Method of Averaging and the Quantum Anharmonic Oscillator

A. Ben Lemlih and J. A. Ellison
Department of Mathematics, University of New Mexico, Albuquerque, New Mexico 87131
(Received 5 August 1985)

The Krylov-Bogoliubov method of averaging is applied to the time-dependent quantum anharmonic oscillator. A regular perturbation expansion contains secular terms. The averaging approximation does not, and as a result has a validity over larger time intervals. A new variant of the usual averaging transformation is used and *rigorous* error bounds are derived. Rigorous averaging methods have been applied extensively to ordinary differential equations but our work appears to be the first generalization to partial differential equations.

PACS numbers: 03.65.Ge, 02.70.+d, 03.20.+i, 61.80.Mk

The Krylov-Bogoliubov method of averaging is a powerful perturbation procedure for ordinary differential equations which, unlike most other perturbation techniques, allows for a rigorous estimate of the difference between the solution of an exact problem and an approximate solution based on an averaged problem. In this Letter, we generalize the method, complete with error bounds, to the quantum anharmonic oscillator

$$i\psi' = H\psi, \quad \psi(0) = \psi_0. \tag{1}$$

Here the Hamiltonian $H = H_0 + \frac{1}{4}\epsilon q^4$ where $H_0 = \frac{1}{2}(p^2 + q^2)$ is the harmonic-oscillator Hamiltonian and the perturbation parameter, ϵ, is a nonnegative real number. A regular perturbation procedure, $\psi(t) = w_0(t) + \epsilon w_1(t) + w_2(t;\epsilon)$, applied to (1) gives rise to a secular term in w_1. This gives incorrect qualitative behavior and limits the validity of $w_0 + \epsilon w_1$, as an approximation to ψ, to $O(1)$ time intervals. In fact, it is easy to prove that $\| w_2(t;\epsilon) \| = O(\epsilon^2)$, but only on $O(1)$ time intervals. As we show, the method of averaging provides an approximation $\psi_1(t;\epsilon)$ to ψ, which has no secular terms and satisfies $\| \psi(t) - \psi_1(t) \| = O(\epsilon^2 t)$ on $L_2(R)$. Thus ψ_1 is an $O(\epsilon^2)$ approximation on $O(1)$ time intervals and an $O(\epsilon)$ approximation on $O(\epsilon^{-1})$ time intervals. It has the added feature that the $O(\epsilon)$ term in the expansion of the eigenvalues of H naturally appears in the approximation.

Averaging has been applied extensively to ordinary differential equations,[1] although its importance, with a few exceptions, is just beginning to be appreciated in the physics community. In the classical case for H above, the equation of motion is the so-called Duffing equation and the method is easily applied complete with error bounds. To our surprise, we found no rigorous work generalizing averaging to partial differential equations.[2]

We became interested in the extension to partial differential equations in a problem involving the Hamiltonian

$$H = (2m)^{-1}p^2 + V(x,y,z)$$

governing the motion of charged particles in perfect crystals. Here V is the periodic crystal potential and z is taken along a major crystal axis. Averaging has been used[3] in the classical case to discuss rigorously the channeling problem of replacing V by its z average to determine the motion of a particle with a large z momentum. In our attempt to treat the quantum case we ran into a problem and decided to try the above anharmonic oscillator which is a widely used test case for perturbation procedures. This Letter is the result of that work.

Let $\mathscr{H}$ be the Hilbert space of functions $f:R \to C$ which are Lebesgue square integrable; then it is well known that H and H_0 are self-adjoint operators[4] on $\mathscr{H}$ where q is the multiplicative operator and $p = -i\, d/dz$. The evolution of the wave function is governed by the Schrödinger equation (1) and the unique solution of ψ_0 in the domain of self-adjointness of H is given by[5]

$$\psi(t) = e^{-iHt}\psi_0. \tag{2}$$

Here the exponential operator is a strongly continuous one-parameter unitary group with generator $-iH$.

