Nekhoroshev type estimates for quantum propagators

1990 J. Phys. A: Math. Gen. 23 2915

(http://iopscience.iop.org/0305-4470/23/13/028)

J. Phys. A: Math. Gen. **23** (1990) 2915-2926. Printed in the UK

Nekhoroshev type estimates for quantum propagators

Jan Herczyński

Dipartimento di Matematica, Università di Bologna, Piazza di Porta S. Donato, 5, 40 127 Bologna, Italy, on leave from (permanent address): Instytut Matematyki Stosowanej i Mechaniki, Uniwersytet Warszawski, PKiN, IXp., 00 901 Warszawa, Poland

Received 10 August 1989

Abstract. We consider the Schrödinger operator for polynomially perturbed d-dimensional non-resonant harmonic oscillators. We adapt to quantum mechanics the argument of Giorgilli and Galgani, based on the Lie perturbation method and leading to Nekhoroshev type estimates. As a consequence we show how to use the Rayleigh-Schrödinger series to describe the quantum propagator for exponentially large times.

1. Introduction and statement of results

Classical perturbation theory can boast two major results, namely the KAM theorem and the Nekhoroshev theorem. The former gives estimates of (eternally) almost periodic motions, the existence of which is proved for a high fraction of the phase space. The latter gives estimates for motions occurring in all of phase space for finite times, exponentially large as the perturbation parameter ε decreases to zero. The theorem of Nekhoroshev establishes a third time scale for the perturbed motion, intermediate between the usual time scale $O(\varepsilon^{-1})$ and 'long' times (or 'eternity'). Some applications of Nekhoroshev estimates in classical mechanics are discussed in [1-3].

The aim of the present paper is to study this intermediate time scale for quantum propagators. We consider the restricted but physically significant example of polynomially perturbed harmonic oscillators with d degrees of freedom and non-resonant frequencies. We study in $L^2(\mathbb{R}^d)$ the Schrödinger operator

$$H_\varepsilon = H_0 + \varepsilon V = \sum_{j=1}^{d} \left(-\frac{\hbar^2}{2} \frac{\partial^2}{\partial q_j^2} + \omega_j q_j^2 \right) - \tfrac{1}{2}\hbar|\omega| + \varepsilon V(q) \tag{1}$$

where the potential V is a real, bounded-below polynomial of degree k and $\omega = (\omega_1, \ldots, \omega_d) \in \mathbb{R}_+^d$ satisfies the condition

$$|\omega \cdot \nu|^{-1} \leq C_1 |\nu|^{\gamma} \tag{2}$$

for any $\nu \neq 0$, $\nu \in \mathbb{Z}^d$, with $\omega \cdot \nu = \sum_{j=1}^{d} \omega_j \nu_j$, $|\nu| = \sum_{j=1}^{d} |\nu_j|$, C_1, γ positive. The shift by $\frac{1}{2}\hbar|\omega|$ in (1) is for reasons of simplicity, it gives $\inf \sigma(H_0) = 0$. The eigenvalues of H_0 are all simple, equal to $E_0(\nu) = \hbar\omega \cdot \nu$, $\nu \in \mathbb{N}^d$, where we let $0 \in \mathbb{N}$. We denote the corresponding normalised eigenfunctions e_ν, and let P_E be the projection onto the finite-dimensional span$\{e_\nu : E_0(\nu) \leq E\}$. As is well known [4], for each $\nu \in \mathbb{N}^d$ there exist a Rayleigh-Schrödinger perturbation series $\sum_{j=0}^{\infty} \varepsilon^j E_j(\nu)$, divergent but Borel summable to the perturbed eigenvalue $E_\varepsilon(\nu)$ of H_ε. Let K_j be the self-adjoint operator defined by $K_j e_\nu = E_j(\nu) e_\nu$. The following is the main result of this paper.

Theorem 1. Let $\alpha=(\gamma+d+2+k/2)^{-1}$, fix $E\geq 1$ and assume $\hbar<E(k\sup_{j=1,\ldots,d}\omega_j)^{-1}$. There exist positive constants A, B, ε_* such that for any $0<\varepsilon<B\varepsilon_*$ there exist a unitary operator U_ε and a self-adjoint operator K_ε of the form $K_\varepsilon=\Sigma_{j=0}^{r(\varepsilon)}\varepsilon^j K_j$ for appropriate integer $r(\varepsilon)$ satisfying

$$\|(U_\varepsilon H_\varepsilon U_\varepsilon^* - K_\varepsilon)P_E\| \leq AE^{k/2}\exp[-(\varepsilon_*/\varepsilon)^\alpha]. \tag{3}$$

Here $r(\varepsilon)$ is of the order of $(\varepsilon_*/\varepsilon)^\alpha$, $\varepsilon_*=D\hbar E^{-k/2}$, and A, B and D are independent of E and $\hbar$.

One should perhaps remark here that defining the unitary operator $\tilde{U}_\varepsilon$ and the self-adjoint operator $\tilde{K}_\varepsilon$ by

$$\tilde{U}_\varepsilon e_{\varepsilon,\nu}=e_\nu \qquad \tilde{K}_\varepsilon e_\nu = E_\varepsilon(\nu)e_\nu$$

where $\{e_{\varepsilon,\nu}\}_{\nu\in\mathbb{N}^d}$ are the orthogonal eigenfunctions of H_ε associated with $E_\varepsilon(\nu)$, $(H_\varepsilon - E_\varepsilon(\nu))e_{\varepsilon,\nu}=0$, we obtain immediately

$$\tilde{U}_\varepsilon H_\varepsilon \tilde{U}_\varepsilon^* = \tilde{K}_\varepsilon. \tag{4}$$

Note, however, that usually one knows neither $e_{\varepsilon,\nu}$ nor $E_\varepsilon(\nu)$ except through their divergent Rayleigh–Schrödinger series. In other words, up to rth order of perturbation theory we can only achieve an approximation of (4) with error $O(\varepsilon^{r+1})$. Since, however, the operators $\tilde{U}_\varepsilon$, $\tilde{K}_\varepsilon$ are defined abstractly we have no control over the constants in these estimates and so cannot put $r=r(\varepsilon)=\varepsilon^{-\alpha}$ to obtain the exponential estimate (3). The proof of (3) consists, in fact, of a concrete iterative procedure with explicit estimates of the remainder.

