PHYSICAL REVIEW
LETTERS

VOLUME 74 | 27 FEBRUARY 1995 | NUMBER 9

Superconvergent Perturbation Method in Quantum Mechanics

Wolfgang Scherer
Institut für Theoretische Physik A, TU Clausthal, Leibnizstrasse 10, D-38678 Clausthal-Zellerfeld, Germany
(Received 13 July 1994)

An analog of Kolmogorov's superconvergent perturbation theory in classical mechanics is constructed for self-adjoint operators. It is different from the usual Rayleigh-Schrödinger perturbation theory and yields expansions for eigenvalues and eigenvectors in terms of functions of the perturbation parameter.

PACS numbers: 03.65.–w, 02.30.Mv, 02.90.+p, 31.15.+q

About a century ago Lindstedt [1], Poincaré [2], and von Zeipel [3] developed a perturbation theory for Hamiltonian systems in classical mechanics (CM). The method was mainly used in celestial mechanics but in most cases failed to converge due to the appearance of small divisors. It was only in the fifties that Kolmogorov [4] proposed a new quadratically convergent perturbation method for Hamiltonian systems in CM and this (nowadays called superconvergent) method was essential in Arnold's proof of the Kolmogorov-Arnold-Moser theorem [5].

In this Letter an exact analog of Kolmogorov's superconvergent method is constructed for self-adjoint operators. So far "superconvergent" is only a name for this new method and a detailed functional analytic investigation will have to determine how much the new perturbation theory constructed here is an improvement on existing schemes. There are, however, indications that this will be so. The first comes from the fact that we have also constructed an analog of the Poincaré–von Zeipel perturbation theory for self-adjoint operators and have shown that this analog is identical to the usual Rayleigh-Schrödinger perturbation theory [6]. Since in CM the superconvergent method is a vast improvement on the Poincaré–von Zeipel method, we may expect the same in quantum mechanics (QM). The second comes from initial numerical studies.

(a) General algorithm of the superconvergent method.—Let H_0^0 be the unperturbed Hamiltonian (CM, function on phase space; QM, operator) and let

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

be the perturbed Hamiltonian, where the H_p^0 do not depend on the perturbation parameter ϵ. Moreover, let

$$W^n(\epsilon) := \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} W_{p+1}^n , \tag{2}$$

and let $-W(\epsilon)$ be generators (CM, functions on phase space; QM, operators) of (CM, canonical; QM, unitary) flows $[\Phi^n(\epsilon)]^{-1}$ with "time" ϵ. This means that the transformations $\Phi^n(\epsilon)$ satisfy the initial-value problem

$$\frac{d}{d\epsilon}\Phi^n(\epsilon)^* = \mathrm{ad}W^n(\epsilon) \circ \Phi^n(\epsilon)^*, \quad \Phi^n(0) = \mathrm{id}, \tag{3}$$

where in CM the $\Phi^n(\epsilon)^*$ act on any phase space function A via

$$\Phi^n(\epsilon)^*(A) := A \circ \Phi^n(\epsilon), \tag{4}$$

and where in CM

$$\mathrm{ad}W^n(\epsilon)(A) := \{W^n(\epsilon), A\} \tag{5}$$

and $\{\cdot,\cdot\}$ denotes the Poisson bracket. In QM, self-adjoint operators $-W^n(\epsilon)$ generate one-parameter groups of unitary transformations $[\Phi^n(\epsilon)]^{-1}$ such that the $\Phi^n(\epsilon)$ solve the same initial-value problem as given in (3) but where the $\Phi^n(\epsilon)^*$ act on operators A via

$$\Phi^n(\epsilon)^*(A) := \Phi^n(\epsilon)^{-1} A \Phi^n(\epsilon), \tag{6}$$

and where now

$$\mathrm{ad}W^n(\epsilon)(A) := \frac{i}{\hbar}[W^n(\epsilon), A] \tag{7}$$

and $[\cdot,\cdot]$ denotes the commutator. With the understanding that the appropriate definition (5) or (7) (depending on

0031-9007/95/74(9)/1495(5)$06.00

whether one is dealing with CM or QM) is chosen we proceed to present Kolmogorov's superconvergent method. In the following it will be useful to have an expansion for $\Phi^n(\epsilon)^*$ in terms of operators T_p^n independent of ϵ. Writing

$$\Phi^n(\epsilon)^* = \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} T_p^n, \tag{8}$$

one finds that the T_p^n are recursively determined by $T_p^n = \mathrm{id}$ and

$$T_{p+1}^n = \sum_{l=0}^{p} \binom{p}{l} \mathrm{ad} W_{l+1}^n \circ T_{p-l}^n. \tag{9}$$

In CM the T_p^n are differential operators such that $T_p^n(A)$ means application of T_p^n to the phase space function A, whereas in QM they are operators acting on operators.

