta\rangle\langle\beta,j| = P_j\,. \tag{106}$$

Since $K(\epsilon) = \Phi(\epsilon)^{-1} H(\epsilon)\Phi(\epsilon)$, one has

$$P_j(\epsilon) = \Phi(\epsilon)\, P_j\, \Phi(\epsilon)^{-1} \tag{107}$$

which together with (102) proves (106). Equation (106) states that

$$Z(\epsilon) := \Phi(\epsilon)^{-1} U(\epsilon) \tag{108}$$

is a unitary transformation commuting with all P_j. On the other hand, from

$$\Phi(\epsilon)\, K(\epsilon)\, \Phi(\epsilon)^{-1} = H(\epsilon) = U(\epsilon)\, \tilde{H}(\epsilon)\, U(\epsilon)^{-1} \tag{109}$$

it follows that

$$\tilde{H}(\epsilon) = Z(\epsilon)^{-1} K(\epsilon) Z(\epsilon) \tag{110}$$

which implies that the roots of equation (105) are identical to the roots of

$$\det\left(P_j(K(\epsilon) - E)P_j\right) = 0 \tag{111}$$

and this proves the equivalence of the two methods for the eigenvalue-expansions in the presence of degenerate eigenvalues.

Moreover, as can be seen from (111) the eigenvectors of $\tilde{H}(\epsilon)$ are related to those of $K(\epsilon)$ by the unitary transformation $Z(\epsilon)$ which preserves each eigenspace of H_0. It is likely that the non-uniqueness in the choice of W_p may be exploited to make $Z(\epsilon)$ trivial [11] but this is still under investigation.

6. Discussion and conclusion

Kummer [12] was the first to discuss the averaging method for quantum systems. Based on ideas from classical averaging he constructed a perturbation method called the normal form approach [13] which is equivalent to 'time averaging' [12] but instead of using averaging to solve (47) and (48) it employs algebraic constructions.

Motivated by the Birkhoff–Gustavson normal form in classical mechanics Ali [14] developed a quantum analogue of this and his construction yields the same expansion as that of Kummer. Working explicitly with an algebra of destruction and creation operators Eckhardt [15] also constructed a quantum analogue of the Birkhoff–Gustavson normal form. A quantization of the classical Birkhoff–Gustavson normal form was attempted by Robnik [20] but this is necessarily plagued by ordering problems which do not affect our work and the other contributions cited above (this is only partly true for [15]).

The constructions of Kummer, Ali and Eckhardt have in common that the existence of the generators of the unitary transformation has to be assumed or assured by certain additional conditions. In the present paper the necessary generators W_p are (at least formally) explicitly constructed.

In fact, it can be shown [11] that the method of quantum averaging as presented here provides explicit solutions in terms of the time averaging integrals for the algebraic constructions of Kummer. The algebraic constructions have the advantage of rigorous validity but lack constructive procedures needed to execute the algorithm. The approach presented here has, apart from its conceptual proximity to the classical situation, the advantage of providing explicit constructions. Due to the analytic character of these constructions, however, technical problems which are absent in the algebraic approach may arise.

Using a slightly modified quantum averaging (in our sense) for the particular example of the quantum anharmonic oscillator, Ben Lemlih and Ellison [16] have derived rigorous error bounds on approximations to the quantum time evolution.

Their work also contains a suggestion to compare the approximation to the eigenvalues of this specific problem to the usual perturbative corrections, i.e. Rayleigh–Schrödinger theory. Ali [14] and and Kummer [13] have found that in all the examples they have treated the normal form perturbative results agree with Rayleigh–Schrödinger theory (incidentally this is not true for Robnik's expansions [20]) and Eckhardt also suggests that the Birkhoff–Gustavson perturbation expansion is identical to the usual Rayleigh–Schrödinger pertubation theory. Our work then provides an explicit proof of this assertion since, as Kummer has shown, his normal form approach is equivalent to the averaging method in quantum

mechanics, and we have shown that averaging is completely equivalent to the Rayleigh–Schrödinger theory yielding the sums over intermediate states in closed forms.

A very important aspect related to this work (but not discussed here) is the fact that just as in classical mechanics a superconvergent perturbation theory can be constructed with the help of averaging, this can be done in quantum mechanics as well and yields a perturbation theory explicitly distinct from the usual Rayleigh–Schrödinger theory [9, 21].

Work is in progress to establish the technical conditions necessary to put quantum averaging on a rigorous mathematical footing and to determine how the non-uniqueness of the W_p may be used to trivialize $Z(\epsilon)$. It may also be possible that standard time-dependent perturbation techniques (e.g. sudden approximation) can be formulated as analogues of classical time-dependent averaging.

Acknowledgments

I am greatly indebted to M Kummer for a critical reading of the first version of this paper and very instructive remarks in which he pointed out some errors in the original version. Helpful discussions with H D Doebner and D Mayer are also gratefully acknowledged.

References

[1] Giacaglia G E O 1972 *Perturbation Methods in Non-Linear Systems* (Berlin: Springer)
[2] Lindstedt A 1882 *Abh. K. Akad. Wiss. St. Petersburg* **31** No 4
[3] Poincaré H *Les Méthodes Nouvelles de la Mécanique Célèste I, II, III* [Reprinted 1957 (New York: Dover)]
[4] von Zeipel H 1916–17 *Ark. Astron. Mat. Phys.* **11–13**
[5] Schrödinger E 1926 *Ann. Phys.* **80** 437
[6] Kato T 1966 *Perturbation Theory for Linear Operators* (Berlin: Springer)
[7] Rellich F 1937 *Math. Ann.* **113** 600, 677; 1939 *Math. Ann.* **116** 555; 1940 *Math. Ann.* **117** 356; 1942 *Math. Ann.* **118** 462
[8] Kolmogorov A N 1954 *Dokl. Akad. Nauk SSSR* **98** 527
[9] Scherer W 1994 Superconvergent perturbation method in quantum mechanics *TU-Clausthal preprint* ASI-TPA/17/94
[10] Reed M and Simon B 1978 *Methods of Modern Mathematical Physics IV: Analysis of Operators* (New York: Academic)
[11] Kummer M private communication
[12] Kummer M 1971 *Nuovo Cim.* B **1** 123
[13] Kummer M 1994 Anharmonic oscillators in classical and quantum mechanics with applications to the perturbed Kepler problem *Proc. Workshop on Conservative Systems and Quantum Chaos (Waterloo, Ont., 1992)* in press
[14] Ali M K 1985 *J. Math. Phys.* **26** 2565
[15] Eckhardt B 1986 *J. Phys. A: Math. Gen.* **19** 2961
[16] Ben Lemlih A and Ellison J A, 1985 *Phys. Rev. Lett.* **55** 1950
[17] Lichtenberg A J and Liebermann M A 1983 *Regular and Stochastic Motion* (Berlin: Springer)
[18] Moser J 1970 *Commun. Pure Appl. Math.* **23** 609
[19] Heisenberg W 1925 *Z. Phys.* **33** 879
[20] Robnik M 1984 *J. Phys. A: Math. Gen.* **17** 109
[21] Scherer W 1994 Quantum averaging II: superconvergent method *TU-Clausthal preprint* ASI-TPA/18/94 (in preparation)