The result of theorem 1 is useful for perturbation theory, i.e. for taking the limit $\varepsilon\to 0$ for $\hbar$ fixed. Unfortunately, the dependence of ε_* on $\hbar$ excludes the semiclassical limit $\hbar\to 0$ for ε fixed. We also note that in the classical case Giorgilli and Galgani [5, 6] obtain a better exponent α, essentially $\alpha=(\gamma+2)^{-1}$, independent of V.

Theorem 1 has a consequence for propagators, which seems to be the first result establishing the intermediate time scale in the framework of quantum perturbation theory.

Theorem 2. With constants as in theorem 1, and A_1 independent of E and $\hbar$, we have

$$\|(\exp(-\mathrm{i}tH_\varepsilon)-\exp(-\mathrm{i}tK_\varepsilon))P_E\| \leq A_1\hbar^{-\alpha}E^{k/2}\varepsilon^\alpha$$

for $|t|\leq\varepsilon^\alpha\exp[(\varepsilon_*/\varepsilon)^\alpha]$.

The proof of theorem 1 is an adaptation to quantum perturbation theory of a proof of Nekhoroshev theorem given by Giorgilli and Galgani [5], see also [7]. Note that the rigorous implementation of classical perturbation algorithms in quantum mechanics was initiated by Graffi and Paul [8], where a WKB-type ansatz in the Bargmann representation reduced the Schrödinger equation to a Hamilton–Jacobi equation with quantum corrections and allowed the calculation of the semiclassical limit for all terms of Rayleigh–Schrödinger series to all orders in $\hbar$. These results were re-established by Degli Esposti, Graffi and Herczyński [9] by implementing the perturbation theory based on the Lie method, where no generating functions in mixed variables and no Hamilton–Jacobi equations appear. Since Giorgilli and Galgani use the Lie method in their proof, we will follow [9].

We observe that Ali [10] considered the Lie perturbation method in quantum mechanics. His results are formal in that he does not give the estimates of the operators involved, and thus cannot obtain, for instance, the unitary operator U_ε. He provides, however, a numerical test of the quantisation procedures. We point out also that a recent interesting application of Nekhoroshev like perturbation technique to the one-dimensional Schrödinger operator by Benettin *et al* [11] is based on rewriting the Schrödinger equation as a dynamical system (one-dimensionality is essential here) and then applying the Nekhoroshev approach in this classical situation. In contrast, we implement here the classical perturbation algorithm directly in the Hilbert space.

The plan of the paper is as follows. In section 2 we formulate the Lie perturbation algorithm in quantum mechanics and give the appropriate estimates. The proof of theorem 1 and theorem 2 is carried out in section 3. In the appendix we prove a lemma on polynomial perturbations necessary for our analysis.

2. The Lie method in quantum mechanics

Our first aim is to derive equations for perturbation theory of $H_\varepsilon = H_0 + \varepsilon V$, so we proceed formally and will consider the convergence problem later. Let $W_\varepsilon = \sum_{j=0}^{\infty} \varepsilon^j W_{j+1}$, where W_j are self-adjoint operators. We will consider the equation $U_\varepsilon H_\varepsilon U_\varepsilon^* = K_\varepsilon$, where U_ε is the unitary solution of $(\mathrm{d}/\mathrm{d}\varepsilon) U_\varepsilon = \mathrm{i}\, U_\varepsilon W_\varepsilon$, $U_0 = I$, so denote $T_\varepsilon A = U_\varepsilon A U_\varepsilon^*$ for any self-adjoint operator A and expand T_ε as a power series in ε: $T_\varepsilon A = \sum_{j=0}^{\infty} \varepsilon^j T_j A$, where $T_0 A = A$ (here W_ε, W_j are operators acting in the Hilbert space, and T_ε, T_j are operators acting on operators). We want to express T_j in terms of W_j. Differentiating $T_\varepsilon A = U_\varepsilon A U_\varepsilon^*$ with respect to ε we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon} T_\varepsilon A = U_\varepsilon \mathrm{i}[W_\varepsilon, A] U_\varepsilon^* \tag{5}$$

so letting $\tilde{L}_j A = \mathrm{i}[W_j, A]$ we obtain

$$n T_n A = \sum_{j=1}^{n} T_{n-j} \tilde{L}_j A$$

whence we can inductively find T_n in terms of commutators with W_j, similarly to the Lie method in classical mechanics as discussed, for instance, in Lieberman and Lichtenberg [12]. These expressions for T_n were used by Degli Esposti, Graffi and Herczyński [9]. It turns out, however, that for the purposes of perturbation theory a different formulation is more convenient. We rewrite (5) as

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon} T_\varepsilon A = \mathrm{i}[U_\varepsilon W_\varepsilon U_\varepsilon^*, T_\varepsilon A] = \mathrm{i}[X_\varepsilon, T_\varepsilon]$$

where $X_\varepsilon = U_\varepsilon W_\varepsilon U_\varepsilon^* = \sum_{j=0}^{\infty} \varepsilon^j X_{j+1}$, and letting $L_j A = \mathrm{i}[X_j, A]$ we obtain

$$T_n A = \frac{1}{n} \sum_{j=1}^{n} L_j T_{n-j} A \tag{6}$$

similar to expressions considered by Giorgilli and Galgani [5] (see also [6]).

