ACKNOWLEDGMENTS

It is a pleasure to thank Professor Geoffrey F. Chew for suggesting this investigation and for several stimulating discussions. We have also benefited from discussions with several colleagues, particularly Dr. John Charap and Dr. M. Froissart. One of us (B.M.U.) wishes to thank Dr. David L. Judd for his hospitality at the Lawrence Radiation Laboratory.

PHYSICAL REVIEW VOLUME 123, NUMBER 4 AUGUST 15, 1961

Hypervirial Theorems for Variational Wave Functions*

SAUL T. EPSTEIN
Department of Physics, University of Nebraska, Lincoln, Nebraska

AND

JOSEPH O. HIRSCHFELDER
Theoretical Chemistry Laboratory, University of Wisconsin, Madison, Wisconsin

(Received March 24, 1961)

It is shown that a sufficient condition for an optimal energy variational wave function ψ_0 to satisfy the hypervirial relation $(\psi_0,[H,W]\psi_0)=0$ is for the trial function ψ to admit variations of the form $\partial\psi/\partial a = (i/\hbar)W\psi$. Here H is the Hamiltonian, W is a Hermitian operator, and a is a variational parameter. Explicit forms of such trial functions are exhibited for several W's. The case in which W generates a point transformation of the coordinates is discussed in detail. Conditions are given for the existence of simultaneous hypervirial theorems.

I. INTRODUCTION

THE diagonal elements (in the energy representation) of the Heisenberg equations of motion[1] are called the hypervirial relations.[2] If χ is a (bound state) eigenfunction of a Hamiltonian H and if W (which henceforth is assumed to be Hermitian) is a time-independent operator, the hypervirial theorem for W states that

$$(\chi,[H,W]\chi)=0, \tag{1}$$

where $[H,W]\equiv HW-WH$ is the commutator of H and W. Physically, this is, of course, just the statement that, for a stationary state, the expectation value of W is independent of time.[3] For a particular choice of W, Eq. (1) yields the familiar virial theorem.[4] For other choices of W, the hypervirial relations lead to generalizations of the virial theorem.

It is well known that if a parameter is introduced into an *approximate*[5] wave function ψ in such a manner that all distances are scaled, and if the parameter is varied so as to obtain the optimum energy, then the corresponding optimal function ψ_0 satisfies the virial theorem.[6] Analogously, we show that, under certain conditions, it is possible to introduce a parameter into a trial function ψ so that the variationally determined approximate wave function ψ_0 satisfies the hypervirial theorem

$$(\psi_0,[H,W]\psi_0)=0. \tag{2}$$

The general plan of this paper is as follows: In Sec. II the conditions are derived in a formal manner. In Secs. III and IV these conditions are put into explicit form for certain special W's. In Sec. V the satisfaction of simultaneous hypervirial relations is discussed, and in Sec. VI possible applications and extensions of our results are considered. For simplicity of presentation, Cartesian coordinates are used throughout the main body

* The Wisconsin portion of this research is being reported under contract with the U. S. Atomic Energy Commission.

[1] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1955), 2nd ed., p. 140.

[2] J. O. Hirschfelder, J. Chem. Phys. **33**, 1762 (1960).

[3] More formally, Eq. (1) is also a partial expression of the fact that the eigenvalues of H are invariant to unitary transformation. Namely, if we subject H to the unitary transformation generated by W, then the first-order change in H is proportional to $i[H,W]$. Equation (1) then correctly tells us that the first-order energy shift vanishes.

[4] In reference 2 it is shown that if $W=\frac{1}{2}\sum_i(x_ip_i+p_ix_i)$, where the x_i are the Cartesian coordinates of the system and the p_i are the corresponding momentum operators, Eq. (1) is a statement of the quantum mechanical virial theorem, originally derived by M. Born, W. Heisenberg, and F. Jordan [Z. Physik **35**, 557 (1925)] and again by J. C. Slater [J. Chem. Phys. **1**, 687 (1933)].

[5] Throughout this paper, all approximate wave functions are assumed to satisfy the continuity-boundary conditions required of physically acceptable bound stationary state wave functions: (1) The function must be single-valued and analytic in all of its variables at every point in configuration space where the potential energy is analytic. (2) The function and its first derivatives must be absolutely and quadratically integrable over the whole of configuration space. (3) The function must vanish at infinity faster than any negative power of the Cartesian coordinates. See E. C. Kemble, *Fundamental Principles of Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1937), Sec. 32.

[6] E. A. Hylleraas, Z. Physik **54**, 347 (1929); V. Fock, *ibid.* **63**, 855 (1930); J. O. Hirschfelder and J. F. Kincaid, Phys. Rev. **32**, 658 (1937); and P. O. Löwdin, *Advances in Chemical Physics*, edited by I. Prigogine (Interscience Publishers, Inc., New York, 1959), Vol. II. p. 219.

of the text, with the generalizations to orthogonal curvilinear coordinates being given in Appendix I.

