prefer a somewhat different definition and set:

$$T = W + k(G) = W + [K,G] \,. \qquad (5.1)$$

In the scope of a perturbation theory, one can look at T as the "solution" of the problem. If one knows T, one can get easily the level-shift operator W and the generator G of the level-shift transformation,

$$W = \langle T \rangle \,, \qquad G = \frac{1}{k}\,(T) \,. \qquad (5.2)$$

This T-operator is an ideal tool to get a solution of a complicated problem in terms of the solutions of simpler subproblems. Suppose one can split the perturbation V such that

$$V = \textstyle\sum_\alpha V^\alpha \qquad (5.3)$$

and that $K + V^\alpha$ is a soluble subproblem. That is, we suppose we know explicitly the level-shift transformation

$$e^{G^\alpha}(K + V^\alpha)e^{-G^\alpha} = K + W^\alpha \,, \qquad (5.4)$$

and, therefore, the operator

$$T^\alpha = W^\alpha + [K,G^\alpha] \,. \qquad (5.5)$$

Now it is possible to expand the perturbation V in terms of the operators T^α. By standard technique one gets:

$$W = \sum_\alpha W^\alpha + \frac{1}{2} \sum_{\alpha\neq\beta}\sum \langle [G^\alpha,[K,G^\beta]] \rangle + \text{higher-order terms} \,, \qquad (5.6)$$

$$G = \sum_\alpha G^\alpha + \frac{1}{2}\frac{1}{k} \sum_{\alpha\neq\beta}\sum [G^\alpha, T^\beta + W^\beta] + \text{higher-order terms} \,. \qquad (5.7)$$

Therefore, in first-order approximation, both the level-shift operator W and the generator G of the level-shift transformation of the complicated problem are *additively* composed from those of the simpler subproblems. Up to the second order in energy there are no energy denominators, so that the calculation of the second-order level-shift operator is quite trivial. Attention should be paid to the summation restrictions $\alpha \neq \beta$ which says that the solution of each subproblem alone is already exactly contained in the first-order term. In this formalism, many of the results of infinite-order perturbation series are implicitely contained, but the use of any diagram technique or summation method is completely avoided. Furthermore, we have the advantage that, if we stop the development of the series at any order, the resulting approximative level-shift transformation is always exactly unitary.

REVIEWS OF MODERN PHYSICS VOLUME 35, NUMBER 3 JULY 1963

Error Bounds for Expectation Values*

NORMAN W. BAZLEY
Institut Battelle, Geneva, Switzerland

AND

DAVID W. FOX
Applied Physics Laboratory, The Johns Hopkins University, Silver Spring, Maryland

1. INTRODUCTION

WE consider the estimation of expectation values for stationary states of Schrödinger's equation, particularly for those states having the lowest energies. The paper contains no new mathematics, but discusses the significance of known facts in a physical setting. We denote the Hamiltonian by H, energies by E, and wave functions by ψ; and suppose that the Hamiltonian is bounded below, self-adjoint, and that the lowest energy levels are point eigenvalues of finite degeneracy.[1] Thus, we have

$$H\psi_\nu - E_\nu\psi_\nu = 0 \,, \quad \nu = 1,2,\cdots \,,$$

and

$$E_1 \leq E_2 \leq \cdots \,.$$

Our considerations are formulated within the frame-

* This work was supported in part by the Department of the Navy under Contract NORD 7386.

[1] T. Kato, Trans. Am. Math. Soc. **70**, 195 (1951), has shown that the first two properties are enjoyed by the usual Hamiltonians for atomic and molecular systems. The third has been demonstrated by T. Kato, Trans. Am. Math. Soc. **70**, 212 (1951) for the helium atom and by G. M. Zhilin, Mosk. Mat. Obshch. Tr. **9**, 81 (1960), for many other atomic and molecular systems. Degenerate eigenvalues are to be counted repeatedly.