Kolmogorov's method consists of finding the generators $W^n(\epsilon)$ in such a way that

$$K^{n-1}(\epsilon) := \Phi^{n-1}(\epsilon)^* \circ \Phi^{n-2}(\epsilon)^* \circ \cdots \circ \Phi^1(\epsilon)^*[H^0(\epsilon)] \tag{10}$$

has an expansion in ϵ

$$K^{n-1}(\epsilon) := \sum_{p=0}^{\infty} \frac{\epsilon^p}{p!} K_p^{n-1}, \tag{11}$$

with "integrable" terms up to order $2^{n-1} - 1$. (Whereas in CM the term integrable has a well-defined meaning, this fails to be the case in QM. However, we only use the notion of integrability to provide motivation for the superconvergent method; all its equations are well defined in CM without any need to define integrability and will be well-defined operator equations in QM even though the term "integrability" is not.) Then a new perturbed Hamiltonian $H^{n-1}(\epsilon) := K^{n-1}(\epsilon)$ is defined such that its unperturbed part H_0^{n-1} consists of the integrable part of K^{n-1} and its perturbation is of the order 2^{n-1}. We can summarize the iterative procedure as follows (for a more detailed discussion we refer the reader to [7] with the warning that our notation differs from theirs):

Result of Φ^{n-1}: H^{n-1}	Choice of Φ^n: W^n	Result of Φ^n: H^n
H_0^{n-1}	$W_p^n = 0, \quad 1 \le p \le 2^{n-1}$	$H_0^n = H_0^{n-1}$
[int. up to $O(2^{n-1} - 1)$]	$\mathrm{ad} W_p^n(H_0^{n-1})$	$+ \sum_{p=2^{n-1}}^{2^n-1} \frac{\epsilon^p}{p!} \overline{H_p^{n-1}}$
	$= \overline{H_p^{n-1}} - H_p^{n-1}, 2^{n-1} \le p < 2^n$	[int. up to $O(2^n - 1)$]
$H_p^{n-1} = 0, \quad 1 \le p < 2^{n-1}$	with $\mathrm{ad} H_0^{n-1}(\overline{H_p^{n-1}}) = 0$	
(no perturbation)	$W_p^n = 0, \quad 2^n \le p$	
	leads to	$H_p^n = 0, \quad 1 \le p < 2^n$
$H_p^{n-1} = K_p^{n-1}, \quad 2^{n-1} \le p$	$K_0^n = H_0^{n-1}$	(no perturbation)
	$K_p^n = 0, \quad 1 \le p < 2^{n-1}$	
	$K_p^n = \overline{H_p^{n-1}}, \quad 2^{n-1} \le p < 2^n$	$H_p^n = K_p^n, \quad 2^n \le p$
	$K_p^n = H_p^{n-1} + \sum_{j=1}^{p-1} \binom{p-1}{j-1} [\mathrm{ad} W_j^n(K_{p-j}^n) + T_{p-j}^n(H_j^{n-1})], \quad 2^n \le p$	

The meaning of $\bar{\ }$ is to be understood as follows. Assume H^{n-1} is of the form given in the left column and choose $W_p^n = 0$ for $1 \le p < 2^{n-1}$. Then one finds first $K_0^n = H_0^{n-1}$, $K_p^n = 0$ for $1 \le p < 2^{n-1}$, and

$$K_p^n = \mathrm{ad} W_p^n(H_0^{n-1}) + H_p^{n-1}, \qquad 2^{n-1} \le p < 2^n. \tag{12}$$

Since H_0^{n-1} is already integrable up to $O(2^{n-1} - 1)$ and Φ^n should improve this, one would like to have that the K_p^n for $2^{n-1} \le p < 2^n$ are integrable. Consequently, the crucial point in this procedure becomes the construction of the W_p^n and some $\overline{H_p^{n-1}}$ for $2^{n-1} \le p < 2^n$ such that

$$\mathrm{ad} W_p^n(H_0^{n-1}) = \overline{H_p^{n-1}} - H_p^{n-1}, \tag{13}$$

where $\overline{H_p^{n-1}}$ is such that

$$\mathrm{ad} H_p^{n-1}(\overline{H_p^{n-1}}) = 0. \tag{14}$$

Then it follows from (12) that $K_p^n = \overline{H_p^{n-1}}$ and from (14) that they commute with the unperturbed Hamiltonian H_0^{n-1} of the previous step. The table summarizing the method omits these intermediate steps and shows only the results once (13) and (14) have been solved.