In order to put Eq. (1) in a form to apply the method of averaging, we introduce the variation-of-parameters transformation

$$\psi(t) = e^{-iH_0t}\phi(t), \tag{3}$$

where the exponential operator is the unitary group with generator $-iH_0$. If we use Eq. (1), the initial-value problem for ϕ becomes

$$i\phi' = \epsilon A(t)\phi, \quad \phi(0) = \psi_0, \tag{4}$$

where

$$A(t) = e^{iH_0t}\tfrac{1}{4}q^4 e^{-iH_0t}. \tag{5}$$

Equation (4) is in a form for the method of averaging. We now proceed to illustrate this method in a formal way. This will give us a candidate for an approximate solution of the original problem (1). Then we derive, in a rigorous way, error bounds relating the approximate and exact solutions.

It is convenient to introduce the raising and lowering operators

$$a = \tfrac{1}{2}\sqrt{2}(q + ip), \quad a^\dagger = \tfrac{1}{2}\sqrt{2}(q - ip). \tag{6}$$

Equation (5) can then be rewritten as

$$A(t)=\tfrac{1}{16}\{e^{-4it}a^4+4e^{-2it}aH_0a+4e^{2it}a^\dagger H_0a^\dagger+e^{4it}a^{\dagger 4}\}+\tfrac{3}{8}H_0^2+\tfrac{3}{32}. \tag{7}$$

The operator A is a periodic in t with period π and its average $\bar{A}$ is

$$\bar{A}=\tfrac{3}{8}H_0^2+\tfrac{3}{32}. \tag{8}$$

For ϵ small, ϕ as defined by (4) appears to be a slowly varying function of time and one might expect that an approximation to ϕ could be obtained by considering the averaged problem

$$iv'=\epsilon\bar{A}v, \quad v(0)=\psi_0. \tag{9}$$

Since $\bar{A}$ is a self-adjoint operator on $\mathscr{H}$ the solution of (9) is

$$v(t)=e^{-i\epsilon\bar{A}t}\psi_0=\sum_{n=0}^{\infty}\psi_{0n}e^{-i\epsilon\beta_n t}\chi_n, \tag{10}$$

where the χ_n are the Hermite functions which form an orthonormal basis for $\mathscr{H}$,

$$H_0\chi_n=E_n\chi_n, \quad E_n=n+\tfrac{1}{2},$$

$$\bar{A}\chi_n=\beta_n\chi_n, \; \beta_n=\tfrac{3}{8}E_n^2+\tfrac{3}{32},$$

and the ψ_{0n} are the coordinates of the initial data. Notice that $\epsilon\beta_n$ gives the $O(\epsilon)$ correction to the harmonic-oscillator eigenvalues[6] and it would be interesting to compare this with current operator approaches to the approximating of eigenvalues.[7]

The method of averaging is a transformation procedure to quantify the relation between solutions of (9) and (4). The usual way to proceed is to look for a near-identity transformation $\phi(t)=u(t)+\epsilon P(t)u(t)$, where $P(t)$ is a one-parameter family of operators on H, which transforms Eq. (4) into the averaged equation (9) plus a small perturbation term. However, this approach leads to certain difficulties with unbounded operators and their invertibility. A novel feature of our approach is that we look for a near-identity transformation of the form

$$w(t)=v(t)+\epsilon P(t)v(t), \tag{11}$$

with $P(t)$ as before, which transforms the averaged equation (9) into

$$iw'=\epsilon A(t)w-\epsilon^2R(t)v, \quad w(0)=\psi_0, \tag{12}$$

where $R(t)$ is a linear operator to be determined. This eliminates the difficulty and leads to a more straightforward error analysis. It is important to observe that the perturbation term in Eq. (12), i.e., the $O(\epsilon^2)$ term, depends on the solution to the averaged problem, v. This is in contrast to usual averaging approaches and is important for the error analysis.

Thus we are looking for $P(t)$ such that the averaged problem (9) is transformed into (4) plus a small perturbation term [here it is convenient to take $P(0)=0$ so that $w(0)=v(0)$]. Since we expect w to approximate ϕ, Eqs. (3) and (11) suggest the approximation

$$\psi_1(t)=e^{-iH_0t}[1+\epsilon P(t)]e^{-i\epsilon\bar{A}t}\psi_0 \tag{13}$$

to the original problem (1). ψ_1 is easily written in terms of the eigenfunctions χ_n of H_0. A second, less refined approximation, ψ_2, is obtained by ignoring the P term in (13):

