Use of a General Virial Theorem with Perturbation Theory

W. J. CARR, JR.
Westinghouse Research Laboratories, Pittsburgh, Pennsylvania
(Received January 3, 1957)

If the Hamiltonian of a system is broken up into a series of homogeneous operators in the space coordinates, the sum of their mean values multiplied by the degree of homogeneity equals zero for a bound state.

Assume that the Hamiltonian contains three groups of homogeneous terms: for example, kinetic energy, Coulomb energy, and a third interaction, U, which is small. By the "virial" theorem, the energy can be expressed in the mean values of any two of these terms. From perturbation theory, another relationship is derived which allows various orders of the energy to be obtained from the mean value of any one of the three terms of the Hamiltonian.

The perturbation relation is $\bar{U}=\sum nE_n$, where E_n is the nth order correction to the energy, defined by $E=\sum E_n$.

INTRODUCTION

FOR a bound system of particles having kinetic energy T and potential energy V, it is well known that the virial theorem

$$\bar{T}=\tfrac{1}{2}\sum_j\langle\mathbf{r}_j\cdot\nabla_j V\rangle_{\text{Av}}, \tag{1}$$

holds both in classical and quantum mechanics. In the former case the bar indicates a time average, while in the latter a quantum mechanical average. The term on the right, where $\mathbf{r}_j$ is the position coordinate of the jth particle, and ∇_j the gradient operator, is the virial of Clausius for the special case where the force can be described by a potential gradient. This term by Euler's theorem, becomes equal to $-\bar{V}/2$ if the particles interact only via Coulomb forces. Then

$$2\bar{T}+\bar{V}=0. \tag{2}$$

Hereafter, the symbol V will be reserved for Coulomb interactions.

A variety of proofs of the quantum virial law has been given by various authors,[1–5] and also the relativistic analog for a Dirac particle has been derived.[3,6,7] In the following, a relation corresponding to the virial theorem for a more general Hamiltonian is considered, along with some applications from the use of perturbation theory

It will be assumed for simplicity that the particles involved are subject to no constraints not included in the Hamiltonian. Of particular interest is a many-body system of electrons and nuclei, which will be considered first as a concrete example. For most purposes the Hamiltonian, in the absence of an applied field, is given to sufficient accuracy by

$$H=T+V+U_{\text{s-o}}+U_{\text{s-s}}, \tag{3}$$

where $U_{\text{s-o}}$ is the spin-orbit interaction and $U_{\text{s-s}}$ the spin-spin term.

If $\mathbf{r}_{ij}=\mathbf{r}_i-\mathbf{r}_j$, c is the velocity of light, and e_i, m_i, $\boldsymbol{\mu}_i$, and $\mathbf{p}_i$ are respectively the charge, mass, spin magnetic moment operator, and momentum operator of the ith particle,

$$U_{\text{s-o}}=\sum_{i\neq j}\sum\frac{e_j\boldsymbol{\mu}_i}{c}\cdot\frac{\mathbf{r}_{ij}}{r_{ij}^3}\times\left(\frac{\mathbf{p}_j}{m_j}-\frac{\mathbf{p}_i}{2m_i}\right), \tag{4}$$

$$U_{\text{s-s}}=\tfrac{1}{2}\sum_{i\neq j}\sum\frac{1}{r_{ij}^3}\left[\boldsymbol{\mu}_i\cdot\boldsymbol{\mu}_j-\frac{3(\boldsymbol{\mu}_i\cdot\mathbf{r}_{ij})(\boldsymbol{\mu}_j\cdot\mathbf{r}_{ij})}{r_{ij}^2}\right]. \tag{5}$$

It will be shown that for the Hamiltonian (3), the virial law becomes

$$2\bar{T}+\bar{V}+3\bar{U}=0, \tag{6}$$

where $U=U_{\text{s-o}}+U_{\text{s-s}}$.

Thus the energy

$$E=\bar{T}+\bar{V}+\bar{U} \tag{7}$$

can be expressed in terms of any two of the mean values $\bar{T}$, $\bar{V}$, or $\bar{U}$. By treating U as a perturbation, the various orders of E can be written in terms of the various orders of $\bar{T}$, $\bar{V}$, or $\bar{U}$ alone.