In CM (13) and (14) can be satisfied with the help of the averaging method and $\overline{H_p^{n-1}}$ turns out to be the average of H_p^{n-1} over the angle variables of the tori, which are assumed to be compact. It should be noted, however, that the $\overline{H_p^{k-1}}$ in general depend on ϵ such that (15) is actually an expansion in terms of functions of the perturbation parameter. Since H_0^{n-1} is assumed to be integrable (14) implies that $\overline{H_p^{n-1}}$ and hence H_0^n are integrable as well.

Hence, after the nth transformation one has a Hamiltonian $H^n(\epsilon)$, where

$$H_0^n = \sum_{k=0}^{n}\left(\sum_{p=2^{k-1}}^{2^k-1}\frac{\epsilon^p}{p!}\overline{H_p^{k-1}}\right) \tag{15}$$

is integrable and the perturbations H_p^n are of order 2^n or higher.

(b) Superconvergent method for self-adjoint operators.—With our previous notation Eqs. (13) and (14) become operator equations in QM which have to be solved for each step. We can construct operators W_p^n, $\overline{H_p^{n-1}}$ satisfying (13) and (14) with the help of the following quantum analog of the classical averaging procedure [6], which is a modified version of an idea used by Weinstein [8] in the context of pseudodifferential operators. Let A, B be self-adjoint operators such that

$$B^{(A)}(t) := \exp\left(-\frac{i}{\hbar}tA\right)B\exp\left(\frac{i}{\hbar}tA\right), \tag{16}$$

$$\overline{B}^{(A)} := \lim_{T\to\infty}\frac{1}{T}\int_0^T B^{(A)}(t)dt, \tag{17}$$

$$S^{(A)}(B) := \lim_{T\to\infty}\frac{1}{T}\int_0^T dt\int_0^t ds[B^{(A)}(s) - \overline{B}^{(A)}] \tag{18}$$

exist and

$$\lim_{T\to\infty}\frac{B^{(A)}(T) - B}{T} = 0. \tag{19}$$

Then the following holds:

$$\mathrm{ad}\overline{B}^{(A)}(A) = \frac{i}{\hbar}[\overline{B}^{(A)}, A] = 0 \tag{20}$$

and

$$\mathrm{ad}S^{(A)}(A) = \frac{i}{\hbar}[S^{(A)}(B), A] = \overline{B}^{(A)} - B. \tag{21}$$

Proof:

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

by assumption, and thus

$$\begin{aligned}\mathrm{ad}S^{(A)}(B)(A) &= \lim_{T\to\infty}\frac{1}{T}\int_0^T dt\int_0^t ds\,\frac{i}{\hbar}[B^{(A)}(s), A]\\ &= \lim_{T\to\infty}\frac{1}{T}\int_0^T dt\int_0^t ds\frac{d}{ds}B^{(A)}(s)\\ &= \lim_{T\to\infty}\frac{1}{T}\int_0^T dt[B^{(A)}(s) - B] = \overline{B}^{(A)} - B,\end{aligned}$$

which completes the proof. Setting now $W_p^n = S^{(H_0^{n-1})} \times (H_p^{n-1})$ solves (13) and (14) in QM with $\overline{H_p^{n-1}} = \overline{H_p^{n-1}}^{(H_0^{n-1})} := \overline{H_p}^{(n-1)}$, where the last equation introduces a simplified notation.

(c) Interpretation of the algorithm for self-adjoint operators.—The results of the preceding section show that we can execute Kolmogorov's superconvergent perturbation algorithm in QM as well. But what has been gained? By construction $H^n(\epsilon)$ and the original perturbed Hamiltonian $H^0(\epsilon)$ are unitarily equivalent:

$$H^n(\epsilon) = \Phi^n(\epsilon)^{-1}\cdots\Phi^1(\epsilon)^{-1}H^0(\epsilon)\Phi^1(\epsilon)\cdots\Phi^n(\epsilon) \tag{22}$$

and thus have identical spectra. Since after the first transformation $[H_0^0, \overline{H_1}^{(0)}] = 0$, we can diagonalize H_0^0 and $\overline{H_1}^{(0)}$ simultaneously which permits us to diagonalize $H_0^1 := H_0^0 + \epsilon\overline{H_1}^{(0)}$. After the second transformation we have