$$\psi_2(t)=e^{-i(H_0+\epsilon\bar{A})t}\psi_0. \tag{14}$$

Plugging (11) into (12) and making use of (9) gives

$$iP'(t)=A(t)-\bar{A}, \tag{15a}$$

$$R(t)=A(t)P(t)-P(t)\bar{A}, \tag{15b}$$

where (15a) is the $O(\epsilon)$ term in the resulting equation and (15b) the remaining term. Equation (15a) is easily solved to give

$$P(t)=-\tfrac{1}{16}i\{\mathscr{I}e^{-4it}a^4+4\mathscr{I}e^{-2it}aH_0a+4\mathscr{I}e^{2it}a^\dagger+H_0a^\dagger+\mathscr{I}e^{4it}a^{\dagger 4}\}, \tag{16}$$

where I denotes the time integral from 0 to t. $R(t)$ is then determined from Eqs. (7), (8), and (16).

In order to justify the approximations to $\psi(t)$ given in (13) and (14), we need to introduce some notation. The inner product (f,g) for f and g in $\mathscr{H}$ will be the integral over R of the complex conjugate of f times g. The norm of f will be denoted $\|f\|$ and will equal the square root of (f,f). A function $f(t)$ which belongs to $\mathscr{H}$ for all t in the domain of interest will be said to be t-diff if, and only if, $h^{-1}[f(t+h)-f(t)]\equiv h^{-1}\Delta f$ converges in $\mathscr{H}$ as $h\to 0$. Its limit will be denoted by f'.

If T is an operator defined on the Hermite functions and $f_n=(\chi_n,f)$, then we define T on $\mathscr{H}$ by

$$D(T)=\{f\in\mathscr{H}\,|\lim_{N\to\infty}\sum_{n=0}^{N}f_nT\chi_n \text{ exists in } \mathscr{H}\},$$

$$Tf=\sum_{n=0}^{\infty}f_nT\chi_n, \quad f\in D(T). \tag{17}$$

This defines all our operators: H, H_0, a, $a^\dagger$, $A(t)$, $\bar{A}$, $P(t)$, $R(t)$ and the exponential operators associated with H, H_0, and $\bar{A}$. The domain of self-adjointness of H is larger than $D(H)$, whereas the domains of self-adjointness for H_0 and $\bar{A}$ are as above. Finally, let the domain D_l be defined as

$$D_l = \{ f \in \mathscr{H} | \sum_{n=0}^{\infty} |n^l f_n|^2 < \infty \}, \tag{18}$$

where l is a nonnegative integer. In the following, $\psi(t)$ is defined by Eq. (2) and $\psi_1(t)$ and $\psi_2(t)$ are defined by Eqs. (13) and (14).

Our main result is the following:

Averaging theorem.—If $\psi_0 \in D_4$ then there exists a constant C independent of t and ϵ such that

$$\| \psi(t) - \psi_1(t) \| \leq \epsilon^2 \int_0^t \| R(s) e^{-i\epsilon \bar{A} s} \psi_0 \| \, ds \leq C\epsilon^2 t. \tag{19}$$

Since $A(t)$, $\bar{A}$, and $P(t)$ map D_l into D_{l-2} for $l \geq 2$ and $e^{-i\epsilon \bar{A} t}$ maps D_l into D_l it is easy to see that the integrand makes sense for $\psi_0 \in D_4$.

Corollary.—If $\psi_0 \in D_4$ then there exist constants C and C_1 independent of t and ϵ such that

$$\| \psi(t) - \psi_2(t) \| \leq C\epsilon^2 t + C_1 \epsilon.$$

The proof of the corollary follows simply from the theorem using $\psi - \psi_2 = \psi - \psi_1 + \psi_1 - \psi_2$. The proof of the theorem requires the *Lemma:* If $\psi_0 \in D_4$ then $\psi_1(t)$ is t-diff and

$$i\psi_1'(t) = H\psi_1(t) - \epsilon^2 e^{-iH_0 t} R(t) e^{-i\epsilon \bar{A} t} \psi_0.$$