PROOF OF EQUATION (6)

The proof of (6) easily may be obtained from the simple variational argument of Fock.[3] One difficulty in this proof, however, is in knowing under what circumstances, as regards to boundary conditions, the variational argument holds. An alternative proof that is more general in application is given below.

Consider the commutator of the operator $\sum_j\mathbf{r}_j\cdot\nabla_j$ with an operator

$$F_n=f_{n+m}(x_1,y_1,z_1\cdots x_N,y_N,z_N)\frac{\partial^m}{\partial x_i^m}, \tag{8}$$

where f_{n+m} is a homogeneous function of degree $n+m$ in the position coordinates of the N particles. The operator

[1] Born, Heisenberg, and Jordan, Z. Physik **35**, 557 (1926).
[2] B. N. Finkelstein, Z. Physik **50**, 293 (1928).
[3] V. Fock, Z. Physik **63**, 855 (1930).
[4] J. C. Slater, J. Chem. Phys. **1**, 687 (1933).
[5] T. L. Cottrell and S. Paterson, Phil. Mag. **42**, 391 (1951).
[6] M. E. Rose and T. A. Welton, Phys. Rev. **86**, 432 (1952).
[7] N. H. March, Phys. Rev. **92**, 481 (1953).

F_n is homogeneous of degree n, in the sense that

$$f_{n+m}(\lambda x_1,\lambda y_1,\lambda z_1\cdots\lambda x_N,\lambda y_N,\lambda z_N)\frac{\partial^m}{\partial(\lambda x_i)^m}$$
$$=\lambda^n f_{n+m}(x_1,y_1,z_1\cdots x_N,y_N,z_N)\frac{\partial^m}{\partial {x_i}^m}. \quad (9)$$

Since by Euler's theorem, $\sum_j \mathbf{r}_j\cdot\nabla_j$ operating on f_{n+m} gives $(n+m)f_{n+m}$, it follows that

$$\sum_j \mathbf{r}_j\cdot\nabla_j F_n=(n+m)F_n+f_{n+m}\sum_j \mathbf{r}_j\cdot\nabla_j\frac{\partial^m}{\partial {x_i}^m}$$
$$=(n+m)F_n+f_{n+m}\left[\mathbf{r}_i\cdot\nabla_i\frac{\partial^m}{\partial {x_i}^m}+\frac{\partial^m}{\partial {x_i}^m}(\sum_j \mathbf{r}_j\cdot\nabla_j-\mathbf{r}_i\cdot\nabla_i)\right]. \quad (10)$$

Therefore, the commutator can be expressed

$$[\sum_j \mathbf{r}_j\cdot\nabla_j, F_n]=(n+m)F_n+f_{n+m}\left[\mathbf{r}_i\cdot\nabla_i,\frac{\partial^m}{\partial {x_i}^m}\right]$$
$$=(n+m)F_n+f_{n+m}\left[x_i\frac{\partial}{\partial x_i},\frac{\partial^m}{\partial {x_i}^m}\right]$$
$$=(n+m)F_n-f_{n+m}m\frac{\partial^m}{\partial {x_i}^m}$$
$$=nF_n. \quad (11)$$

By the same kind of argument, this result also can be proved for operators of the type $(\partial^m/\partial {x_i}^m)f_{n+m}$ or $f_{n+m-p}(\partial^m/\partial {x_i}^m)f_p'$.

All the terms of the Hamiltonian of interest here can be expressed as linear combinations in terms of the type (8). It will be observed that T, V, and U are homogeneous of degree -2, -1, and -3, respectively.

Thus, for the Hamiltonian (3)

$$[\textstyle\sum_j \mathbf{r}_j\cdot\nabla_j, H]=-2T-V-3U. \quad (12)$$

It is worth noting that (12) is simply a relationship among operators, and free, at this point, from the criticism[5] of previous types of proof[2,4] involving partial integrations over wave functions.