However, before embarking on this program, we should emphasize that the conditions which we find are all *sufficient* conditions, and that for some W's it may be possible to satisfy the hypervirial theorem Eq. (2) in other, simpler, ways. In particular, the hypervirial theorem may be satisfied simply because $W\psi_0$ is orthogonal to $H\psi_0$ for reasons of *symmetry*. For example, if H is invariant with respect to some symmetry group and ψ_0 is chosen to transform according to a nondegenerate irreducible representation of this group, the hypervirial relation is trivially satisfied unless W has a component which remains invariant under the group operations, since $(\psi_0,HW\psi_0)$ and $(\psi_0,WH\psi_0)$ each will vanish separately.[7] Thus, if H remains invariant under inversion, ψ_0 has a definite parity, and W is odd with respect to the inversion, then Eq. (2) is satisfied. Such symmetry arguments suggest that those W's will be most useful which transform in accordance with the irreducible representations of the symmetry group of H.

As another example, if H and W are invariant with respect to *time reversal*,[8] then the hypervirial theorem is satisfied by any trial wave function ψ_0 which is also invariant with respect to time reversal. In this case, the operator $\Omega=(i/\hbar)[H,W]$ (which corresponds to the time-derivative of W) is Hermitian and changes sign under time reversal. Thus, if ψ_0 is invariant with respect to time reversal, in accordance with a theorem due to Wigner, $(\psi_0,\Omega\psi_0)=0$ and the hypervirial relation is satisfied. Time reversal changes the direction of all motions and the sign of the spins. *Spin-free* operators and wave functions are invariant with respect to time reversal if they are real.[9] Thus, if we can neglect spins and if H is real, the hypervirial theorem for every real Hermitian W is satisfied by *any real* trial wave function ψ_0.

II. FORMAL VARIATIONAL FUNCTIONS

Let us suppose that the approximate energy of the system, $E=(\psi,H\psi)/(\psi,\psi)$, is a function of a parameter a. Here the Hamiltonian may be function of a or else the parameter may simply be a variational parameter embedded in ψ. Differentiating the definitional equation for the energy with respect to a and making use of the Hermitian character of H,

$$\frac{\partial E}{\partial a}(\psi,\psi)-\left(\psi,\frac{\partial H}{\partial a}\psi\right)=\left(\frac{\partial\psi}{\partial a},(H-E)\psi\right)+\left((H-E)\psi,\frac{\partial\psi}{\partial a}\right). \quad (3)$$

For exact wave functions, for which $(H-E)\psi=0$, the right-hand side of Eq. (3) is zero and Eq. (3) becomes a statement of the *generalized Hellmann-Feynman theorem*.[10,11] For approximate wave functions, let us suppose that

$$\partial\psi/\partial a=(i/\hbar)W\psi. \quad (4)$$

Then, making use of the Hermitian property of the H and W, Eq. (3) becomes

$$\frac{\partial E}{\partial a}(\psi,\psi)-\left(\psi,\frac{\partial H}{\partial a}\psi\right)=(i/\hbar)(\psi,[H,W]\psi). \quad (5)$$

Thus, a sufficient condition that an approximate wave function should satisfy the generalized Hellmann-Feynman theorem is that it satisfy Eq. (4) and the corresponding hypervirial theorem, Eq. (2). The remainder of the present paper has no further bearing on the Hellmann-Feynman theorem.

Instead, we seek the variational functions which lead to the hypervirial relations. Hence, in the remainder of this paper, a is considered to be a variational parameter and H is independent of a. Thus, if a_0 is the value of a for which $\partial E/\partial a=0$ and if ψ_0 is the approximate wave function for this optimal value of a, Eq. (5) becomes

$$(\psi_0,[H,W]\psi_0)=0. \quad (2)$$

This result may be summarized: *If, among other variational parameters, we introduce a parameter a such that Eq. (4) is satisfied, then the optimal ψ satisfies the hypervirial theorem for W. Or conversely, if ψ_0 satisfies the hypervirial theorem for W, then the energy is stationary for a variation as prescribed by Eq. (4).*

Approximate wave functions can be *selectively* improved by using the hypervirial theorem in the converse sense. A parameter (or set of parameters) in the approximate wave function can be adjusted so as to

[7] If ψ_0 belongs to a degenerate irreducible representation, then W need not have a component which remains invariant under the group operations. For example, in an atomic problem, if ψ_0 has the symmetry of a P function, then (in accordance with the Clebsch-Gordan theorem), $W\psi_0$ can have the same symmetry as ψ_0 if W transforms in the manner of the S, P, or D representations.

[8] E. P. Wigner, *Group Theory*, translated from German by J. J. Griffin (Academic Press, New York, 1959), see Chap. 26; also, E. P. Wigner, Nachr. Akad. Wiss Göttingen Math.-physik Kl. **31**, 546 (1932).

[9] The effect of time reversal on operators and wave functions involving spin is somewhat more complicated. For example, if ψ_0 is an electronic wave function having the form

$$\psi_0=\sum_{s_1=\pm\frac{1}{2}}\cdots\sum_{s_n=\pm\frac{1}{2}}\Phi_{s_1\cdots s_n}(x_1,y_1,z_1,s_1;\cdots;x_n,y_n,z_n,s_n),$$

and θ is the time-reversal operator, then

$$\theta\psi_0=\sum_{s_1=\pm\frac{1}{2}}\cdots\sum_{s_n=\pm\frac{1}{2}}i^{-2(s_1+\cdots+s_n)}\times\Phi_{s_1\cdots s_n}{}^*(x_1,y_1,z_1,-s_1;\cdots;x_n,y_n,z_n,-s_n).$$

Thus, ψ_0 is invariant with respect to time reversal if $\psi_0=\theta\psi_0$.

