Method of Moments in Quantum Mechanics*†

FRANCIS R. HALPERN‡
Radiation Laboratory, University of California, Berkeley, California
(Received April 29, 1957)

An approximation technique for quantum-mechanical problems based on the expectation values of the powers of the Hamiltonian is developed. The mathematical foundations on which this approach is based are the method of moments employed in probability theory, and the theory of orthogonal polynomials. In practice this method constitutes an extension of the Rayleigh-Ritz principle and gives a systematic method of improving the trial function.

I. INTRODUCTION

THE general problem in quantum mechanics is the diagonalization of the Hamiltonian operator. In the method of moments the procedure is to assume that an arbitrarily chosen state vector is expanded in the complete set, the members of which are the eigenfunctions of the Hamiltonian. The problem is then to determine the unknown eigenfunctions occurring in this expansion, their eigenvalues, and the associated expansion coefficients. To achieve this end, a step-by-step method is prescribed that will remove all but one term in the expansion.

It is not necessary in the course of this calculation to assume any separation of the Hamiltonian into perturbed and unperturbed portions. It will frequently be convenient, however, from the computational point of view, to take the initial vector to be an eigenfunction of a portion of the Hamiltonian. The basic numerical quantities that enter into a calculation are the matrix elements of the powers of the Hamiltonian in the chosen initial state. Because these quantities are simple to calculate, it is feasible to carry the calculations to quite high orders.

II. DESCRIPTION OF THE METHOD

The system to be treated is described by a Hamiltonian H. The system is assumed to be enclosed in a box in order to assure that the state vectors occurring are normalizable to unity. The normalized eigenfunctions of the Hamiltonian are $\psi(E_i,a_j)$. The a's are the eigenvalues of additional operators A that commute with the Hamiltonian and that are necessary to completely describe the states. The ψ's satisfy the equations

$$H\psi(E_i,a_j)=E_i\psi(E_i,a_j),$$
$$A\psi(E_i,a_j)=a_j\psi(E_i,a_j).$$

The ψ's and the E_i's are of course unknown, although the a_j's will in general be known. To proceed it is necessary to choose a trial function ϕ. It is most convenient if ϕ is an eigenfunction of the auxiliary variables A. In this event the expansion of ϕ in the set of ψ's is

$$\phi=\sum\alpha_i\psi(E_i). \tag{1}$$

The dependence of the ψ's on the a's has been dropped, as ϕ is an eigenfunction of the A's and only one set of a's can appear for each E.

It is convenient to introduce the function

$$F_\phi(E)=\sum_{E_i\leq E}|\alpha_i|^2 \tag{2}$$

associated with this expansion. $F_\phi(E)$ is a nondecreasing function of bounded variation that vanishes for sufficiently small E. The first property holds because $F_\phi(E)$ changes only by amounts $|\alpha_i|^2$ at the points E_i. The second property is a result of taking ϕ to be normalizable. The eigenvalues E_i of H describe the allowable energy levels of a physical system, therefore there must be a smallest one and hence a smallest point of increase for $F_\phi(E)$.

The function $F_\phi(E)$ contains essentially all the information about the physical system. The eigenvalues are immediately evident as points of discontinuity of F and the eigenfunction that belongs to E_k is given by the formula

$$\psi(E_k)=\prod_{i\neq k}(H-E_i)\phi.$$

The matrix element of any function of H in the state ϕ is given by the expression

$$(\phi|G(H)|\phi)=\int G(E)dF_\phi(E). \tag{3}$$

Thus the quantum-mechanical problem is equivalent to determining the function F.

The function F is a probability distribution function, and there are procedures for the determination of such a function. The method to be employed is called the method of moments, and consists in developing an approximating sequence of functions $F_\phi^{(n)}(E)$ that it is hoped will converge to $F_\phi(E)$. The quantities employed to compute $F_\phi^{(n)}(E)$ are the moments H_n of the distribution $F_\phi(E)$. These are defined by

$$H_n=(\phi|H^n|\phi)=\int E^n dF_\phi(E). \tag{4}$$

* Based on a dissertation submitted in partial fulfillment of the requirements for the Ph.D. degree at the University of California.
† This work was performed under the auspices of the U. S. Atomic Energy Commission.
‡ Present address: Palmer Physical Laboratory, Princeton University, Princeton, New Jersey.

There is a difficulty in principle that should be noted at this point. The moments may not uniquely determine the distribution. A very simple way in which this may occur is if some of them are infinite.