work of a Hilbert space[2] $\mathfrak{H}$ in which the inner product is designated by (φ,ψ) and the resulting norm by $||\varphi||$. The expectation values have the form $(B\psi_\nu,\psi_\mu)$, in which B is a symmetric (Hermitian) operator. Since, in general, the true wave functions are not known, the best that can be done is to calculate approximations to them, compute the expectation values from the approximate wave functions, and then to estimate bounds on the errors. Thus, what is needed are bounds on the quantities

$$|(B\varphi_\nu,\varphi_\mu) - (B\psi_\nu,\psi_\mu)| \,,$$

in which φ_ν and φ_μ are approximations to ψ_ν and ψ_μ, respectively. Wave functions and approximations to them are assumed to be normalized. The ingredients from which such estimates can be constructed are these: the Hamiltonian H, the operator B, the sense in which φ approximates ψ, and the relationships that may exist between B and H. In our discussion we restrict attention to approximate wave functions that result from a variational calculation (Rayleigh–Ritz) using the Hamiltonian H and a family of trial functions, and show how error bounds can be obtained for expectation values when the operator B is *bounded relative to a norm generated by H*. The error bounds contain only upper and lower bounds for the energies and the constant involved in the bound of B relative to H.

2. EXAMPLES

Some typical examples of expectation operators and brief comments on each of these are given below. The order of listing them is, roughly speaking, that of increasing complexity as is clear from the following section.

A. B Is the Hamiltonian H

Since for a wave function ψ_ν and the corresponding Rayleigh–Ritz vector φ_ν we have[3]

$$(H\varphi_\nu,\varphi_\nu) = E_\nu^{\mathrm{U}} \geq E_\nu = (H\psi_\nu,\psi_\nu) \,,$$

the problem of estimating the error,

$$(H\varphi_\nu,\varphi_\nu) - (H\psi_\nu,\psi_\nu) \,,$$

is exactly the problem of determining a lower bound E_ν^{L} for E_ν.

B. B Is the Projection on a Given Wave Function

Since $B\varphi = (\varphi,\psi_\nu)\psi_\nu$ for any φ in $\mathfrak{H}$, we observe that

$$(B\varphi_\nu,\varphi_\nu) \leq 1 \,, \quad (B\psi_\nu,\psi_\nu) = 1 \,,$$

and

$$0 \leq (B\psi_\nu,\psi_\nu) - (B\varphi_\nu,\varphi_\nu) = 1 - |(\varphi_\nu,\psi_\nu)|^2 \,,$$

so that in this case the problem is that of finding positive lower bounds for $|(\varphi_\nu,\psi_\nu)|^2$. This, in turn, is equivalent to estimating error bounds for $||\varphi_\nu - \psi_\nu||^2$, for we may assume that φ_ν has been adjusted by the introduction of a complex multiplier of magnitude one so that

$$\mathrm{Re}\,(\varphi_\nu,\psi_\nu) = |(\varphi_\nu,\psi_\nu)| = (\varphi_\nu,\psi_\nu) \,,$$

and, hence,

$$||\varphi_\nu - \psi_\nu||^2 = 2[1 - |(\varphi_\nu,\psi_\nu)|] \,.$$

C. Bounded Operators

These are the operators that are defined for all elements φ of $\mathfrak{H}$. For each such B there is a nonnegative real constant c such that

$$|(B\varphi,\varphi)| \leq c(\varphi,\varphi) \tag{1}$$

holds for every φ of $\mathfrak{H}$. In particular, these operators include multiplication by real functions $f(q)$ that satisfy $\int |f(q)|^2 d\tau < \infty$, and hence they include such expectation values as the mean value of $|\psi|^2$ over a fixed neighborhood of a given point.

D. Unbounded Operators

For such operators, no constant c as in (1) can exist. We assume that these operators are defined on at least those functions of $\mathfrak{H}$ for which H is defined. In particular, such operators as momenta ($B = i(\partial/\partial q)$), kinetic energies ($B = -\frac{1}{2}\nabla^2$), and dipole moments ($B = z$) are examples of unbounded operators of frequent occurrence.

3. DISCUSSION[4]

In recent years effective methods for systematic calculation of lower bounds for energies of quantum mechanical Hamiltonians have been developed.[5]

Consequently, whenever B is the Hamiltonian H

---

[2] We have in mind the Hilbert space of square integrable functions over the particle coordinate space.

[3] Here and later we continue to write $(H\varphi,\varphi)$ even for φ's on which H may not be defined; we understand by this the value of the closure of the form generated by H as is usual in the Rayleigh–Ritz procedure.

[4] To avoid complication we will assume that the wave functions for which the expectation values are wanted are nondegenerate.