[10] H. Hellmann, *Einfuhrung in die Quantenchemie* (Franz Deuticke, Leipzig, 1937), p. 285; R. P. Feynman, Phys. Rev. **56**, 340 (1934); and T. Berlin, J. Chem. Phys. **19**, 208 (1951). For a related discussion, see A. A. Frost and P. G. Lykos, J. Chem. Phys. **25**, 1299 (1956).

[11] A. C. Hurley, Proc. Roy. Soc. (London) **A226**, 170, 179, and 193 (1954), points out that the generalized Hellmann-Feynman theorem is also satisfied by approximate wave functions which do not depend upon the parameter a, so that $\partial\psi/\partial a=0$. He used such "floating" wave functions to calculate molecular energies and obtained results which were not encouraging.

enable the resulting ψ_0 to satisfy the hypervirial theorem for a particular W (or simultaneously, for a particular set of W's). The resulting energy of the system is stationary with respect to variations of the approximate wave function of the form $\psi_0+\epsilon iW\psi_0$ so that the optimum value of the parameter ϵ is zero. Thus, the choice of the W's determines in what sense the approximate wave function is improved.

It is easy to generalize our results so as to apply to cases, such as the Hartree-Fock calculations, where whole functions are determined on a variational basis. If $\delta\psi=\epsilon iW\psi$ is included in the set of possible variations, then the energetic optimum ψ_0 satisfies the hypervirial theorem for W.

A *formal* solution to Eq. (4) is, of course,[12]

$$\psi=\exp(iaW/\hbar)\Phi, \tag{6}$$

where Φ is a function independent of a and where, in order that this formal solution should be acceptable, ψ must satisfy the continuity-boundary conditions as given in footnote 5. In the next two sections, we determine the *explicit* functional form of ψ for certain special W's.

Equation (6) leads to a simple criterion that the energy corresponding to a_0 should not only be stationary with respect to the type of variation indicated by Eq. (4), but should actually be a *minimum*. Writing ψ in the form $\psi=\exp[i(a-a_0)W/\hbar]\psi_0$ and making use of the unitary property of the exponential operator,

$$E=(\psi_0,\exp[-i(a-a_0)W/\hbar]H \times\exp[i(a-a_0)W/\hbar]\psi_0)/(\psi_0,\psi_0). \tag{7}$$

Expanding the exponential operator in Eq. (7) in powers of $(a-a_0)$ and making use of Eq. (2), we obtain

$$E=E_0-\frac{(a-a_0)^2}{2\hbar^2}(\psi_0,[W,[W,H]]\psi_0)/(\psi_0,\psi_0)+\cdots. \tag{8}$$

Thus the energy corresponding to ψ_0 is *stable* with respect to the variation as indicated by Eq. (4) if

$$(\psi_0,[W,[W,H]]\psi_0)<0. \tag{9}$$

III. THE SPECIAL CASE: $W=\frac{1}{2}(pf+fp)$

The most important class of hypervirial relations corresponds to setting $W=\frac{1}{2}(pf+fp)$. Here $p=(\hbar/i)\partial/\partial x$ is the momentum operator corresponding to a Cartesian coordinate x and $f(x)$ is an arbitrary real function of x and the other coordinates of the system, which need not be indicated explicitly. The $f(x)$ should not involve any momenta. For this case,[13] W is most conveniently written in the form $W=(\hbar/i)f^{\frac{1}{2}}(\partial/\partial x)(f^{\frac{1}{2}})$ so that Eq. (4) becomes

$$\partial\psi/\partial a=f^{\frac{1}{2}}\partial(f^{\frac{1}{2}}\psi)/\partial x. \tag{10}$$

Multiplying Eq. (10) by $f^{\frac{1}{2}}$ and changing variables from x to

$$S(x)=\int^x dx'/f(x'), \tag{11}$$

we obtain

$$(\partial/\partial a)(f^{\frac{1}{2}}\psi)=(\partial/\partial S)(f^{\frac{1}{2}}\psi). \tag{12}$$

The general solution to Eq. (12) is

$$\psi(x,a)=f^{-\frac{1}{2}}(x)\Theta(a+S(x)). \tag{13}$$

Here Θ is an arbitrary function of $a+S(x)$.

Let us define the new variable $\sigma(x,a)$ *by the requirement that as*

$$a\rightarrow 0,\quad \sigma(x,a)\rightarrow x, \tag{14}$$

and by either of the equivalent relations

$$\partial\sigma/\partial a=f(\sigma), \tag{15}$$

or

$$S(\sigma)=a+S(x). \tag{16}$$

Equation (13) can then be written in terms of σ,

$$\psi(x,a)=f^{\frac{1}{2}}(\sigma)f^{-\frac{1}{2}}(x)\Phi(\sigma). \tag{17}$$

Moreover, the Jacobian of the transformation from σ to x is

$$\partial\sigma/\partial x=f(\sigma)/f(x). \tag{18}$$

Thus, ψ is derived from Φ by the point transformation[14,15] $x\rightarrow\sigma(x,a)$. This is indeed a generalization of the coordinate scaling in the case of the usual virial theorem for which (in one dimension) $f(x)=x$, $S(x)=\log_e|x|$, and $\sigma(x,a)=xe^a$.