All our estimates of U_ε will be expressed in terms of X_j, therefore to ensure that the operator U_ε is then well defined, we need the following result.

Lemma 1. If $\|X_j\| \leqslant \frac{1}{2}\beta^j$ for any j, then $\|W_j\| \leqslant \frac{1}{2}(2\beta)^j$ and W_ε is convergent for $|\varepsilon| < 1/2\beta$. Moreover, if X_j are self-adjoint, then so are W_j.

Proof. The proof uses the result of lemma 2 below, which is independent of lemma 1. Since $X_\varepsilon = U_\varepsilon W_\varepsilon U_\varepsilon = T_\varepsilon W_\varepsilon$, we have $W_1 = X_1$ and

$$W_n = X_n - \sum_{j=1}^{n-1} T_{n-j} W_j \tag{7}$$

for $n \geq 2$. Now we prove the lemma inductively. For $j = 1$ it is immediate. Suppose we have $\| W_j \| \leq \frac{1}{2}(2\beta)^j$ for $j = 1, \ldots, n-1$. Then by (7) and lemma 2 we find

$$\| W_n \| \leq \| X_n \| + \sum_{j=1}^{n-1} \| T_{n-j} W_j \| \leq \tfrac{1}{2}\beta^n + \tfrac{1}{2}\beta^n \sum_{j=1}^{n-1} 2^j \leq \tfrac{1}{2}(2\beta)^n.$$

The self-adjointness statement is obvious from (7), and thus the lemma is proved. □

We next want to consider the convergence of $\Sigma_{j=0}^{\infty} \varepsilon^j T_j$. Although X_j are bounded operators, we will have to consider expressions of the type $T_n H_0$, which are not bounded. We therefore introduce the following definitions. We say that an operator A is in the class F_s, $A \in F_s$, if A is relatively bounded with respect to some power of H_0 and, moreover, $(e_\nu, Ae_\mu) = 0$ whenever $|\nu - \mu| > s$, where $(\ , \ )$ denotes the scalar product in $L^2(\mathbb{R}^d)$ (as in section 1, we put $|\nu| = \Sigma_{j=1}^{d} |\nu_j|$). Moreover, we let $\|A\|_E$ denote $\|AP_E\|$, which is well defined since AP_E is, for any A, a finite-rank operator. Let $\omega^* = \sup_{j=1,\ldots,d} \omega_j$. The inequality

$$\|AB\|_E = \|AP_{E+sh\omega^*} BP_E\| \leq \|A\|_{E+sh\omega^*} \|B\|_E$$

holds for $B \in F_s$ and arbitrary A. If, furthermore, B is bounded, we obtain

$$\|[A, B]\|_E \leq 2\|A\|_{E+sh\omega^*} \|B\|. \tag{8}$$

Lemma 2. Suppose $X_j \in F_{jk}$, $\|X_j\| \leq \frac{1}{2}\beta^j$ for $j = 1, 2, \ldots$. Then

$$\| T_j A \|_E \leq \beta^j \|A\|_{E-j\mathscr{E}}$$

where $\mathscr{E} = \hbar k \omega^*$.

Proof. For $j = 0$ the result of the lemma is true. Suppose we have proved that $\| T_j A\|_E \leq \beta^j \|A\|_{E+j\mathscr{E}}$ for $j = 0, \ldots, n-1$. Then by (6) and (8)

$$\| T_n A\|_E \leq \frac{1}{n} \sum_{j=1}^{n} \|[X_j, T_{n-j} A]\|_E \leq \frac{2}{n} \sum_{j=1}^{n} \|X_j\| \| T_{n-j} A\|_{E+j\mathscr{E}} \leq \beta^n \|A\|_{E+n\mathscr{E}}.$$

The lemma is proved. □

Lemma 2 can be used to estimate the remainder of the series $\Sigma_{j=0}^{\infty} \varepsilon^j T_j A$.

Lemma 3. Suppose X_j are as in lemma 2. If A is bounded then $\Sigma_{j=0}^{\infty} \varepsilon^j T_j A$ is convergent for $|\varepsilon| \leq 1/\beta$ and

$$\|R_r(A)\| \leq (1 - \varepsilon\beta)^{-1} (\varepsilon\beta)^{r+1} \|A\|$$

where $R_r(A) = U_\varepsilon A U_\varepsilon^* - \Sigma_{j=0}^{r} \varepsilon^j T_j A$. If A is such that $\|A\|_E \leq CE^\kappa$, then $\Sigma_{j=0}^{\infty} \varepsilon^j T_j(A) P_E$ is convergent for $|\varepsilon| \leq 1/2\beta$, for any E, and

$$\|R_r(A)\|_E \leq CC_\kappa (E + \mathscr{E})^\kappa (1 - 2\varepsilon\beta)^{-1} (2\varepsilon\beta)^{r+1}$$

where C_κ depends only on κ.