$$[H_0^1, \overline{H_2}^{(1)}] = 0 = [H_0^1, \overline{H_3}^{(1)}], \tag{23}$$

and thus we can diagonalize

$$H_0^2 = H_0^0 + \epsilon\overline{H_1}^{(0)} + \frac{\epsilon^2}{2!}\overline{H_2}^{(1)} + \frac{\epsilon^3}{3!}\overline{H_3}^{(1)}, \tag{24}$$

etc. In this way we arrive after n transformations at a diagonal $H_0^n(\epsilon)$ whose eigenvalues $E_j^n(\epsilon)$ approximate those of $H^n(\epsilon)$ and thus of $H^0(\epsilon)$. The eigenvectors $|j\rangle^n$ of H_0^n are known by construction, and they approximate those of H^n. Because of (22) it follows that $\Phi^1(\epsilon)\cdots\Phi^n(\epsilon)|j\rangle^n$ are the appropriately approximated eigenvectors of the original perturbed $H^0(\epsilon)$. In formulas, if $H^0(\epsilon)|j\rangle(\epsilon) = E_j(\epsilon)|j\rangle(\epsilon)$ then

$$E_j(\epsilon) = E_j^n(\epsilon) + O(2^n), \tag{25}$$

$$|j\rangle(\epsilon) = \Phi^1(\epsilon)\cdots\Phi^n(\epsilon)|j\rangle^n + O(2^n). \tag{26}$$

Since the $\overline{H_p}^{(n)}$ depend on ϵ, the $E_j^n(\epsilon)$ are no longer a pure power series expansion but should be viewed as an expansion in terms of functions of ϵ. These functions themselves have in general infinite series expansions in ϵ so that already low orders of $E_j^n(\epsilon)$ contain contributions of all order in ϵ. The example treated in the next section will make this point clearer.

(d) An example: discrete and nondegenerate spectrum.—It remains to show that the method constructed here is truly distinct from the usual Rayleigh-Schrödinger perturbation theory. For this purpose we sketch the results for the case of a Hamiltonian $H_0^0 = \sum_j |j\rangle E_j^0\langle j|$ with a purely discrete and nondegenerate spectrum and a perturbation only linear in ϵ, i.e., $H_1^0 = \sum_{j,k}|j\rangle V_{jk}\langle k|$ and $H_p^0 = 0$ for $p \geq 2$. In this case the eigenvectors $|j\rangle^1$ of $H_0^1 = H_0^0 + \epsilon\overline{H_1}^{(0)}$ and $|j\rangle$ of H_0^0 coincide because H_0^0

and $\overline{H_1}^{(0)}$ commute and H_0^0 has a nondegenerate spectrum. Continuing this argument inductively it is easy to see that H_0^0 and all H_0^n have the same eigenvectors $|j\rangle$. By straightforward calculation one finds

$$W_p^n = \frac{\hbar}{i} \sum_{j \neq k} |j\rangle \frac{\langle j|H_p^{n-1}|k\rangle}{E_j^{n-1} - E_k^{n-1}} \langle k|, \tag{27}$$

$$\overline{H_p}^{(n-1)} = \sum_j |j\rangle\langle j|H_p^{n-1}|j\rangle\langle j|, \tag{28}$$

$$E_j^n = E_j^0 + \sum_{k=1}^{n} \left(\sum_{l=2^{k-1}}^{2^k-1} \frac{\epsilon^l}{l!} E_j^{(l)} \right), \tag{29}$$

where $E_j^{(l)} := \langle j|H_l^{k-1}|j\rangle$, $2^{k-1} \leq l \leq 2^k - 1$, and the E_j^n are the eigenvalues of H_0^n. From the algorithm presented in Sect. (*a*) one obtains

$$H_2^1 = \frac{i}{\hbar}[W_1^1, \overline{H_1}^{(0)} + H_1^0], \tag{30}$$

$$H_3^1 = \left(\frac{i}{\hbar}\right)^2 [W_1^1, [W_1^1, \overline{H_1}^{(0)} + 2H_1^0]], \tag{31}$$

$$H_4^2 = \left(\frac{i}{\hbar}\right)^3 [W_1^1, |W_1^1, [W_1^1, \overline{H_1}^{(0)} + 3H_1^0]|] + 3\frac{i}{\hbar}[W_2^2, \overline{H_2}^{(1)} + H_2^1]. \tag{32}$$