The idea of using a point transformation to improve approximate wave functions is not new. Both Löwdin[16] and Hall[17] have used point transformations to improve atomic wave functions. However, their transformation functions were not derivable from the hypervirial relations and, as a result, they experienced considerable difficulty in optimizing the energy.

Only a very limited class of functions can satisfy our definition requirements for the transformation function $\sigma(x,a)$. Indeed, it would be difficult to select, at random, a function $\sigma(x,a)$ whose partial derivative with respect to a is a function only of σ and does not otherwise involve either x or a. Although any function $f(x)$

[12] If, in place of Eq. (4), we had assumed that $\partial\psi/\partial a=W\psi$, then, instead of the hypervirial relation, we would have obtained the following equation which is satisfied by the optimal function,

$$(\psi_0,WH\psi_0)+(\psi_0,HW\psi_0)=2E(\psi_0,W\psi_0).$$

In this case, the formal solution is $\psi=\exp(aW)\Phi$.

[13] The generalization to orthogonal curvilinear coordinates is given in Appendix I.

[14] An alternate derivation of Eq. (16) is given in Appendix II which exhibits this point transformation from the outset.

[15] B. S. DeWitt, Phys. Rev. **85**, 653 (1952), discusses the relationship between point transformations in quantum and classical mechanics.

[16] P. O. Löwdin, *Proceedings of the Nikko Symposium on Molecular Physics* (Maruzen Ltd., Tokyo, 1954), p. 116.

[17] G. G. Hall, Proc. Phys. Soc. (London) **75**, 575 (1960). Hall used the radial coordinate transformation function,

$$\sigma(r,a)=[-1+(1+4ar)^{\frac{1}{2}}]/2a.$$

will generate a transformation function (for at least some values of a and some values of x) most of these σ's have pathological properties[18] (such as discontinuities and values of x for which σ is not defined) which make them unusable for our purposes.

For the most satisfactory transformations, $\sigma(x,a)$ and $\partial\sigma/\partial x$ are continuous functions of x and, also, $\sigma(x,a)$ should be a monotonically increasing function of x and have the same range as x, that is, from $-\infty$ to ∞. In this case, the transformation $x\rightarrow\sigma(x,a)$ is a simple 1-1 mapping and it is clear that if $\Phi(x)$ is a suitable trial wave function, then $\psi(x,a)$ is equally suitable. One example of this class of transformations is

$$\sigma(x,a)=x_0+\sinh^{-1}[a+\sinh(x-x_0)], \tag{19}$$

which results from taking

$$S(x)=\sinh(x-x_0) \quad \text{or} \quad f(x)=1/\cosh(x-x_0). \tag{20}$$

Another equally satisfactory transformation function is

$$[\sigma(x,a)]^2=\mathrm{Ei}^{(-1)}\{a+\mathrm{Ei}(x^2)\}. \tag{21}$$

Here $\mathrm{Ei}^{(-1)}(z)$ is the inverse of the exponential integral $\mathrm{Ei}(z)$. This transformation results from

$$S(x)=\mathrm{Ei}(x^2) \quad \text{or} \quad f(x)=\tfrac{1}{2}x\exp(-x^2). \tag{22}$$

The sufficient conditions for the function $f(x)$ to generate satisfactory transformations are:

(1) $f(x)$ is continuous and finite for finite values of x.

(2) At any point x_0 where $f(x_0)=0$, df/dx is zero or finite.

(3) $f(x)/x$ is finite or zero as $x\rightarrow\pm\infty$.

With these conditions it is easy to discover a wide variety of satisfactory transformations.

Sometimes it is desirable to confine the distortion of the wave function to a specified region of space. For example, suppose that it is desired to leave the wave function undisturbed for values of x greater than x_1. This may be accomplished by selecting a function $f(x)$ such that $f(x)$ and df/dx approach zero as $x\rightarrow x_1$. Then at x_1, $\sigma=x_1$ and $\partial\sigma/\partial x=1$. Thus, a physically satisfactory transformation would correspond to using such a function $f(x)$ for $x\leq x_1$ and $f(x)=0$, corresponding to $\sigma=x$, for $x\geq x_1$.

It is also possible to use the 1-1 transformations which result from the satisfaction of conditions (1) and (2) but not (3). In this case, the $-\infty$ to ∞ range of x is mapped into the $-\alpha(a)$ to $\beta(a)$ range of $\sigma(x,a)$. However, the $\Phi(\sigma)$ is required to be identically equal to zero outside of the range $-\alpha'<\sigma<\beta'$, where $\alpha'<\alpha(a)$ and $\beta'<\beta(a)$ for the desired range of values of a. In order that $\Phi(\sigma)$ should satisfy the requirements for an acceptable trial wave function, both $\Phi(\sigma)$ and its first derivative must approach continuously the value of zero in the vicinity of both $\sigma=\alpha'$ and $\sigma=\beta'$. The method of convolutions[19] can be used to construct approximate wave functions which satisfy these boundary conditions. A typical example of this type of transformation is

$$\sigma=x[1+(x/b)^{n-1}]^{-1/(n-1)}. \tag{23}$$

Here, n is an odd integer and $b^{n-1}=-1/a$. This transformation is generated by $f(x)=(n-1)x^n$, for which $S(x)=-1/x^{(n-1)}$. In this case, $\alpha(a)=\beta(a)=b$.