[5] N. W. Bazley, Phys. Rev. **120**, 144 (1960); J. Math. Mech. **10**, 289 (1961); N. W. Bazley and D. W. Fox, J. Res. Natl. Bur. Std. **65B**, 105 (1961); Phys. Rev. **124**, 483 (1961); J. Math. Phys. **3**, 469 (1962); Arch. Rat. Mech. Anal. **10**, 352 (1962); J. Math. Phys., (to be published). T. Kato, J. Phys. Soc. Japan **4**, 334 (1949); Math. Ann. **126**, 253 (1953). H. F. Weinberger, Tech. Note BN-183, Institute for Fluid Dynamics and Applied Mathematics, University of Maryland (1959) (unpublished).

itself, these may be used to give the needed estimates:

$$0 \leq (H\varphi_\nu,\varphi_\nu) - (H\psi_\nu,\psi_\nu) \leq E_\nu^{\mathrm{U}} - E_\nu^{\mathrm{L}} .$$

We observe that always $(H\varphi_\nu,\varphi_\mu) = (H\psi_\nu,\psi_\mu) = 0$, whenever $\nu \neq \mu$.

When B is the projection on a wave function ψ_ν that is known to be nondegenerate, then estimates can be made in terms of upper and lower bounds to the energies alone. To know that ψ_ν is nondegenerate, it is essential to have sufficiently exact upper and lower bounds so that

$$E_{\nu-1}^{\mathrm{U}} < E_\nu^{\mathrm{L}} \leq E_\nu^{\mathrm{U}} < E_{\nu+1}^{\mathrm{L}} ;$$

further, this information is sufficient to insure that $(B\varphi_\nu,\varphi_\nu)$ is strictly positive,[6] that is, that

$$(B\varphi_\nu,\varphi_\nu) = |(\varphi_\nu,\psi_\nu)|^2 > 0 .$$

The simplest estimate is the well-known Eckart criterion for the ground-state trial function,

$$1 \geqslant |(\varphi_1,\psi_1)|^2 \geqslant 1 - \frac{E_1^{\mathrm{U}} - E_1^{\mathrm{L}}}{E_2^{\mathrm{L}} - E_1^{\mathrm{L}}} . \tag{2}$$

Generalizations of inequalities of this type have been given by Löwdin and Shull[7] and Weinberger.[6] Here we give two special cases of the Weinberger result, which seems to be little known outside of mathematical circles. In the first, we assume that the upper and lower bounds to the energies are sufficiently good to show that the first m states ($m \geq \nu$) are nondegenerate, that is,

$$E_1^{\mathrm{L}} \leq E_1^{\mathrm{U}} < E_2^{\mathrm{L}} \leq E_2^{\mathrm{U}} < \cdots \leq E_m^{\mathrm{U}} < E_{m+1}^{\mathrm{L}} .$$

Then the Weinberger result gives

$$1 \geqslant |(\varphi_\nu,\psi_\nu)|^2 \geqslant \left\{1 - \frac{E_\nu^{\mathrm{U}} - E_\nu^{\mathrm{L}}}{E_{m+1}^{\mathrm{L}} - E_\nu^{\mathrm{L}}}\right\} \times \prod_{\substack{\mu=1\\ \mu\neq\nu}}^{m} \left\{1 - \frac{E_\nu^{\mathrm{U}} - E_\nu^{\mathrm{L}}}{E_\nu^{\mathrm{U}} - E_\mu^{\mathrm{U}}} \cdot \frac{E_\mu^{\mathrm{U}} - E_\mu^{\mathrm{L}}}{E_\nu^{\mathrm{L}} - E_\mu^{\mathrm{L}}}\right\} . \tag{3}$$

If we use only the information

$$E_1^{\mathrm{L}} \leq E_{\nu-1}^{\mathrm{U}} < E_\nu^{\mathrm{L}} \leq E_\nu^{\mathrm{U}} < E_{\nu+1}^{\mathrm{L}} ,$$

then for $\nu > 1$, the Weinberger result gives

$$1 \geqslant |(\varphi_\nu,\psi_\nu)|^2 \geqslant \left\{1 - \frac{E_\nu^{\mathrm{U}} - E_\nu^{\mathrm{L}}}{E_{\nu+1}^{\mathrm{L}} - E_\nu^{\mathrm{L}}}\right\} \times \left\{1 - \frac{E_\nu^{\mathrm{U}} - E_\nu^{\mathrm{L}}}{E_\nu^{\mathrm{U}} - E_{\nu-1}^{\mathrm{U}}} \cdot \frac{E_{\nu-1}^{\mathrm{U}} - E_1^{\mathrm{L}}}{E_\nu^{\mathrm{L}} - E_1^{\mathrm{L}}}\right\} . \tag{4}$$

[6] See H. F. Weinberger, J. Res. Natl. Bur. Std. **64B**, 217 (1960).