To obtain (6) for a bound state, the mean value of both sides of (12) is evaluated for an eigenfunction ψ of H. The left hand side is

$$\langle[\textstyle\sum_j \mathbf{r}_j\cdot\nabla_j, H]\rangle_{Av}=E\langle\sum_j \mathbf{r}_j\cdot\nabla_j\rangle_{Av}$$
$$-\int\psi^* H\textstyle\sum_j \mathbf{r}_j\cdot\nabla_j\psi d\tau, \quad (13)$$

which, because of the Hermitian property of H, is zero.[8,9]

The case of an arbitrary Hamiltonian may be treated in a similar manner by breaking it up into homogeneous parts. The extension of the theorem to include an applied electromagnetic field involves calculating the commutator of $\sum_j \mathbf{r}_j\cdot\nabla_j$ with the additional terms in the Hamiltonian

$$-\sum_i\frac{e_i}{2m_ic}(\mathbf{p}_i\cdot\mathbf{A}_i+\mathbf{A}_i\cdot\mathbf{p}_i)$$
$$+\sum_i\frac{e_i^2}{2m_ic^2}A_i^2-\sum_i \mathbf{u}_i\cdot\mathrm{curl}\mathbf{A}_i+\sum_i e_i\varphi_i, \quad (15)$$

and depends upon the manner in which the vector and scalar potentials, $\mathbf{A}$ and φ, vary with position. For uniform static electric and magnetic fields, $\mathbf{A}$ and φ are homogeneous of degree one. Thus for this case of uniform fields, to the right side of (12) must be added

$$2\sum_i\frac{e_i^2}{2m_ic^2}A_i^2+\sum_i e_i\varphi_i, \quad (16)$$

where $\mathbf{A}_i=\frac{1}{2}\mathfrak{H}\times\mathbf{r}_i$ and $\varphi_i=\mathfrak{E}\cdot\mathbf{r}_i$, $\mathfrak{H}$ and $\mathfrak{E}$ being the magnetic and electric fields.

USE OF PERTURBATION THEORY

Both U of (4) and (5) and the field-dependent terms, (15), in the Hamiltonian ordinarily may be treated as small perturbations compared with $T+V$. In the following example consider the Hamiltonian (3) with U the perturbing term. Let the wave function and energy be written as

$$\psi=\textstyle\sum_n\lambda^n\psi_n,$$
$$E=\textstyle\sum_n\lambda^n E_n, \quad (17)$$

where $\lambda=1$ and the different orders of approximation are given, as usual, by the coefficients of powers of λ.

In the standard Rayleigh-Schrödinger perturbation treatment, results for E_n are given in terms of matrix elements of U with eigenfunctions of the unperturbed Hamiltonian. These results are obtained by starting with the Schrödinger equation.

$$(T+V+\lambda U)\textstyle\sum_n\lambda^n\psi_n=\sum_n\sum_m\lambda^{n+m}E_n\psi_m, \quad (18)$$

and expanding each order, ψ_i, in the unperturbed eigenfunctions. It is shown in the Appendix that the rather complicated results of this treatment are expressed very

[8] For states which are not bound, $\mathbf{r}_j\cdot\nabla_j$ becomes infinite as $r_j\to\infty$ and (13) is not necessarily zero. The case of box normalization, not considered here, is an example of an external constraint, and leads to a term in (13) involving pressure.

[9] From (12), a relation also can be written for the off-diagonal matrix elements. In a representation in which H is diagonal,

$$(E_\alpha-E_\beta)(\textstyle\sum_j \mathbf{r}_j\cdot\nabla_j)_{\alpha\beta}=2T_{\alpha\beta}+V_{\alpha\beta}+3U_{\alpha\beta}, \quad (14)$$

which is obtained by replacing the commutator with the time derivative of $\sum_j \mathbf{r}_j\cdot\nabla_j$.

simply in terms of the ψ_i, before an expansion in eigenfunctions is made: the relationship being[10]

$$E_{n+1}=\frac{1}{n+1}\sum_{m=0}^{n}\int\psi_m^*U\psi_{n-m}d\tau. \tag{19}$$

By summing both sides over all n, one obtains for the mean value of U in the perturbed system,

$$\bar{U}=\int\psi^*U\psi d\tau=\sum_{n=0}^{\infty}(n+1)E_{n+1}=\sum_{n=0}^{\infty}nE_n. \tag{20}$$

It is observed from the right hand side of (20) that the successive approximations for $\bar{U}$ do not converge as fast as those for the energy, $E=\sum E_n$. It nevertheless is assumed in the following that the perturbing term is small enough so that $\sum nE_n$ converges.