Proof. The first part of the lemma follows by geometric series estimates, and the second from the inequality

$$C(E+j\mathscr{E})^{\kappa} \leqslant CC_{\kappa}(E+\mathscr{E})^{\kappa} 2^{j}. \qquad \square$$

With the above lemmas we can now proceed to write down the perturbation theory algorithm. Consider the equation

$$U_{\varepsilon}(H_0+\varepsilon V)U_{\varepsilon}^{*} = \sum_{j=0}^{\infty} \varepsilon^{j} K_{j}$$

where we want K_j to commute with H_0, i.e. to belong to F_0. Expanding in ε, we rewrite the above as $K_0 = H_0$ and

$$T_n H_0 + T_{n-1} V = K_n. \tag{9}$$

Using (6), we rewrite (9) as

$$\frac{\mathrm{i}}{n}[X_n, H_0] + V^{(n)} = K_n \tag{10}$$

where

$$V^{(n)} = T_{n-1} V + \frac{1}{n} \sum_{j=1}^{n-1} L_j T_{n-j} H_0. \tag{11}$$

Equation (10) is solved for X_n, K_n, assuming $X_1, \ldots, X_{n-1}$, and hence also $V^{(n)}$, known, as follows. We have $(e_\nu, [X_n, H_0]e_\nu) = 0$ for any $\nu \in \mathbb{N}^d$, and we assume $K_n \in F_0$. Therefore

$$(e_\nu, K_n e_\nu) = (e_\nu, V^{(n)} e_\nu) \tag{12}$$

and

$$(e_\nu, X_n e_\mu) = n\mathrm{i}\,\frac{(e_\nu, V^{(n)} e_\mu)}{\hbar\omega\cdot(\mu-\nu)} \tag{13}$$

for $\mu \neq \nu$. We put, by convention, $(e_\nu, X_n e_\nu) = 0$. We will be able to estimate the norm of X_n using the Diophantine condition (2) and the fact that $V^{(n)} \in F_{nk}$. But before proceeding we will follow Giorgilli and Galgani in deriving another expression for $V^{(n)}$, more useful than (11). We find, using (6) and (9), that

$$\begin{aligned}
V^{(n)} &= \frac{1}{n-1} \sum_{j=1}^{n-1} L_j T_{n-1-j} V + \frac{1}{n} \sum_{j=1}^{n-1} L_j T_{n-j} H_0 \\
&= \frac{1}{n} \frac{1}{n-1} \sum_{j=1}^{n-1} L_j (n T_{n-1-j} V + (n-1) T_{n-j} H_0) \\
&= \frac{1}{n} \frac{1}{n-1} \sum_{j=1}^{n-1} L_j (T_{n-1-j} V + (n-1) K_{n-j}) \\
&= \frac{1}{n} T_{n-1} V + \frac{1}{n} \sum_{j=1}^{n-1} L_j K_{n-j}.
\end{aligned} \tag{14}$$

From our point of view the main value of the formula (14) is that it allows one to estimate the norm of $V^{(n)}$ when V is bounded. Moreover, we note that

$$(e_\nu, K_n e_\nu) = \frac{1}{n}(e_\nu, T_{n-1} V e_\nu)$$

which may be more useful than (12), though we will not use it.

Equations (10), (12), (13) and (14) constitute the perturbation algorithm based on the Lie method implemented directly in the Hilbert space. Because of the uniqueness of the Rayleigh-Schrödinger series for simple eigenvalues, $(e_\nu, K_n e_\nu)$ will turn out to be $E_n(\nu)$. Note, however, that if V is a polynomial in q, then it is an unbounded operator and so will be X_1, K_1 and in general all X_j, and we will not be able to obtain the estimates needed for lemmas 3 and 4. Therefore we proceed in the following manner: instead of V we consider $\Phi = P_{E_0} V P_{E_0}$ for appropriately chosen E_0. This is a bounded operator. Thus our perturbation equations become

$$\frac{\mathrm{i}}{n}[X_n, H_0] + V^{(n)} = \tilde{K}_n \tag{15}$$

$$V^{(n)} = \frac{1}{n} T_{n-1}\Phi + \frac{1}{n}\sum_{j=1}^{n-1} L_j \tilde{K}_{n-j} \tag{16}$$

for $n \geqslant 2$, $V^{(1)} = \Phi$, where $(e_\nu, \tilde{K}_n e_\nu)$ is the nth term of the Rayleigh-Schrödinger series for $H_0 + \varepsilon\Phi$ for perturbation of $E_0(\nu)$. Equation (15) is solved as in (12), (13). Since now all operators X_j, $\tilde{K}_j$ are bounded, we may estimate the norm of $V^{(n)}$. This will depend on the norm of Φ and hence on the choice of E_0, to be discussed in the next section.

It is easy to see inductively that $V^{(n)}$, $X_n \in F_{kn}$, and also that $T_n A \in F_{s+nk}$ for $A \in F_s$. Observe also that since $V^{(1)}$ is self-adjoint, so are $V^{(n)}$ and X_n for any $n \geqslant 1$. Our next step is to provide estimates for the iterative procedure (15), (16).

Lemma 4. For any $j = 1, 2, \ldots$ we have

$$\|V^{(j)}\| \leqslant \|\Phi\| \mathscr{E}(j)^{j-1} \tag{17}$$

$$\|T_{j-1}\Phi\| \leqslant \|\Phi\| \mathscr{E}(j)^{j-1} \tag{18}$$

where $\mathscr{E}(1) = 1$, $\mathscr{E}(j) = (1/\hbar) 2^{d+1} C_1 k^{\gamma+d} \|\Phi\| j^{\gamma+d+1}$, for $j \geqslant 2$.