The analysis in this section is not limited to one-dimensional problems. In a many-dimensional problem, the x would represent the jth Cartesian coordinate and p would be the corresponding momentum. The $f(x)$ might represent a function of all of the coordinates and $S(x)$ becomes the integral of $1/f(x)$ over the coordinate x_j (holding all of the other coordinates constant). Thus, σ can involve all of the coordinates. For example, in the transformation given by Eq. (19), the x_0 might be a function of all of the coordinates except x_j. The function $f(\sigma)$ is the same as $f(x)$ with the exception that wherever x_j occurs it is replaced by σ. The same interpretation is to be placed on $\Phi(\sigma)$.

The special case of spherical coordinates[13] with $W=(f^{\frac{1}{2}}/r)(\hbar/i)(\partial/\partial r)(rf^{\frac{1}{2}})$ merits special attention. Here f is a function of r, θ, and φ. Defining σ again by Eqs. (11), (14), and (16), with the exception that now r everywhere replaces the variable x, the approximate wave function can be written in the form

$$\psi(r,\theta,\varphi;a)=r^{-1}f^{\frac{1}{2}}(\sigma,\theta,\varphi)f^{-\frac{1}{2}}(r,\theta,\varphi)\Phi(\sigma,\theta,\varphi). \tag{24}$$

For the most satisfactory transformations, $\sigma(r,\theta,\varphi;a)$ and $\partial\sigma/\partial r$ should be continuous functions, and also $\sigma(r,\theta,\varphi)$ should be a monotonically increasing function of r and have the same range as r, i.e., from 0 to ∞. With this simple 1-1 mapping it is clear that if $\Phi(r,\theta,\varphi)/r$ is a suitable approximate wave function, then $\psi(r,\theta,\varphi;a)$ is equally suitable. This type of transformation is generated by functions $f(r,\theta,\varphi)$ which satisfy the conditions: (1) $f(r)$ is continuous and finite for finite values of r; (2) At any point r_0 where $f(r_0)=0$, $\partial f/\partial r$ is finite or zero; (3) $f(r)/r$ is finite or zero both in the limit as $r\rightarrow 0$ and as $r\rightarrow\infty$. One example of this type of transformation is

$$2\sigma=(a+r-\alpha r^{-1}) +[a^2+2a(r-\alpha r^{-1})+(r+\alpha r^{-1})^2]^{\frac{1}{2}}. \tag{25}$$

Here α is positive and either a constant or else a function of θ and φ. This transformation is generated by $f(r,\theta,\varphi)=r^2(\alpha+r^2)^{-1}$ or $S(r,\theta,\varphi)=r-\alpha r^{-1}$. Another example is

$$\sigma=\mathrm{Ei}^{(-1)}\{a+\mathrm{Ei}(r)\}. \tag{26}$$

Here again $\mathrm{Ei}^{(-1)}(z)$ is the inverse of the exponential integral $\mathrm{Ei}(z)$. This transformation is generated by $f(r)=re^{-r}$ or $S(r)=\mathrm{Ei}(r)$.

[18] Some topological observations regarding the transformation functions and their properties are given in Appendix III.

[19] J. O. Hirschfelder and G. V. Nazaroff, J. Chem. Phys. 34, 1666 (1961).

IV. OTHER SOLUBLE W's

(A) The analysis of the last section is readily extended to the case

$$W=\tfrac{1}{2}\sum_{k=1}^{n}(p_k f_k+f_k p_k), \tag{27}$$

where now f_k depends only on x_k (and, if the system is more than n dimensional, on the coordinates not represented in the summation and which are not indicated explicitly). One easily ascertains that the solution in this case is[19a]

$$\psi(x_1,\cdots,x_n;a)$$
$$=\left[\prod_{k=1}^{n} f_k^{\frac{1}{2}}(\sigma_k)f_k^{-\frac{1}{2}}(x_k)\right]\Phi(\sigma_1,\cdots,\sigma_n), \tag{28}$$

where σ_k is defined by the requirement that as

$$a\to 0,\ \sigma_k\to x_k, \tag{29}$$

and

$$S_k(\sigma_k)=a+S_k(x_k), \tag{30}$$

where

$$S_k(x_k)=\int^{x_k} dx'/f_k(x'). \tag{31}$$

The usual virial theorem, for which $f_k=x_k$, $S_k(x_k)=\log_e|x_k|$, and $\sigma_k=x_k e^a$, is an example.

(B) The explicit functional form of the solution of Eq. (4) is known to be for W a quadratic function[20] of the p_k and x_k. In this case, ψ and Φ are related by an *integral transform*. However, if the W does not contain any terms linear in the p_k so that W is real and invariant with respect to time reversal, as noted in Sec. I, the hypervirial theorem is satisfied by any *real* trial function.

(C) The case where $W=f(x_1,\cdots,x_n)$ yields a simple gauge transformation

$$\psi(x_1,\cdots,x_n;a)=\exp(iaf/\hbar)\Phi(x_1,\cdots,x_n). \tag{32}$$

However, as noted in Sec. I, the hypervirial theorem for this W is trivially satisfied by *any* real trial function.