By making use of (6) and (7)

$$E=-\bar{T}-2\bar{U}, \tag{21}$$

or

$$E=(\bar{V}-\bar{U})/2. \tag{22}$$

A substitution of (17) for E, and (20) for $\bar{U}$ gives

$$\bar{T}=-\sum(1+2n)E_n, \tag{23}$$

and

$$\bar{V}=\sum(2+n)E_n. \tag{24}$$

Finally (20), (23), and (24) may be rearranged to give

$$E_m=\frac{\bar{U}}{m}-\frac{1}{m}\sum_{n\neq m}nE_n, \quad m\neq 0 \tag{25}$$

$$E_m=-\frac{\bar{T}}{1+2m}-\frac{1}{1+2m}\sum_{n\neq m}(1+2n)E_n, \tag{26}$$

or

$$E_m=\frac{\bar{V}}{2+m}-\frac{1}{2+m}\sum_{n\neq m}(2+n)E_n. \tag{27}$$

To demonstrate the usefulness of (25), (26), and (27), let it be assumed that from ordinary perturbation theory it is known through symmetry arguments that E_m is the perturbed energy term of interest for a particular problem. The second group of terms on the right in (25), (26), and (27) therefore may be dismissed from consideration. Thus if $\bar{U}'$, $\bar{T}'$, and $\bar{V}'$ are the parts of $\bar{U}$, $\bar{T}$, and $\bar{V}$ of the desired symmetry,

$$E_m\simeq(\bar{U}'/m)m\neq 0, \tag{28}$$

or

$$E_m\simeq-\bar{T}'/(1+2m), \tag{29}$$

or

$$E_m\simeq\bar{V}'/(2+m). \tag{30}$$

[10] It is interesting to note that if A_n is defined as $\Sigma_{m=0}^{n}\int\psi_m^*A\psi_{n-m}d\tau$, then by multiplying both sides of (18) by ψ^*, integrating, and equating equal coefficients of λ, one obtains $E_{n+1}=(T+V)_{n+1}+U_n$. But with the help of (19), $(T+V)_{n+1}=-nU_n/(n+1)$. Thus in higher orders, the contributions of the perturbing and unperturbed parts of the Hamiltonian tend more and more to cancel one another.

If any of the three quantities $\bar{U}$, $\bar{T}$, or $\bar{V}$ can be approximated in some manner, E_m is obtained by one of the above three equations. An approximate equality is written in case higher order perturbations, E_{m+i}, also contain the appropriate symmetry.

It readily may be verified that (28), (29), and (30) satisfy the relation $E_m\simeq\bar{T}'+\bar{V}'+\bar{U}'$ as they must; but the three equations show in addition that $\bar{T}$, $\bar{V}$, and $\bar{U}$ contribute in the ratio $-(1+2m):(2+m):m$, $m\neq 0$.

Equation (30) is generally the most useful since in order to approximate $\bar{V}$ (by an electrostatic calculation) it is necessary to approximate only the physically meaningful quantity $|\psi|^2$ (or in fact the probability density for two electrons). On the other hand, to calculate E_m by usual perturbation methods, all the unperturbed eigenfunctions germane to the problem must be known. In a many-body system, for high order perturbations, it is often simpler to approximate $|\psi|^2$ of the perturbed system than to approximate the complete spectrum of eigenfunctions of the unperturbed state.

AN ELEMENTARY EXAMPLE

As a check on the preceding results, a simple example, the Stark effect in a linear oscillator, may be worked out, since the exact results for this problem are known. The example also illustrates the use of a different Hamiltonian. If v is the potential energy, $kx^2/2$, of the oscillator, and $w=x\mathcal{E}$ is the perturbing term of the electric field $\mathcal{E}$, the energy is

$$E=\bar{T}+\bar{v}+\bar{w}. \tag{31}$$

As T, v, and w are respectively homogeneous of degree -2, 2, and 1, it follows that

$$0=-2\bar{T}+2\bar{v}+\bar{w}. \tag{32}$$

From (31) and (32)

$$\bar{v}=\tfrac{1}{2}E-\tfrac{3}{4}\bar{w}. \tag{33}$$

Making use of $E=\sum E_n$ and the perturbation expression (20) for $\bar{w}$, one obtains