Proof. The proof is by induction. For $j = 1$, (16) and (17) are obvious, so suppose we have them for $j = 1, \ldots, n-1$. Note first that by (12)

$$\|\tilde{K}_j\| \leqslant \|V^{(j)}\| \leqslant \|\Phi\| \mathscr{E}(j)^{j-1} \tag{19}$$

and by (13), (2) and $V^{(j)} \in F_{jk}$ we have

$$\begin{aligned}
\|X_j\| &\leqslant \sup_{\|f\|=\|g\|=1} |(f, X_j g)| \\
&\leqslant \sup_{\|f\|=\|g\|=1} \sum_{\mu, |\nu| \leqslant jk} |(f, e_\nu)(e_\nu, X_j e_{\nu+\mu})(e_{\nu+\mu}, g)| \\
&\leqslant \sup_{\|f\|=\|g\|=1} \|V^{(j)}\| \frac{1}{\hbar} C_1 j (jk)^\gamma \sum_{\mu, |\nu| \leqslant jk} |(f, e_\nu)(e_{\nu+\mu}, g)| \\
&\leqslant \|V^{(j)}\| \frac{1}{\hbar} C_1 j (jk)^\gamma \sum_{|\nu| \leqslant jk} 1 \\
&\leqslant \|\Phi\| \mathscr{E}(j)^{j-1} \frac{1}{\hbar} C_1 j (jk)^\gamma (2jk)^d \\
&= \tfrac{1}{2} \mathscr{E}(j)^j.
\end{aligned} \tag{20}$$

We now consider (18) for $j = n$. We find, by (6) and (20), that

$$\begin{aligned}\|T_{n-1}\Phi\| &\leqslant \frac{1}{n-1}\sum_{j=1}^{n-1}\|L_j T_{n-1-j}\Phi\| \\ &\leqslant \frac{2}{n-1}\sum_{j=1}^{n-1}\|X_j\|\|T_{n-1-j}\Phi\| \\ &\leqslant \frac{1}{n-1}\sum_{j=1}^{n-1}\mathscr{E}(j)^j\|\Phi\|\mathscr{E}(n-j)^{n-1-j} \\ &\leqslant \|\Phi\|\mathscr{E}(n)^{n-1}.\end{aligned}$$

We pass to (17) for $j = n$ and use the elegant formula (14) to obtain

$$\begin{aligned}\|V^{(n)}\| &\leqslant \frac{1}{n}\|T_{n-1}\| + \frac{1}{n}\sum_{j=1}^{n-1}\|L_j K_{n-j}\| \\ &\leqslant \frac{1}{n}\|T_{n-1}\| + \frac{2}{n}\sum_{j=1}^{n-1}\|X_j\|\|K_{n-j}\| \\ &\leqslant \frac{1}{n}\|\Phi\|\mathscr{E}(n)^{n-1} + \frac{n-1}{n}\|\Phi\|\mathscr{E}(n)^{n-1} \\ &= \|\Phi\|\mathscr{E}(n)^{n-1}.\end{aligned}$$

The lemma is proved. □

For our purposes the estimate (20) is essential. It shows that if we carry out the perturbation theory to all orders we will not get the estimate $\|X_j\| \leqslant \frac{1}{2}\beta^j$, needed for lemma 2. We will, instead, again follow Giogilli and Galgani [5] by considering the perturbation theory up to order r, putting $\beta = \mathscr{E}(r)$, and then by choosing the best $r(\varepsilon)$ to minimise the remainder.

3. The proof of theorem 1

We carry out the perturbation theory as described in the previous section up to order r and set $X_j = 0$ for $j \geqslant r$. We then get $\|X_j\| \leqslant \frac{1}{2}\mathscr{E}(r)^j$ for any j and we can apply lemma 3 to estimate the remainder. The heart of the argument is the optimisation of this estimate for a given ε by choosing the right r. In contrast to Giorgilli and Galgani [5], however, we also have to be careful to make the right choice of E_0 in the definition of Φ. We put $E_0 = E + (r+1)\mathscr{E}$ where, we recall, $\mathscr{E} = \hbar k\omega^*$. The following two lemmas justify this choice.

Lemma 5. If $E_0 = E + (r+1)\mathscr{E}$, then

$$\tilde{K}_j P_E = K_j P_E \tag{21}$$

for $j = 1, \ldots, r$.

Lemma 6. Assume that $\varepsilon < 1/2\beta$ and $E > \mathscr{E}$. If E_0 is as above, then

$$\|U_\varepsilon(V-\Phi)U_\varepsilon^*\|_E \leqslant C_2 E^{k/2} r^{k/2}(2\varepsilon\beta)^r$$

where C_2 depends only on V and $\beta = \mathscr{E}(r)$.

Lemma 5 is essential if we want to have the Rayleigh–Schrödinger series for $H_0+\varepsilon V$ instead of $H_0+\varepsilon\Phi$ in theorem 1. Lemma 6 gives the estimate similar to that of lemma 3, which will be important in the optimisation step.

Proof of lemma 5. For the purposes of this proof only, we introduce the following notation: X_j, $V^{(j)}$, T_j for perturbation theory of $H_0+\varepsilon V$, $\tilde{X}_j$, $\tilde{V}^{(j)}$, $\tilde{T}_j$ for perturbation theory of $H_0+\varepsilon\Phi$, $V^{(1)}=V$, $\tilde{V}^{(1)}=\Phi$. We will show a stronger statement than (21), namely

$$\tilde{V}^{(j)}P_{E_0-j\mathscr{E}}=V^{(j)}P_{E_0-j\mathscr{E}} \tag{22}$$

for $j\leqslant r$, from which (21) follows immediately. We prove (22) by a kind of 'finite' induction for $j\leqslant r$, together with

$$\tilde{T}_{j-1}\Phi P_{E_0-j\mathscr{E}}=T_{j-1}VP_{E_0-j\mathscr{E}}\,. \tag{23}$$