The requirement on the approximating function $F_\phi^{(n)}(E)$ is that it be a step function with n points of increase such that its first $2n$ moments agree with those of $F_\phi(E)$. The nth approximating function depends on $2n$ numbers; the n values of the argument $\epsilon_1^{(n)}, \epsilon_2^{(n)}, \cdots, \epsilon_n^{(n)}$, at which $F_\phi^{(n)}(E)$ is discontinuous and the n real positive numbers $b_1^{(n)}, b_2^{(n)}, \cdots, b_n^{(n)}$, which are the magnitudes of the discontinuities. These numbers must satisfy the $2n$ relations

$$H_k=\int E^k dF_\phi(E)=\int E^k dF_\phi^{(n)}(E)=\sum \epsilon_i^{(n)k} b_i^{(n)} \qquad (k=0, 1, 2, \cdots, 2n-1).$$

It has been shown that the following prescription gives the unique function satisfying these requirements.[1] The polynomial $P_n(E)$ defined by

$$P_n(E)=\begin{vmatrix} 1 & E & E^2 & \cdots & E^n \\ H_0 & H_1 & H_2 & \cdots & H_n \\ H_1 & H_2 & H_3 & \cdots & H_{n+1} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ H_{n-1} & \cdot & \cdot & \cdots & H_{2n-1} \end{vmatrix}$$

is constructed. The polynomial $P_n(E)$ has n real distinct roots. These are the correct values for the $\epsilon^{(n)}$'s. It is now possible to solve the n linear equations

$$H_k=\sum_{i=1}^{n} \epsilon_i^{(n)k} b_i^{(n)} \quad (k=0, 1, 2, \cdots, n-1)$$

for the $b^{(n)}$'s.

The approximating function $F_\phi^{(n)}(E)$ is now completely determined. It has the following useful properties. All the $\epsilon_i^{(n)}$ lie between the greatest and least points of increase of $F_\phi(E)$, or the smallest $\epsilon_i^{(n)}$ is an upper bound for the lowest eigenvalue of H and the largest $\epsilon_i^{(n)}$ is a lower bound for the largest eigenvalue. Between any two points of increase of $F_\phi^n(E)$ there is at least one point of increase of $F_\phi(E)$, and at each point of increase ϵ_i^n the inequality

$$F_\phi^{(n)}(\epsilon_i^-)\leq F_\phi(\epsilon_i^{(n)})\leq F_\phi^{(n)}(\epsilon_i^+)$$

holds. Thus there is an eigenvalue of H between any successive ϵ's. These remarks also apply if $F_\phi(E)$ is replaced by $F_\phi^{(N)}(E)$ for $n<N$.

The lowest-order approximation is given by $F_\phi^{(1)}(E)$. This function has a single point of increase given by the root of the equation

$$P_1(E)=\begin{vmatrix} 1 & E \\ H_0 & H_1 \end{vmatrix}=0.$$

This solution is

$$\epsilon^{(1)}=H_1/H_0,$$

and $b^{(1)}$ is just H_0.

The inequality stated above is to this order the Rayleigh-Ritz principle. That is, the smallest eigenvalue of H, $E_{\min}$ satisfies the inequality

$$E_{\min}\leq E=H_1/H_0.$$

A slightly different approach is also possible. From the set of functions 1, E, E^2, $\cdots$, it is possible by the usual Schmidt orthogonalization procedure to construct a set of polynomials $p_n(E)$ with the properties

$$\int p_n(E)p_{n'}(E)dF_\phi(\epsilon)=\delta_{nn'}.$$

Except for a normalization constant we have $p_n=P_n$, and the determinantal form is just a convenient method of writing the Schmidt process. Most of the results quoted above then follow from the properties of the roots of sets of orthogonal polynomials and their associated distribution.[2]

If all that is desired is an estimate of the eigenvalues and their spacing, it is sufficient to calculate the roots of the determinants. If wave functions are desired, then appropriate polynomials in the Hamiltonian of the form $\prod(H-\epsilon_i^{(n)})$ are used to operate on the initial vector.

It is possible to construct n vectors $\psi^{(n)}$ in this form. The $\psi^{(n)}$ are given by

$$\psi_k^{(n)}=\prod_{i\neq k} (H-\epsilon_i^{(n)})\phi.$$

The vectors $\psi^{(n)}$ are orthogonal,

$$(\psi_k^{(n)},\psi_{k'}^{(n)})=\delta_{kk'}(\psi_k^{(n)},\psi_k^{(n)}).$$

This result may be proved by expressing the matrix element in terms of the original vector ϕ:

$$(\psi_k^{(n)},\psi_{k'}^{(n)})=(\phi|\prod_{i\neq k} (H-\epsilon_i^{(n)}) \prod_{i\neq k'} (H-\epsilon_i^{(n)})|\phi).$$

The properties of the energy-distribution function $F_\phi(E)$ are employed to write this expression as an integral,

$$(\psi_k^{(n)},\psi_{k'}^{(n)}) = \begin{cases} \int \prod (E-\epsilon_i^{(n)}) \prod_{i\neq k,k'} (E-\epsilon_i^{(n)})dF_\phi(E)=0, & k\neq k' \\ \int \prod_{i\neq k} (E-\epsilon_i^{(n)})^2 dF_\phi(E)>0, & k=k' \end{cases}$$

[1] J. V. Uspensky, *An Introduction to Mathematical Probability* (McGraw-Hill Book Company, Inc., New York, 1937). The theorem stated above is a trivial generalization of those proved in this book. References are given to the original work here and in reference 2.