$$\bar{v}=\sum(\tfrac{1}{2}-\tfrac{3}{4}n)E_n. \tag{34}$$

The problem is to find the term in the energy proportional to $\mathcal{E}^2$. From Rayleigh-Schrödinger perturbation theory this term obviously comes from the second order perturbation, which according to (34) is given by

$$E_2\simeq-\bar{v}(\mathcal{E}^2), \tag{35}$$

where $\bar{v}(\mathcal{E}^2)$ is the part of $\bar{v}$ depending on $\mathcal{E}^2$. Since the exact wave functions for the Hamiltonian $T+v+w$ are known, $\bar{v}(\mathcal{E}^2)$ can be calculated precisely.[11] The result for $\bar{v}$ is

$$\bar{v}=E_0/2+\mathcal{E}^2/2k, \tag{36}$$

[11] E. U. Condon and P. M. Morse, *Quantum Mechanics* (McGraw-Hill Book Company, New York, 1929), p. 122.

or $\bar{v}(\mathcal{E}^2)=\mathcal{E}^2/2k$. This quantity substituted into (35) checks the results of perturbation theory, in this case exactly.

It is not possible in this particular problem to find E_2 from $\bar{T}$, for from (31), (32), and (20),

$$\bar{T}=\sum\left(\frac{2-n}{4}\right)E_n, \tag{37}$$

and therefore $\bar{T}$ does not contain E_2, in agreement with the exact calculation showing $\bar{T}$ does not involve an $\mathcal{E}^2$ term.

In a paper to appear later, a more useful calculation is made of the fourth-order spin-orbit perturbation in a ferromagnetic solid. The potential energy is approximated by a multipole expansion.

APPENDIX. RAYLEIGH-SCHRÖDINGER THEORY

One proof of (20) can be obtained by solving for the various orders of ψ and E and substituting into (19). The method of the following much simpler proof was pointed out to the author by E. N. Adams and T. D. Holstein.

The mean value of the Hamiltonian $H_0+\lambda U$ is given by

$$E(\lambda)=\int\psi^*(\lambda)(H_0+\lambda U)\psi(\lambda)d\tau, \tag{38}$$

where $E(\lambda)$ and $\psi(\lambda)$ are given by (17). From the fact ψ is normalized and $H\psi=E\psi$, it is easily established that

$$\int\frac{\partial\psi^*}{\partial\lambda}H\psi d\tau+\int\psi^*H\frac{\partial\psi}{\partial\lambda}d\tau=0, \tag{39}$$

and thus

$$\frac{\partial E}{\partial\lambda}=\int\psi^*\frac{\partial H}{\partial\lambda}\psi d\tau=\bar{U}. \tag{40}$$

But $E=\sum\lambda^nE_n$, so (20) immediately follows from (40) when λ is set equal to unity.

LENNARD-JONES–BRILLOUIN–WIGNER THEORY

With Lennard-Jones–Brillouin–Wigner perturbation theory, Eq. (20) also holds. In this case, the direct proof is simpler since the energy and wave function easily can be written explicitly in the matrix elements U_{ij} and eigenvalues w_i of the unperturbed functions, u_i:

$$E=w_0+\lambda U_{00}+\lambda^2\sum_i{}'\frac{U_{0i}U_{i0}}{(E-w_i)}+\lambda^3\sum_i{}'\sum_j{}'\frac{U_{0i}U_{ij}U_{j0}}{(E-w_i)(E-w_j)}+\cdots, \tag{41}$$

$$\psi=u_0+\lambda\sum_i{}'\frac{U_{i0}u_i}{E-w_i}+\lambda^2\sum_i{}'\sum_j{}'\frac{U_{ij}U_{j0}u_i}{(E-w_i)(E-w_j)}+\cdots. \tag{42}$$

By a direct calculation using (42)

$$\bar{U}=U_{00}+2\lambda\sum_i{}'\frac{U_{0i}U_{i0}}{(E-w_i)}+3\lambda^2\sum_i{}'\sum_j{}'\frac{U_{0i}U_{ij}U_{j0}}{(E-w_i)(E-w_j)}+\cdots, \tag{43}$$

which may be compared with (41).