Generalized Perturbation Theory in Operator Form

H. PRIMAS

Swiss Federal Institute of Technology, Laboratory of Physical Chemistry, Switzerland

1. INTRODUCTION

We propose to formulate any quantum-mechanical perturbation theory in operator form and to take advantage of the fact that such a theory can be completely formulated in the domain of a Lie algebra. Further we generalize the concept of a T-operator and use this theory to compose a solution of a complicated quantum-mechanical problem out from simpler subproblems connected therewith. For the sake of simplicity, we assume that the Hamiltonian is an operator in a finite dimensional space. Let the Hamiltonian H be split in an arbitrary way in two Hermitian operators K and V,

$$H = K + V\,. \tag{1.1}$$

The operator K should describe a simple, but physically meaningful system and we consider V as perturbation of this system, but it is not assumed that V should be "small" in whatever sense. The unperturbed operator K may have any degeneracy of its eigenvalues, and we will assume that the spectral resolution of K is explicitly known,

$$K = \sum \kappa_n K_n \quad \text{with } K_n K_m = \delta_{nm} K_n\,, \quad \sum K_n = 1$$
$$\kappa_n \neq \kappa_m \quad \text{for } n \neq m \quad \text{(by definition)} \tag{1.2}$$
$$\mathrm{tr}\,(K_n) = \text{multiplicity of } \kappa_n\,.$$

Any perturbation theory can be formulated in the domain of Lie algebra generated by the operators K and V (i.e., in the domain of all operators K; V; $[K,V]$; $\{K,[K,V]\}$; $\{V,[K,V]\}$; $\{[K,V],[K,V]\}$, etc.) The solution of a perturbation problem is closely connected with the solution of commutator equations of the type

$$[K,X] = Y\,. \tag{1.3}$$

Using the spectral resolution one gets immediately the identity

$$X - \sum_n K_n X K_n = \sum_{n\neq m}\sum (\kappa_n - \kappa_m)^{-1} K_n Y K_m \tag{1.4}$$

the result of which can be written more conveniently with the aid of super-operators.

2. SUPER-OPERATORS[1]

Let $\mathfrak{A}$ be the operator algebra of all linear operators $A,B,C,\ldots$ over the state space $\mathfrak{U}$ of quantum mechanics (which we assumed to be of finite dimension). The operator algebra $\mathfrak{A}$ is a vector space over which we can again define operators that we call super-operators. So a super-operator is a function with the operator algebra $\mathfrak{A}$ as its domain and its range contained in $\mathfrak{A}$. From Eq. (1.4) it can be seen that two super-operators are of special importance to perturbation theory. First, the super-operator, which projects from any operator that part which commutes with K, we define as

$$\langle X\rangle = \sum_n K_n X K_n \quad \text{for any } X \in \mathfrak{A} \tag{2.1}$$

and call the operator $\langle X\rangle$ the "diagonal part" (relative to K) of the operator X. If all eigenvalues of K are different, the matrix representation of $\langle X\rangle$ is actually diagonal in the eigenbasis of K. The super-operator, which forms the commutator from any operator with K, will be denoted by k,

$$k(X) = [K,X] \quad \text{for any } X \in \mathfrak{A}\,, \tag{2.2}$$

and we call k the "derivation super-operator generated by K." A derivation super-operator has no unique inverse, but we can define a special inverse of k that we denote by $\frac{1}{k}$ (in contrast to the general k^{-1}). The super-operator $\frac{1}{k}$ has the whole algebra as its domain and is defined by

$$\frac{1}{k}(X) = \sum_{n\neq m}\sum (\kappa_n - \kappa_m)^{-1} K_n X K_m \quad \text{for any } X \in \mathfrak{A} \tag{2.3}$$

and is in contrast to other inverses characterized by

$$\left\langle \frac{1}{k}(X) \right\rangle = 0 \quad \text{for any } X \in \mathfrak{A}\,. \tag{2.4}$$

The general solution of the commutator Eq. (1.3) is

$$X = \frac{1}{k}(Y) + \langle A\rangle, \text{ where } A \text{ is an arbitrary operator}\,. \tag{2.5}$$

[1] A fuller account of the theory of super-operators can be found in: M. Rosenblum, Duke Math. J. **23**, 263 (1956); J. A. Crawford, Nuovo Cimento **10**, 698 (1958); A. Meckler, Nuovo Cimento Suppl. **12**, 1 (1959); G. Lumer and M. Rosenblum, Proc. Am. Math. Soc. **10**, 32 (1959); J. Schwinger, Proc. Nat. Acad. Sci. **46**, 257, 570 (1960); H. Primas, Helv. Phys. Acta **34**, 331 (1961); C. N. Banwell and H. Primas, Mol. Phys. (to be published).