For $j=1$ (22) and (23) are obvious. Suppose we have them for $j=1,\ldots,n<r$. Then also $\tilde{X}_jP_{E_0-j\mathscr{E}}=X_jP_{E_0-j\mathscr{E}}$ for $j\leqslant n$, and

$$\begin{aligned}\tilde{T}_n\Phi P_{E_0-n\mathscr{E}}&=\frac{\mathrm{i}}{n}\sum_{j=1}^{n}(\tilde{X}_j\tilde{T}_{n-j}\Phi-\tilde{T}_{n-j}\Phi\tilde{X}_j)P_{E_0-n\mathscr{E}}\\&=\frac{\mathrm{i}}{n}\sum_{j=1}^{n}(\tilde{X}_jP_{E_0-j\mathscr{E}}\tilde{T}_{n-j}\Phi-\tilde{T}_{n-j}\Phi P_{E_0-(n-j)\mathscr{E}}\tilde{X}_j)P_{E_0-n\mathscr{E}}\\&=\frac{\mathrm{i}}{n}\sum_{j=1}^{n}(X_jP_{E_0-j\mathscr{E}}T_{n-j}V-T_{n-j}VP_{E_0-(n-j)\mathscr{E}}X_j)P_{E_0-n\mathscr{E}}\\&=T_nVP_{E_0-n\mathscr{E}}\end{aligned}$$

and similarly we prove, using (14), that (22) holds for $j=n+1$. The proof is thus complete. □

Proof of lemma 6. For $\Omega\subset\mathbb{R}$ let P_Ω denote the orthogonal projection onto $\mathrm{span}\{e_\nu\colon E_0(\nu)\in\Omega\}$. Note that

$$\begin{aligned}V&=P_{E_0}VP_{E_0}+(I-P_{E_0})V+V(I-P_{E_0})\\&=\Phi+P_{(E_0,E_0+\mathscr{E})}VP_{(E_0-\mathscr{E},E_0]}+VP_{(E_0,\infty)}\end{aligned}$$

since $V\in F_k$, therefore putting $E_1=E_0-\mathscr{E}$, we have

$$\begin{aligned}&\|U_\varepsilon(V-\Phi)U_\varepsilon^*\|_E\\&\quad\leqslant 2\|U_\varepsilon VP_{(E_1,\infty)}U_\varepsilon^*\|_E\\&\quad\leqslant 2\sum_{j=1}^{\infty}\|U_\varepsilon VP_{(jE_1,(j+1)E_1]}U_\varepsilon^*\|_E\\&\quad\leqslant 2\sum_{j=1}^{\infty}\left(\sum_{t=0}^{t_j}\|T_t(VP_{(jE_1,(j+1)E_1]})\|_E+\|R_{t_j}(VP_{(jE_1,(j+1)E_1]})\|_E\right).\end{aligned}$$

We use now lemma 2 to note that we have $\|T_t(VP_{(jE_1,(j+1)E_1]})\|_E=0$ for $E+t\mathscr{E}<jE_1$, so we let $t_j=\mathrm{Int}((jE_1-E)/\mathscr{E})=\mathrm{Int}(jr+(j-1)E/\mathscr{E})$ and obtain, using lemma 3 and

(A1),

$$\|U_\varepsilon(V-\Phi)U_\varepsilon^*\|_E \leqslant 2\sum_{j=1}^{\infty}\|R_{t_j}(VP_{(jE_1,(j+1)E_1]})\|_E$$

$$\leqslant 2(1-\varepsilon\beta)^{-1}\sum_{j=1}^{\infty}(\varepsilon\beta)^{t_j+1}\|V\|_{(j+1)E_1}$$

$$\leqslant 2C_V E_1^{k/2}(1-\varepsilon\beta)^{-1}\sum_{j=1}^{\infty}(\varepsilon\beta)^{t_j+1}(j+1)^{k/2}.$$

Recall that $\varepsilon\beta<\frac{1}{2}$ and $E>\mathscr{E}$, and observe that $t_j+1>jr+(j-1)$ so letting C_κ be as in lemma 3, we get

$$\|U_\varepsilon(V-\Phi)U_\varepsilon^*\|_E \leqslant 4C_{k/2}C_V E_1^{k/2}\sum_{j=1}^{\infty}(2\varepsilon\beta)^{jr}$$

$$\leqslant 8C_{k/2}C_V E_1^{k/2}(2\varepsilon\beta)^r.$$

Note now that $E_1=E+r\mathscr{E}\leqslant 2Er$, and the result of lemma 6 follows. □

We finally come to the proof of theorem 1 and observe that

$$U_\varepsilon(H_0+\varepsilon V)U_\varepsilon^*-\sum_{j=0}^{r}\varepsilon^jK_j=\varepsilon U_\varepsilon(V-\Phi)U_\varepsilon^*+R_r(H_0)+\varepsilon R_{r-1}(V)+\sum_{j=0}^{r}\varepsilon^j(\tilde K_j-K_j).$$

Using lemma 3, lemma 5 and lemma 6 we find, for $r\geqslant k$,

$$\left\|U_\varepsilon(H_0+\varepsilon V)U_\varepsilon^*-\sum_{j=0}^{r}\varepsilon^jK_j\right\|_E \leqslant C_3E^{k/2}(2\varepsilon\beta r)^r \tag{24}$$

where, by lemma 4 and (A1),

$$\beta=\mathscr{E}(r)=\frac{2^{d+1}C_1k^{\gamma+d}}{\hbar}r^{\gamma+d+1}\|P_{E_0}VP_{E_0}\|$$

$$\leqslant\frac{2^{d+1}C_1C_Vk^{\gamma+d}}{\hbar}E^{k/2}r^{\gamma+d+1+k/2}. \tag{25}$$