Proof of theorem.—Let $e(t) = \psi(t) - \psi_1(t)$; then by the lemma, e is t-diff and $ie'(t) = He(t) + g(t)$ where $g(t) = \epsilon^2 e^{-iH_0 t} R(t) e^{i\epsilon \bar{A} t} \psi_0$. It follows that

$$(e(t), e(t))' = (He + g, ie) + (ie, He + g) = -2\mathrm{Im}(g, e),$$

where the second equality uses the symmetry of H. Integrating this equality and using the Schwartz inequality gives $\| e \|^2 \leq 2 \int_0^t \| g(s) \| \, \| e(s) \| \, ds$. Denoting the right-hand side of this inequality by $r(t)$ yields the differential inequality $r'(t) \leq 2 \| g(t) \| r(t)^{1/2}$ which can be solved to give $r(t)^{1/2} \leq \int_0^t \| g(s) \| \, ds$. The first inequality in (19) easily follows. The second inequality follows after an eigenfunction expansion for ψ_0 shows that the integrand can be bounded independent of s and ϵ. The argument here is similar to part three in the proof of the lemma.

Proof of lemma.—First, we show that $\psi_1(t)$ is in D_2. $e^{-i\epsilon \bar{A} t}$ maps D_4 into D_4 as is easily seen by expanding ψ_0 in terms of the Hermite functions. From Eq. (16), $P(t)$ is a linear combination of a^4, $aH_0 a$, $a^\dagger H_0 a^\dagger$, and $a^{\dagger 4}$ and, for example, a^4 acting on a function f in D_4 gives a function in D_2 because $a^4 f$ equals $\sum f_n a^4 \chi_n$ by definition and this sum is in D_2 since $a^4 \chi_n = O(n^2) \chi_{n-4}$. The other terms in $P(t)$ are similar. Thus $\psi_1(t) \in D_2$ because $e^{-iH_0 t}$ maps D_2 into D_2.

Second, we assume that $w(t) = [1 + \epsilon P(t)] e^{-i\epsilon \bar{A} t} \psi_0$ is t-diff; then it is easy to show that

$$i\psi_1'(t) = H_0 \psi_1(t) + e^{-iH_0 t} i w'(t).$$

In fact,

$$h^{-1} \| i\Delta\psi_1 - h(H_0\psi_1 + e^{-H_0 t} i w') \| \leq h^{-1} \| (e^{-iH_0 h} - 1 + ihH_0) w \| + h^{-1} \| \Delta w - hw' \| + h^{-1} \| (e^{-iH_0 h} - 1)(\Delta w - hw') \| + \| (e^{-iH_0 h} - 1) w' \|.$$

The first and fourth terms go to zero as $h \to 0$ since $w' \in \mathscr{H}$ and $w(t) \in D_2 \subset D(H_0)$ which is contained in the domain of self-adjointness of H_0.[5] The second and third terms go to zero because w is assumed to be t-diff.

Third, we show that for $\psi_0 \in D_4$, $w(t)$ is t-diff and

$$iw'(t) = \epsilon A(t) w(t) - \epsilon^2 R(t) e^{-i\epsilon \bar{A} t} \psi_0.$$

Clearly, $(e^{-i\epsilon \bar{A} t} \psi_0)' = -i\epsilon \bar{A} e^{-i\epsilon \bar{A} t} \psi_0$ for $\psi_0 \in D_4$ since it is easy to check that D_4 is contained in the domain of self-adjointness of $\bar{A}$. The result follows if

$$[iP(t) e^{-i\epsilon \bar{A} t} \psi_0]' = [A(t) - \bar{A} + \epsilon P(t) \bar{A}] e^{-i\epsilon \bar{A} t} \psi_0.$$

But $P(t)$ is a linear combination of the operators mentioned in the first step and it suffices to consider only one of the terms, for example,

$$u(t) = e^{-2it} a H_0 a e^{-i\epsilon \bar{A} t} \psi_0 = \sum_{n=0}^{\infty} \psi_{0n} \exp[-i(2 + \epsilon\beta_n) t] \alpha_n \chi_{n-2}$$

and its formal derivative

$$u_1(t) = -i \sum_{n=0}^{\infty} \alpha_n (2 + \epsilon\beta_n) \psi_{0n} \exp[-i(2 + \epsilon\beta_n) t] \chi_{n-2}(x),$$

where $\alpha_n = (n-1)^{1/2}(n - \frac{1}{2})\sqrt{n} = O(n^2)$. Since $\beta_n = O(n^2)$ and $\psi_0 \in D_4$, $u_1(t)$ makes sense in $\mathscr{H}$. Also the series for $h^{-2} \| u(t+h) - u(t) - u_1(t) h \|^2$ converges uniformly for $h \in (0, \infty)$ and the limit as $h \to 0$ can be taken inside the sum giving zero. Therefore $u_1(t)$ is the t derivative of $u(t)$. Repeating this argument for the other terms of $P(t)$ gives the desired result.