[2] G. Szegö, *Orthogonal Polynomials* (American Mathematical Society Colloquium Publications, New York, 1939), Vol. 23.

where the usual properties of orthogonal functions have been used.

A similar argument is used to compute the expectation value of H in the state $\psi_k^{(n)}$:

$$(\psi_k,H\psi_k)=(\phi|\prod_{i\neq k}(H-\epsilon_i^{(n)})H\prod_{i\neq k}(H-\epsilon_i^{(n)})|\phi)$$

$$=\int\prod_{i\neq k}(E-\epsilon_i^{(n)})^2(E-\epsilon_k+\epsilon_k)dF_\phi(E)$$

$$=\epsilon_K^{(n)}\int\prod_{i\neq k}(E-\epsilon_i^{(n)})^2dF_\phi(E)$$

$$=\epsilon_k^{(n)}(\psi_k^{(n)},\psi_k^{(n)}).$$

To test the adequacy of the state $\psi_k^{(n)}$ as approximations to the eigenfunctions of H, the quantity σ defined by

$$\sigma_k^{(n)}=\left((\psi_k,H^2\psi_k)-\frac{\epsilon_k^2}{(\psi_k,\psi_k)}\right)^{\frac{1}{2}}$$

is introduced. Here $\sigma_k^{(n)}$ is always non-negative and vanishes only for an eigenfunction.

III. EXAMPLE AND CONCLUSIONS

A simple example of this technique is the following application to the problem of neutral scalar mesons interacting with static nucleons. The notation is the same as that by Wentzel.[3] A cut-off K is employed, since the moments are divergent. The first four moments of H in the state $\phi_0(a_k\phi_0=0)$ are

$$H_0=1,$$

$$H_1=0,$$

$$H_2=\frac{g^2}{(2\pi)^2}\left[\frac{K\omega_K}{2}-\frac{\mu^2}{2}\ln\left(\frac{K+\omega_K}{\mu}\right)\right],$$

$$H_3=\frac{g^2}{(2\pi)^2}\left(\frac{K^3}{3}\right).$$

In the linear Rayleigh-Ritz approximation there is no change from the noninteracting system. The quadratic approximation is determined by solving the equation

$$\begin{vmatrix}1 & E & E^2\\ 1 & 0 & \frac{g^2}{(2\pi)^2}\left[\frac{1}{2}K\omega_K-\mu^2\ln\left(\frac{K+\omega_K}{\mu}\right)\right]\\ 0 & \frac{g^2}{(2\pi)^2}\left[\frac{1}{2}K\omega_K-\mu^2\ln\left(\frac{K+\omega_K}{\mu}\right)\right] & \frac{g^2}{(2\pi)^2}\frac{K^3}{3}\end{vmatrix}=0.$$

The solution of this equation is

$$E=\frac{\frac{g^2}{(2\pi)^2}\left(\frac{K^3}{3}\right)\pm\left\{\left[\frac{g^2}{(2\pi)^2}\left(\frac{K^3}{3}\right)\right]^2+\left(\frac{g^2}{(2\pi)^2}\left[\frac{K\omega_K}{2}-\mu^2\ln\left(\frac{K+\omega_K}{\mu}\right)\right]\right)^3\right\}^{\frac{1}{2}}}{2\left\{\frac{g^2}{(2\pi)^2}\left[\frac{K\omega_K}{2}-\mu^2\ln\left(\frac{K+\omega_K}{\mu}\right)\right]\right\}}.$$

For large K, where the logarithmic terms and the difference between K and ω_K can be neglected, this becomes

$$E=\tfrac{1}{3}K\left[1\pm\left(1+\frac{9}{8}\frac{g^2}{(2\pi)^2}\right)^{\frac{1}{2}}\right].$$

The negative root diverges linearly to $-\infty$ just as the correct value of the self-energy does. The constant $1-[1+(9/8)g^2/(2\pi)^2]^{\frac{1}{2}}$ is smaller than the correct value.

The advantages of this method are that it is independent of the magnitude of the interactions, and the basic quantities are relatively simple to compute. The chief shortcoming is the requirement that a sufficiently good initial state ϕ be chosen so that all the moments are finite. A detailed calculation of the phonon-polaron problem is being carried out and will be published shortly.

[3] G. Wentzel, *Quantum Theory of Fields* (Interscience Publishers, Inc., New York, 1949).