The appearance of the denominators $E_j^{n-1} - E_k^{n-1}$ in W_p^n leads to a perturbation theory substantially different from the usual Rayleigh-Schrödinger theory. However, this difference only shows up in the fourth and higher order terms, since in H_2^1 and H_3^1 only commutators with W_1^1 appear, which contains only the usual Rayleigh-Schrödinger denominator $E_j^0 - E_k^0$. Indeed, one finds

$$E_j^{(1)} = V_{jj}, \tag{33}$$

$$\frac{1}{2!} E_j^{(2)} = \sum_{j \neq k} \frac{|V_{jk}|^2}{E_j^0 - E_k^0}, \tag{34}$$

$$\frac{1}{3!} E_j^{(3)} = \sum_{m \neq j \neq k} \frac{V_{jk} V_{km} V_{mj}}{(E_j^0 - E_k^0)(E_j^0 - E_m^0)} - \sum_{j \neq k} \frac{|V_{jk}|^2 V_{jj}}{(E_j^0 - E_k^0)^2}, \tag{35}$$

showing that up to $O(3)$ the new method coincides with the usual Rayleigh-Schrödinger perturbation theory. But H_4^2 contains W_2^2 with denominators of the form $E_j^1 - E_k^1 = E_j^0 - E_k^0 + \epsilon(V_{jj} - V_{kk})$, which are functions of ϵ and which appear as denominators in

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

As an example we quote the result for the ground state correction up to $E_0^{(4)}$ with the quartic perturbation $H_1^0 = x^4$ of the harmonic oscillator $H_0^0 = -d^2/dx^2 + x^2$ (with $\hbar = 1$)

$$E_0^0(\epsilon)_{\mathrm{SU}} = 1 + \frac{3}{4}\epsilon - \frac{21}{16}\epsilon^2 + \frac{333}{64}\epsilon^3 - \frac{3(1317760 + 12935472\epsilon + 36433368\epsilon^2 + 25183305\epsilon^3)}{2048(4 + 9\epsilon)(4 + 15\epsilon)(4 + 21\epsilon)}\epsilon^4 + O(\epsilon^5),$$

which approximates the numerically computed eigenvalues [9] much better than the standard fourth order Rayleigh-Schrödinger correction

$$E_0^0(\epsilon)_{\mathrm{RS}} = 1 + \frac{3}{4}\epsilon - \frac{21}{16}\epsilon^2 + \frac{333}{64}\epsilon^3 - \frac{30885}{1024}\epsilon^4 + O(\epsilon^5). \tag{36}$$

This example also illustrates the statements made earlier that the lower order corrections $E_j^{(n)}$ already contain infinite power series of ϵ, which may be the reason for improved convergence.

In conclusion, it should be emphasized that the method presented here is very general and, in principle, applicable to any self-adjoint operator A and perturbation B provided they (and the higher order operators) satisfy the conditions needed to construct $\overline{B}^{(A)}$ and $S^{(A)}(B)$ in (17) and (18). It is also designed to include analytic perturbations which are not necessarily only linear in the perturbation parameter. Moreover, it gives the corrections as integrals of the form $\overline{B}^{(A)}$ which, if desired, can be exhibited as sums over intermediate states as was done in Sect. (*d*), but which may be evaluated directly, circumventing the calculation of these sums, which may sometimes be impossible. This was already shown to be an advantage of the formulation of the usual Rayleigh-Schrödinger theory as an analog of the classical Poincaré–von Zeipel perturbation method in [6].

The method will be presented in more detail in a forthcoming longer paper [10].

[1] A. Lindstedt, Abh. K. Akad. Wiss. St. Petersburg **31**, No. 4 (1882).

[2] H. Poincaré, *Les Méthodes Nouvelles de la Mécanique Célèste I, II, III* (Dover Publ., New York, 1957).

[3] H. von Zeipel, Ark. Astron. Mat. Phys. **11–13** (1916–17).

[4] A. N. Kolmogorov, Dokl. Akad. Nauk SSSR **98**, 527 (1954).

[5] V. I. Arnold, Russian Math. Surveys **18**, 9 (1963).

[6] W. Scherer, TU-Clausthal Report No. ASI-TPA/16/94 [J. Phys. A (to be published)].

[7] A. J. Lichtenberg and M. A. Liebermann, *Regular and Stochastic Motion* (Springer, New York, 1983).

[8] A. Weinstein, Duke Math. J. **44**, 883 (1977).

[9] A. Galindo and P. Pascal, *Quantum Mechanics II* (Springer, New York, 1991).

[10] W. Scherer, TU-Clausthal Report No. ASI-TPA/18/94 (to be published).