[7] P.-O. Löwdin, *Advances in Chemical Physics*, edited by I. Prigogine (Interscience Publishers, Inc., 1959), Vol. 2, p. 207.

For bounded operators, the condition of boundedness,

$$|(B\varphi,\varphi)| \leq c(\varphi,\varphi) ,$$

is equivalent to

$$|(B\varphi_1,\varphi_2)| \leq c||\varphi_1||\cdot||\varphi_2||$$

for every φ_1 and φ_2 of $\mathfrak{H}$. From the identity,

$$(B(\varphi + \psi),\varphi - \psi) = (B\varphi,\varphi) - (B\psi,\psi) + 2i \operatorname{Im} (B\psi,\varphi) ,$$

which leads to the inequality,

$$|(B\varphi,\varphi) - (B\psi,\psi)| \leq |(B(\varphi + \psi),\varphi - \psi)| \leq c||\varphi + \psi||\cdot||\varphi - \psi|| ,$$

we obtain

$$|(B\varphi_\nu,\varphi_\nu) - (B\psi_\nu,\psi_\nu)| \leq c||\varphi_\nu + \psi_\nu||\cdot||\varphi_\nu - \psi_\nu|| \leq 2c||\varphi_\nu - \psi_\nu|| . \tag{5}$$

When we wish an error bound for $|(B\varphi_\nu,\varphi_\mu) - (B\psi_\nu,\psi_\mu)|$, we start from

$$(B\varphi_\nu,\varphi_\mu) - (B\psi_\nu,\psi_\mu) = (B(\varphi_\nu - \psi_\nu),\varphi_\mu) + (B\psi_\nu,\varphi_\mu - \psi_\mu) ,$$

and obtain

$$\begin{aligned} &|(B\varphi_\nu,\varphi_\mu) - (B\psi_\nu,\psi_\mu)| \\ &\leq c\{||\varphi_\mu||\cdot||\varphi_\nu - \psi_\nu|| + ||\psi_\nu||\cdot||\varphi_\mu - \psi_\mu||\} \\ &= c\{||\varphi_\nu - \psi_\nu|| + ||\varphi_\mu - \psi_\mu||\} . \end{aligned} \tag{6}$$

The right side of the inequality (5) coincides with that of (6) when we put $\nu = \mu$. Thus, for bounded expectation values what is needed is a good estimate for the constant c that appears and estimates for $||\varphi - \psi||$, which, as indicated above, can be given in terms of the upper and lower bounds for the energies.

When the operator B that gives rise to the expectation values is not bounded, then it is still possible to make systematic error estimates for expectation values of approximation vectors calculated variationally from the Hamiltonian provided that B *is bounded relative to a norm generated by* H. To explain what this means we introduce an auxiliary bilinear form $[\varphi_1,\varphi_2]$ defined by

$$[\varphi_1,\varphi_2] = (H\varphi_1,\varphi_2) + (\delta - E_1^{\mathrm{L}})(\varphi_1,\varphi_2) , \quad \delta > 0 ,$$

for all vectors φ_1,φ_2 for which H is defined. This form has all of the properties of an inner product, and we may extend it by closure so that it is defined for those vectors φ which can be taken as Rayleigh–Ritz trial vectors for H.[8] We denote the norm induced by this inner product, called a norm generated by H, by $[\varphi]$ so that

$$[\varphi]^2 = [\varphi,\varphi] .$$

The condition that B is to satisfy is that there exists a constant k such that

$$|(B\varphi,\varphi)| \leq k[\varphi,\varphi]$$

for all permissible trial vectors φ. As for the bounded case, this is equivalent to

$$|(B\varphi_1,\varphi_2)| \leq k[\varphi_1][\varphi_2]$$

for all pairs of trial vectors. Following the same lines of argument as before we obtain the inequalities

$$|(B\varphi_\nu,\varphi_\nu) - (B\psi_\nu,\psi_\nu)| \leq k[\varphi_\nu + \psi_\nu]\cdot[\varphi_\nu - \psi_\nu] \tag{7}$$

and

$$|(B\varphi_\nu,\varphi_\mu) - (B\psi_\nu,\psi_\mu)| \leq k\{[\varphi_\mu]\cdot[\varphi_\nu - \psi_\nu] + [\psi_\nu]\cdot[\varphi_\mu - \psi_\mu]\} . \tag{8}$$