(D) Representing the wave function in momentum, rather than coordinate, space, it follows that $W=\frac{1}{2}\sum_k(x_k g_k+g_k x_k)$, where the g_k is a real function of p_k, generates a point transformation in *momentum space*. Also, $W=g(p_1,\cdots,p_n)$ generates a gauge transformation in *momentum space*.

(E) In our discussion we have been concerned with the determination of general solutions to Eq. (4) in closed form. However, it should be noted that there are many W's for which, although we cannot find such a closed form, we can find some, or all, of their eigenfunctions: $W\varphi_{W'}=W'\varphi_{W'}$. If this is possible, we can write down any number of *particular* solutions of Eq. (4), which have the form

$$\psi=\sum_{W'} A_{W'}\varphi_{W'}\exp(iW'a/\hbar), \tag{33}$$

where the constants $A_{W'}$ are independent of a and otherwise arbitrary.

V. SIMULTANEOUS HYPERVIRIAL THEOREMS

We wish to satisfy the two hypervirial relations

$$(\psi_0,[H,W_1]\psi_0)=0 \quad\text{and}\quad (\psi_0,[H,W_2]\psi_0)=0 \tag{34}$$

simultaneously (the generalization to more than two will be obvious). Then we may introduce, possibly among others, two variational parameters a_1 and a_2 such that

$$\partial\psi/\partial a_1=(i/\hbar)W_1\psi \quad\text{and}\quad \partial\psi/\partial a_2=(i/\hbar)W_2\psi. \tag{35}$$

Hence, we must solve the pair of equations (35) simultaneously. If W_1 and W_2 commute, the solution is simple, namely

$$\psi=\exp(ia_1W_1/\hbar)\exp(ia_2W_2/\hbar)\Phi. \tag{36}$$

In particular, if $W_i=\frac{1}{2}[p_i f_i(x_i)+f_i(x_i)p_i]$, then

$$\psi(x_1,x_2;a_1,a_2)$$
$$=f_1^{\frac{1}{2}}(\sigma_1)f_2^{\frac{1}{2}}(\sigma_2)f_1^{-\frac{1}{2}}(x_1)f_2^{-\frac{1}{2}}(x_2)\Phi(\sigma_1,\sigma_2), \tag{37}$$

where the

$$S_i(\sigma_i)=a_i+S(x_i), \tag{38}$$

and the S_i are related to the f_i as before.

We now turn to the case where the W_1 and W_2 *do not commute*. In this case we do not know how to *a priori* select both W_1 and W_2 and obtain a formal trial function such that the function which results from optimizing a_1 and a_2 satisfies Eq. (34). Indeed, if W_1 and W_2 were independent of the variational parameters, then one readily sees that the two relations of Eq. (35) are not integrable, since $\partial^2\psi/\partial a_1\partial a_2$ would not be equal to $\partial^2\psi/\partial a_2\partial a_1$. However, if we let W_2 be a function of a_1, we can obtain hypervirial theorems for W_1 and $W_2(a_{10})$, where a_{10} is the optimum value of a_1. The $W_2(a_1)$ is given by

$$W_2(a_1)=\exp(ia_1W_1/\hbar)\bar{W}_2\exp(-ia_1W_1/\hbar). \tag{39}$$

Here W_1 and $\bar{W}_2$ *are* independent of the variation parameters and hence *can* be selected *a priori*.

Clearly, from the discussion of Sec. II, Eq. (34) is satisfied with W_2 replaced by $W_2(a_{10})$ if Eq. (35) is satisfied with W_2 replaced by $W_2(a_{10})$. But from Eq. (39) we readily see that the relations of Eq. (35) are now integrable and have as their solution

$$\psi(x;a_1,a_2)=\exp(ia_1W_1/\hbar)\exp(ia_2\bar{W}_2/\hbar)\Phi(x)$$
$$=\exp(ia_2W_2/\hbar)\exp(ia_1W_1/\hbar)\Phi(x). \tag{40}$$

To summarize: *If we use a trial function of the form given by Eq. (40), then the optimal function satisfies the hyper-*

[19a] *Note added in proof.* The generalization corresponding to f_k an arbitrary function of the coordinates is considered in a forthcoming paper by J. O. Hirschfelder and C. A. Coulson in J. Chem. Phys.

[20] M. Kolsrud, Phys. Rev. **104**, 1186 (1956).

virial theorems for W_1 and W_2 (a_{10}). Note also that

$$W_2(a_1)=W_2(\bar{x},\bar{p}), \tag{41}$$

where

$$\begin{aligned}\bar{x}&=\exp(ia_1W_1/\hbar)x\,\exp(-ia_1W_1/\hbar),\\ \bar{p}&=\exp(ia_1W_1/\hbar)p\,\exp(-ia_1W_1/\hbar).\end{aligned} \tag{42}$$

As an example, let $W_1=\frac{1}{2}[pf_1(x)+f_1(x)p]$ and $W_2=\frac{1}{2}[pf_2(x)+f_2(x)p]$. Then the first hypervirial leads to the point transformation $x\rightarrow\sigma_1(x)$ and the second leads to the transformation $\sigma_1(x)\rightarrow\sigma_2(\sigma_1(x))$. Thus,

$$\begin{aligned}\psi(x;a_1,a_2)&=\exp(ia_1W_1/\hbar)\,\exp(ia_2W_2/\hbar)\Phi(x)\\ &=\frac{f_1^{\frac{1}{2}}(\sigma_1(x))f_2^{\frac{1}{2}}(\sigma_2(\sigma_1(x)))}{f_1^{\frac{1}{2}}(x)f_2^{\frac{1}{2}}(\sigma_1(x))}\Phi(\sigma_2(\sigma_1(x))).\end{aligned} \tag{43}$$