Finally, we now have

$$i\psi_1'(t) = H_0 \psi_1(t) + e^{-iH_0 t} [\epsilon A(t) w(t) - \epsilon^2 R(t) e^{-i\epsilon \bar{A} t} \psi_0],$$

and it is easy to check that $e^{-iH_0 t} A(t) w(t) = \frac{1}{4} q^4 \psi_1(t)$ which completes the proof of the lemma and the averaging theorem.

This paper shows that the method of averaging complete with error bounds can be extended to partial differential equations. Because averaging is rigorous it can be used to establish properties of the exact solution by use of the averaged system as is done with ordinary differential equations.[8,9] Here we have concentrated on first-order averaging but the details of second-order averaging have been worked out and in principle can be extended to the nth order.[10] Also the perturbation in Eq. (1) can be replaced by polynomials in q which are bounded below.[10] More generally, averaging can be attempted on problems of the form of Eq. (4) where the average of $A(t)$ makes sense and can be adequately characterized. For example, it has been applied[10] to the problems of Cook, Shankland, and Wells[11] and Coutsias and McIver[12] which contain time-dependent perturbations and these turn out to be simpler to treat than Eq. (1). Reference 12 uses the multiple-time-scale perturbation method and, just as in the ordinary differential equation case, this provides an interesting comparison with averaging. We believe that there are many more applications of averaging to partial differential equations.

We gratefully acknowledge discussions with H. S. Dumas and A. W. Sáenz. This work was partially supported by the National Science Foundation through Grant No. DMR-8214301.

[1]See, e.g., N. N. Bogoliubov and Y. A. Mitropolsky, *Asymptotic Methods in the Theory of Nonlinear Oscillations* (Gordon and Breach, New York, 1961), Chaps. 5 and 6; J. K. Hale, *Ordinary Differential Equations* (Wiley-Interscience, New York, 1969), 1st ed., Chap. 5; V. I. Arnold, *Geometrical Methods in the Theory of Ordinary Differential Equations* (Springer, New York, 1983), Chap. 4; T. J. Burns and J. A. Ellison, Phys. Rev. B **29**, 2790 (1984); H. S. Dumas and J. A. Ellison, "Particle Channeling in Crystals and the Method of Averaging," in Proceedings of the Workshop on Local and Global Methods of Dynamics, edited by R. Cawley, A. W. Sáenz, and W. W. Zachary (Springer, New York, to be published); H. S. Dumas, J. A. Ellison, and A. W. Sáenz, to be published. The last preceding work eliminates a restrictive assumption needed in the averaging results of the first two papers. See also C. Robinson, IEEE Trans. Circuits Syst. **30**, 591 (1983); and references therein.

[2]There are, however, formal discussions of averaging for partial differential equations in the literature and M. Kummer, Nuovo Cimento **B1**, 123 (1971), is a good example in relation to this paper.

[3]See Burns and Ellison, Dumas and Ellison, and Dumas, Ellison, and Sáenz, Ref. 1.

[4]See M. Reed and B. Simon, *Methods of Modern Mathematical Physics II: Fourier Analysis, Self-Adjointness* (Academic, New York, 1972), for several proofs of the self-adjointness of H for $\epsilon > 0$.

[5]See, e.g., M. Reed and B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis* (Academic, New York, 1972), p. 265, Theorem VIII.7.

[6]See, e.g., L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Pergamon, New York, 1977), p. 136.

[7]See, e.g., I. D. Feranchuk and L. I. Komarov, Phys. Lett. **88A**, 211 (1982), and references therein.

[8]See Hale, Ref. 1.

[9]See Robinson, Ref. 1.

[10]A. Ben Lemlih, thesis, University of New Mexico (to be published).

[11]R. J. Cook, D. G. Shankland, and A. L. Wells, Phys. Rev. A **31**, 564 (1985).

[12]E. A. Coutsias and J. K. McIver, Phys. Rev. A **31**, 3155 (1985).