All of the quantities on the right can be successfully estimated in terms of upper and lower bounds for the energies. In fact, we find the bounds

$$[\varphi_\mu] = (\delta + E_\mu^{\mathrm{U}} - E_1^{\mathrm{L}})^{\frac{1}{2}} ,$$
$$[\psi_\nu] \leq (\delta + E_\nu^{\mathrm{U}} - E_1^{\mathrm{L}})^{\frac{1}{2}} ,$$

and by use of the triangle inequality

$$[\varphi_\nu + \psi_\nu] \leq 2(\delta + E_\nu^{\mathrm{U}} - E_1^{\mathrm{L}})^{\frac{1}{2}} .$$

With these estimates, (7) coincides with (8) when $\nu = \mu$. Further, the terms involving $[\varphi - \psi]$ become small as the upper and lower bounds for the energies are improved, for we can make the estimate,

$$[\varphi_\nu - \psi_\nu] \leq \{(E_\nu^{\mathrm{U}} - E_\nu^{\mathrm{L}}) + 2(\delta + E_\nu^{\mathrm{U}} - E_1^{\mathrm{L}}) \times [1 - |(\varphi_\nu,\psi_\nu)|]\}^{\frac{1}{2}} ,$$

in which the first term goes to zero as E_ν^{U} and E_ν approach and the second as well by virtue of estimates of the type (2), (3), or (4).[9] It is not difficult to show that the first two examples of unbounded operators that we have given earlier are bounded relative to [ ] and that the third is not.

For those operators that are not bounded relative to [ ] there is no reason to hope that trial functions calculated from the usual variational procedure will lead to good approximations for the expectation values.[10] In fact, if B is not bounded relative to [ ], then for each ψ there exist sequences $\{\varphi^n\}$ of normalized Rayleigh–Ritz approximations for ψ such that

$$[\varphi^n - \psi] \rightarrow 0 ,$$

that is,

$$||\varphi^n - \psi|| \rightarrow 0 , \qquad (H\varphi^n , \varphi^n) = E^{Un} \rightarrow E ,$$

and at the same time

$$|(B\varphi^n,\varphi^n)| \rightarrow \infty .$$

Thus, apparently very good variational approximations for ψ may lead to outrageous errors in supposed approximations for such expectation values.

[8] In fact, the domain $\mathfrak{D}_{[\,]}$ of this new inner product is itself a Hilbert space contained in $\mathfrak{H}$. It is precisely the space in which an infinite family of trial vectors must be complete in order to assure convergence of the Rayleigh–Ritz variational procedure.

[9] We do not consider here the questions of how the constants k and δ which appear may be optimized.

[10] It may be possible, nonetheless, to show that in special cases the wave functions to be approximated share some particular properties, for example, uniform asymptotic behavior at infinity, and that on suitably chosen classes of trial functions that share the same properties, the boundedness of B relative to [ ] holds.

Discussion on Upper and Lower Bounds for Expectation Values

A. WEINSTEIN, *Chairman*

LÖWDIN: Fox's contribution deals certainly with one of the most crucial problems in modern quantum theory of matter, namely the reliability of the calculations we can carry out by limiting procedures. So far, we have often had many equivalent formulas to consider, but it seems now as if we could use with confidence only those where the operators involved are bounded with respect to the Hamiltonian.

Our group has been very interested in the same problems, and there is a rather close analogy between the ideas you have presented so nicely here and my remarks in connection with the discussion of spin densities. Perhaps I may also refer to the papers quoted there for the identities we have used to estimate the errors in the expectation values of physical quantities other than the energy.

T. C. CHEN: The unboundedness of the expectation error can perhaps be removed by evaluating