VI. DISCUSSION

It is clear that one of the principal applications of the hypervirial relations will be to *selectively* improve atomic and molecular wave functions in regions of configuration and momentum space which are important for determining the expectation values of specific physical problems. Thus, many of the variational parameters in the trial functions will be determined by the requirement that an appropriate set of hypervirial relations be satisfied.[21] Hypervirial relations which are designed to improve the electron correlation would be most useful.

As mentioned in connection with Eq. (5), the generalized Hellmann-Feynman theorem[10] is applicable to approximate wave functions which satisfy certain hypervirial relations. However, we do not know how to select the operator W so as to insure that the approximate wave function will give the correct dependency of the energy on the Hamiltonian parameter. Good approximate wave functions which satisfy the Hellmann-Feynman theorem would be very useful in the calculation of intermolecular forces, chemical binding energies, etc.

The hypervirial relations may also be useful in variational approaches to meson theory[22] and we (S.T.E.) hope to initiate some research along these lines.

Further, it may be of interest to indicate some possible generalizations of our results. One obvious generalization would be the introduction of spin and relativity (the Pauli and Dirac equations). The formal results of Sec. II would still apply and the essential problem would be the determination of the soluble W's. Quite apart from spin and relativity, the class of soluble W's may be extended by noting that if $\bar{x}$ and $\bar{p}$ are related to x and p by a unitary transformation, then

$$W=\tfrac{1}{2}\sum[\bar{p}_kf_k(\bar{x}_k)+f_k(\bar{x}_k)\bar{p}_k] \tag{44}$$

generates a point transformation in $\bar{x}$ space. In particular, as we have already remarked, $W=\frac{1}{2}\sum[x_kg_k(p_k)+g_k(p_k)x_k]$ generates a point transformation in momentum space.

It would be useful to establish the conditions for the validity of "off-diagonal" hypervirial theorems.[23] That is, if χ_1 and χ_2 are exact wave functions, then

$$(\chi_1,[H,W]\chi_2)=(E_1-E_2)(\chi_1,W\chi_2). \tag{45}$$

Under what conditions would Eq. (45) apply if χ_1 and χ_2 were replaced by approximate wave functions and the E_1 and E_2 were replaced by the corresponding variationally determined energies?

Finally, there are a wide variety of *generalized hypervirial theorems* of the form

$$(\psi_0,F[G,W]J\psi_0)=0, \tag{46}$$

where F, G, and J are functions (or possibly functionals) of H. Furthermore, G and W are Hermitian. Obviously, any true (bound state) eigenfunction of the Hamiltonian satisfies Eq. (46) for arbitrary F, G, J, and W. Thus, some of the parameters in an approximate wave function might be adjusted by requiring the satisfaction of such generalized hypervirial relations.

APPENDIX I. GENERALIZATIONS TO ORTHOGONAL CURVILINEAR COORDINATES

Let us consider a set of generalized orthogonal coordinates q_j and their corresponding metric scale factors g_j. It is convenient to let $g=g_1g_2\cdots$. The most important class of hypervirials corresponds to $W=\frac{1}{2}(p_kf+fp_k)$, where $p_k=(\hbar/i)(\partial/\partial q_k)$ is the operator for the kth generalized momentum and f may be a function of all of the coordinates. In the following development, only the q_k is indicated explicitly. In order that W be Hermitian, it should be expressed in the form

$$W=(f/g)^{\frac{1}{2}}(\hbar/i)(\partial/\partial q_k)(f^{\frac{1}{2}}g^{\frac{1}{2}}\quad). \tag{A.1}$$

The derivation of the explicit solution to Eq. (4) is very similar to that which was given for the Cartesian coordinates. Thus,

$$\psi(q_k,a)=g^{-\frac{1}{2}}f^{\frac{1}{2}}(\sigma)f^{-\frac{1}{2}}(q_k)\Phi(\sigma). \tag{A.2}$$

The definition of the transformation function σ is the same as in the Cartesian coordinate case, with the exception that q_k replaces x. The most satisfactory transformation functions are continuous monotonically increasing functions of q_k and have the same range as q_k. The transformation $q_k\rightarrow\sigma$, in this case, is a simple 1-1 mapping and if $g^{-\frac{1}{2}}\Phi(q_k)$ is a suitable trial wave function, $\psi(q_k,a)$ is equally suitable.

[21] If the trial function has the form $\psi=c_i\Sigma\varphi_i$, where the φ_i are known functions, then the satisfaction of a set of hypervirial relations leads to a set of simultaneous quadratic forms which must be solved for the coefficients c_i, $\Sigma_{i,j}c_ic_jJ_{ij}(k)=0$ with $k=1, 2, \cdots, n$. If n is three or more, it is difficult to determine the coefficients. An improved procedure for numerical calculations is needed.