Letting $\bar\varepsilon_*=\hbar(2^{d+2}C_1C_Vk^{\gamma+d}E^{k/2})^{-1}$, and $\tilde\gamma=\gamma+d+2+k/2$, we find

$$\left\|U_\varepsilon(H_0+\varepsilon V)U_\varepsilon^*-\sum_{j=0}^{r}\varepsilon^jK_j\right\|_E \leqslant C_3E^{k/2}\left(\frac{\varepsilon r^{\tilde\gamma}}{\bar\varepsilon_*}\right)^r \tag{26}$$

where neither C_3 nor $\bar\varepsilon_*$ depends on r. This means we are now free to choose r as we please, and it is easy to see that the best choice is $r(\varepsilon)=\mathrm{Int}(\bar r(\varepsilon))$, where

$$\bar r(\varepsilon)=\mathrm{e}^{-1}\left(\frac{\bar\varepsilon_*}{\varepsilon}\right)^{1/\tilde\gamma}.$$

We then have

$$\left(\frac{\varepsilon r(\varepsilon)^{\tilde\gamma}}{\bar\varepsilon_*}\right)^{r(\varepsilon)}\leqslant\left(\frac{\varepsilon\bar r(\varepsilon)^{\tilde\gamma}}{\bar\varepsilon_*}\right)^{\bar r(\varepsilon)-1}=\exp(\tilde\gamma)\exp(-\bar r(\varepsilon)\tilde\gamma).$$

Note now that $\bar r(\varepsilon)\tilde\gamma=(\varepsilon_*/\varepsilon)^{\alpha}$ for $\varepsilon_*=\bar\varepsilon_*(\tilde\gamma/e)^{\tilde\gamma}$, hence by (26) we obtain the estimate (3). We now check the consistency of conditions imposed on ε in the above argument.

We need $r \geqslant k$, used in (24), and $2\beta\varepsilon < 1$, used in lemma 1, lemma 3 and lemma 6. The first of these reduces to

$$e^{-1}\left(\frac{\bar{\varepsilon}_*}{\varepsilon}\right)^{\alpha} \geqslant k+1$$

and is satisfied by the choice $B = [(k+1)\tilde{\gamma}]^{-\tilde{\gamma}}$. The second, using $\beta = (2\bar{\varepsilon}_*)^{-1} r(\varepsilon)^{\gamma+d+1+k/2}$, follows from

$$2\beta\varepsilon = \frac{\varepsilon r(\varepsilon)^{\gamma+d+1+k/2}}{\bar{\varepsilon}_*} \leqslant \frac{\varepsilon \bar{r}(\varepsilon)^{\tilde{\gamma}}}{\bar{\varepsilon}_*} = \exp(-\tilde{\gamma}). \tag{27}$$

This completes the proof of theorem 1. □

Using theorem 1, we can now give a simple proof of theorem 2. The identity

$$\exp(-\mathrm{i}tU_\varepsilon H_\varepsilon U_\varepsilon^*) - \exp(-\mathrm{i}tK_\varepsilon)$$
$$= -\mathrm{i}\int_0^t \exp(-\mathrm{i}sU_\varepsilon H_\varepsilon U_\varepsilon^*)(U_\varepsilon H_\varepsilon U_\varepsilon^* - K_\varepsilon)\exp[-\mathrm{i}(t-s)K_\varepsilon]\,\mathrm{d}s$$

the fact that K_ε commutes with P_E, and (3) yield

$$\|[\exp(-\mathrm{i}tU_\varepsilon H_\varepsilon U_\varepsilon^*) - \exp(-\mathrm{i}tK_\varepsilon)]P_E\| \leqslant AE^{k/2}\varepsilon^\alpha \tag{28}$$

for $|t| \leqslant \varepsilon^\alpha \exp[(\varepsilon_*/\varepsilon)^\alpha]$. We use lemma 3 with $r=0$ to note that

$$\exp(-\mathrm{i}tU_\varepsilon H_\varepsilon U_\varepsilon^*) - \exp(-\mathrm{i}tH_\varepsilon) = T_\varepsilon[\exp(-\mathrm{i}tH_\varepsilon)] - \exp(-\mathrm{i}tH_\varepsilon)$$
$$= R_0[\exp(-\mathrm{i}tH_\varepsilon)]$$

can be estimated by

$$\|\exp(-\mathrm{i}tU_\varepsilon H_\varepsilon U_\varepsilon^*) - \exp(-\mathrm{i}tH_\varepsilon)\| \leqslant (1-\beta\varepsilon)^{-1}\beta\varepsilon \leqslant 2\beta\varepsilon \tag{29}$$

since $\varepsilon\beta \leqslant \frac{1}{2}$. Using (25) we give another estimate of $2\beta\varepsilon$, different from (27):

$$2\beta\varepsilon = \frac{\varepsilon}{\bar{\varepsilon}_*}\left(\frac{\bar{\varepsilon}_*}{\varepsilon}\right)^{1-\alpha} \exp[-(\gamma+d+1+k/2)] \leqslant \left(\frac{\varepsilon}{\bar{\varepsilon}_*}\right)^{\alpha}.$$

Putting (28), (29) and the above together, and using the definition of $\bar{\varepsilon}_*$, we obtain theorem 2. □

Acknowledgments

This work was performed during my stay at Dipartimento di Matematica, Università di Bologna. I am indebted to my Italian colleagues for their hospitality, in particular to Sandro Graffi for making my stay here possible and for many useful discussions. I thank Franco Nardini for his critical reading of the manuscript. The financial support of CNR in Firenze, Gruppo di Fisica Matematica, is most gratefully acknowledged.