The Lie character of the functions $\frac{1}{k}(\cdots)$ and $\langle\cdots\rangle$ can be set in evidence by the relations

$$\frac{1}{k}(X) = i \lim_{\alpha\to 0} \int_0^\infty e^{-\alpha t} e^{-iKt}\{X - \langle X\rangle\} e^{iKt} dt\,, \qquad (2.6)$$

$$\langle X\rangle = \lim_{\alpha\to 0} \int_0^\infty e^{-\alpha t} e^{-iKt} X e^{iKt} dt\,, \qquad (2.7)$$

which are often convenient for practical computations [compare also Eqs. (4.1) and (4.2)].

3. LEVEL SHIFT REPRESENTATION[2,3]

A convenient way to formulate any approximative perturbation theory is to perform a unitary transformation of the Hamiltonian in such a way that the transformed Hamiltonian commutes with K,

$$UHU^\dagger = \tilde{H} \quad \text{such that } [K,\tilde{H}] = 0\,, \qquad (3.1)$$
$$UU^\dagger = U^\dagger U = 1\,.$$

It is of great advantage to preserve the Lie structure of the problem and for this we set

$$U = \exp G \quad \text{with } G = -G^\dagger\,. \qquad (3.2)$$

This problem is uniquely defined only if we normalize U in some way, and it is convenient to require a vanishing diagonal part of G, in respect to K,

$$\langle G\rangle = 0\,. \qquad (3.3)$$

Splitting off the trivial part K we get from Eq. (3.1):

$$e^G(K+V)e^{-G} = K + W \quad \text{with } [K,W] = 0 \qquad (3.4)$$
$$\text{and } \langle G\rangle = 0\,.$$

We call this well defined problem the level-shift transformation, G the generator of this transformation, and W level-shift operator. If E_n is a nondegenerated eigenvalue of H and if λ_n and φ_n are the nth eigenvalue and eigenfunction of K, then we have the exact relation

$$E_n = \lambda_n + (\varphi_n|W|\varphi_n)\,,$$

so that in this case W gives actually the exact level shift of the energy by switching on the perturbation.

4. LIE STRUCTURE OF PERTURBATION THEORY[3,4]

Many of the difficulties connected with the various perturbation methods can be avoided by confining the domain of the operator algebra to the Lie algebra generated by the operators K and V. The Hausdorff formula[5]

$$e^g(H) = H + \frac{1}{1!}\,g(H) + \frac{1}{2!}\,g^2(H) + \cdots$$
$$\text{with } g(X) = [G,X]\,, \qquad (4.1)$$

or written more conventionally without super-operators

$$e^G H e^{-G} = H + \frac{1}{1!}\,[G,H] + \frac{1}{2!}\,[G,[G,H]] + \cdots \qquad (4.2)$$

shows clearly that the problem Eq. (3.4) can be written completely in the domain of a Lie algebra. In the same sense that we can formulate a problem of real analysis in the field of complex numbers, we are allowed to go outside of the domain of the Lie algebra. This is usually done and perfectly all right, but one should not wonder about the fact that in the final expressions there are such things as unlinked clusters that all cancel[6] much in the same way that all imaginary terms cancel in the final result of a real problem, formulated in the field of complex numbers. The Rayleigh–Schrödinger perturbation theory gives the best known example. If we expand both sides of Eq. (3.4) in powers of V and use Eq. (4.2) we get:

$$W = \langle V\rangle + \frac{1}{2}\left\langle\left[\frac{1}{k}(V),V\right]\right\rangle + \cdots\,, \qquad (4.3)$$

$$G = \frac{1}{k}(V) + \frac{1}{2}\frac{1}{k}\left[\frac{1}{k}(V),V + \langle V\rangle\right] + \cdots\,. \qquad (4.4)$$

Usually one writes Eq. (4.3) in the identical form

$$W = \langle V\rangle + \left\langle V\,\frac{1}{k}(V)\right\rangle + \cdots\,. \qquad (4.5)$$

Because the product of the two operators V and $\frac{1}{k}(V)$ is *not* within the Lie algebra, there are terms which cancel if the diagonal part of $V\,\frac{1}{k}\,V$ is formed. This annoying canceling is avoided in the formulation Eq. (4.3) from the outset.

5. GENERALIZED T-OPERATOR, DECOMPOSITION OF A COMPLICATED PROBLEM IN SIMPLER SUBPROBLEMS[3]

In order to be able to handle singular potentials in perturbation theory, Watson[7] introduced a t-operator whose diagonal part is our level-shift operator W. We

---

[2] S. Tani, Progr. Theoret. Phys. (Kyoto) **11**, 190 (1959).
[3] H. Primas, Helv. Phys. Acta **34**, 331 (1961).
[4] F. J. Murray, J. Math. Phys. **3**, 451 (1962).
[5] F. Hausdorff, Leipziger Ber. Ges. Wiss., math., phys. Kl. **58**, 19 (1906).
[6] K. A. Brückner, Phys. Rev. **100**, 36 (1955); J. Goldstone, Proc. Roy. Soc. (London) **A239**, 267 (1957); F. Coester, Nucl. Phys. **7**, 421 (1958).
[7] K. M. Watson, Phys. Rev. **89**, 575 (1953).

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

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

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.