Appendix

The aim of this appendix is to prove the following lemma.

Lemma A. Let V be a real polynomial of degree k. Then $V \in F_k$ and there exist $C_V > 0$ such that

$$\|V\|_E \leqslant C_V E^{k/2} \tag{A1}$$

for $E \geqslant 1$, $\hbar < 1$.

Proof. It is easiest to perform the proof in the Bargmann representation [8, 13], which was already used for the purposes of perturbation theory in [8, 9]. We consider the Hilbert space of analytic functions

$$\mathscr{F}_d = \left\{ f \text{ analytic in } \mathbb{C}^d, \int |f(x+iy)|^2 \exp(-(|x|^2+|y|^2)/\hbar)\, dx\, dy < \infty \right\}$$

in which the operator H_0 becomes

$$H_0 = \hbar\omega z \nabla_z = \hbar \sum_{j=1}^{d} \omega_j z_j \frac{\partial}{\partial z_j}$$

with eigenfunction $e_\nu(z) = (\hbar^{|\nu|}\nu!)^{-1/2} z^\nu$, $\nu \in \mathbb{N}^d$. Moreover, the multiplication by q_j becomes $(z_j + \hbar\partial/\partial z_j)/\sqrt{2\omega_j}$, see [8], so we see that V becomes an operator of the form

$$V = \sum_{|\alpha+\beta| \leqslant k} v_{\alpha,\beta} z^\alpha (\hbar\nabla_z)^\beta \tag{A2}$$

whence $V \in F_k$ is immediate. Moreover $v_{\alpha,\beta}$ in (A2) are polynomials in $\hbar$ of degree less than k. Note now that

$$z^\alpha e_\nu = \hbar^{|\alpha|/2} \left(\frac{(\alpha+\nu)!}{\nu!} \right)^{1/2} e_{\nu+\alpha} \qquad (\hbar\nabla_z)^\beta e_\nu = \hbar^{|\beta|/2} \left(\frac{\nu!}{(\nu-\beta)!} \right)^{1/2} e_{\nu-\beta}$$

for $\beta \leqslant \nu$. Letting $\omega_* = \min\{\omega_1, \dots, \omega_d, \frac{1}{2}\}$, we find

$$\|z^\alpha e_\nu\| \leqslant (\hbar|\nu| + \hbar k)^{|\alpha|/2} \leqslant \omega_*^{-|\alpha|/2}(k+1)^{|\alpha|/2} E_0(\nu)^{|\alpha|/2}$$

$$\|(\hbar\nabla_z)^\beta e_\nu\| \leqslant (\hbar|\nu|)^{|\beta|/2} \leqslant \omega_*^{-|\beta|/2} E_0(\nu)^{|\beta|/2}$$

whence we obtain the estimate (A1) with

$$C_V = \sum_{|\alpha+\beta| \leqslant k} \sup_{0<\hbar<1} |v_{\alpha,\beta}| \omega_*^{-|\alpha+\beta|/2} (k+1)^{|\alpha|/2}.$$

The proof is complete. □

References

[1] Benettin G, Galgani L and Giorgilli A 1987 Realization of holonomic constraints and freezing of high-frequency degrees of freedom in the light of classical perturbation theory Part I *Commun. Math. Phys.* **113** 87–103; 1989 Part II *Commun. Math. Phys.* **121** 557–601

[2] Galgani L 1988 Relaxation times and the foundations of classical statistical mechanics in the light of modern perturbation theory *Nonlinear Evolution and Chaotic Phenomena* ed G Gallavotti and P Zweifel (New York: Plenum) pp 147–60

[3] Giorgilli A 1988 Relevance of exponentially large time scales in practical applications: effective fractal dimensions in conservative dynamical systems *Nonlinear Evolution and Chaotic Phenomena* ed G Gallavotti and P Zweifel (New York: Plenum) pp 161–70

[4] Graffi S, Grecchi V and Simon B 1970 Borel summability: applications to the anharmonic oscillator *Phys. Lett.* **32B** 631–4

[5] Giorgilli A and Galgani L 1985 Rigorous estimates for the series expansions of Hamilton perturbation theory *Celest. Mech.* **37** 95–112

[6] Giorgilli A and Galgani L 1978 Formal integrals for an autonomous Hamiltonian system near an equilibrium point *Celest. Mech.* **17** 267–80

[7] Giorgilli A, Delsham A, Fontich E, Galgani L and Simó C 1989 Effective stability for a Hamiltonian system near an elliptic equilibrium point, with an application to the restricted three body problem *J. Diff. Eq.* **77** 167–98

[8] Graffi S and Paul T 1987 The Schrödinger equation and canonical perturbation theory *Commun. Math. Phys.* **108** 25-40

[9] Degli Esposti M, Graffi S and Herczyński J 1990 Quantization of the classical Lie algorithm in the Bargmann representation, in preparation

[10] Ali M K 1985 The quantum normal form and its equivalents *J. Math. Phys.* **26** 2565-72

[11] Benettin G, Chierchia L and Fassó F 1989 Exponential estimates for the one-dimensional Schrödinger equation with bounded analytic potential *CARR Reports in Math. Phys.* 11/89

[12] Lichtenberg A J and Lieberman M A 1983 *Regular and Stochastic Motion* (Berlin: Springer)

[13] Bargmann V 1961 On a Hilbert space of analytic functions and an associated integral transform *Commun. Pure Appl. Math.* **14** 187-214