[22] S. Fubini, Nuovo cimento **3**, 1725 (1956) and Suppl. Nuovo cimento **14**, 283 (1959) used hypervirial theorems to derive various exact results in fixed-source meson theories.

[23] S. Ehrenson and P. E. Phillipson, Bull. Am. Phys. Soc. **5**, 155 (1960).

The hypervirial relations, corresponding to the W of Eq. (A.1), can be expressed in a very significant form[24] if ψ is real,

$$\left(f\psi^2,\frac{\partial}{\partial q_k}(H\psi/\psi)\right)=0. \tag{A.3}$$

The derivation of Eq. (A.3) is easy. Since H and ψ are real, W is pure imaginary, and H and W are Hermitian,

$$(\psi,HW\psi)=(H\psi,W\psi)=-(W\psi,H\psi)=-(\psi,WH\psi). \tag{A.4}$$

Thus, for this case, the hypervirial relation becomes

$$(\psi,[H,W]\psi)=2(H\psi,W\psi)=0. \tag{A.5}$$

But, from Eq. (A.1) it follows that

$$W\psi=(2g\psi)^{-1}\frac{\partial}{\partial q_k}(fg\psi^2). \tag{A.6}$$

Substituting Eq. (A.6) into Eq. (A.5),

$$\left(g^{-1}\psi^{-1}H\psi,\frac{\partial}{\partial q_k}(fg\psi^2)\right)=0. \tag{A.7}$$

Equation (A.3) results from Eq. (A.7) after integration by parts. If ψ is the true wave function satisfying the Schrödinger equation, $H\psi=E\psi$, Eq. (A.3) would be obviously true. Indeed, the hypervirial relation may be obtained by inverting this derivation.

The Hamiltonian for a system of N particles each having the same mass m is

$$H=-\frac{\hbar^2}{2m}\sum_{i=1}^{3N}g^{-1}\frac{\partial}{\partial q_j}\left[gg_j^{-2}\frac{\partial}{\partial q_j}\right]+V. \tag{A.8}$$

Let us define the hypervirial operator,

$$J(W)=(i/\hbar)[W,H].$$

Then, if W is a function of the coordinates, $W=f$,

$$J(f)=\frac{i\hbar}{2m}\sum_{i=1}^{3N}\left\{g^{-1}\frac{\partial}{\partial q_j}\left[gg_j^{-2}\frac{\partial f}{\partial q_j}\right]+2g_j^{-2}\frac{\partial f}{\partial q_j}\frac{\partial}{\partial q_j}\right\}. \tag{A.9}$$

If, on the other hand, W is given by Eq. (A.1) and we indicate this W by the abbreviation fp_k, then

$$J(fp_k)=f\frac{\partial V}{\partial q_k}-\frac{\hbar^2}{2m}\sum_{i=1}^{3N}B_{kj}+\frac{\hbar}{i}J(f)\frac{\partial}{\partial q_k}+\frac{\hbar}{2i}J(t). \tag{A.10}$$

Here,

$$B_{kj}=\left[\frac{\partial}{\partial q_k}\left\{g^{-1}\frac{\partial}{\partial q_k}(gg_j^{-2})\right\}\right]\frac{\partial}{\partial q_k}+\left(\frac{\partial g_j^{-2}}{\partial q_k}\right)\frac{\partial^2}{\partial q_j^2}, \tag{A.11}$$

and the $J(t)$ is the same as the $J(f)$, except that f is now replaced by $t=g^{-1}(\partial/\partial q_k)(fg)$.

[24] This is a generalization of a derivation given by M. S. Wertheim in reference 2.

APPENDIX II. ALTERNATIVE DERIVATION OF EQ. (16)

We will use a one-dimensional notation throughout, indicating the generalization to more than one dimension at appropriate points. We use the Dirac notation $|B=B'\rangle$ to denote an eigenket of the operator B, with eigenvalue B'. Since $B|B=B'\rangle=B'|B=B'\rangle$, we have

$$\begin{aligned}&\Phi(x')=\langle x=x'|\Phi\rangle,\\ &\psi(x',a)=\langle x=x'|\exp(iaW/\hbar)|\Phi\rangle=\langle X_U=x'|\Phi\rangle,\end{aligned} \tag{B.1}$$

where

$$X_U=U^{-1}xU,\quad U=\exp(iaW/\hbar). \tag{B.2}$$

So far this has been quite general. Now let us assume that U generates a point transformation, i.e., that X_U depends only on x but not on p, and indeed that this is a transformation which provides a simple 1-1 mapping with X_U being a continuous monotonically increasing function of x. This implies that when X_U has a sharp value x', the x also has a sharp value; call it σ'. Thus, we infer that

$$\langle X_U=x'|=N\langle x=\sigma'|, \tag{B.3}$$

where N is the normalization factor. To determine N, we have

$$\begin{aligned}\langle X_U=x'|X_U=x''\rangle&=\delta(x'-x'')=\delta(\sigma'-\sigma'')d\sigma'/dx'\\ &=\langle x=\sigma'|x=\sigma''\rangle d\sigma'/dx'.\end{aligned} \tag{B.4}$$

From Eq. (B.4) it is clear that $N=(d\sigma'/dx')^{\frac{1}{2}}$ and hence

$$\psi(x',a)=(d\sigma'/dx')\langle x=\sigma'|\Phi\rangle=(d\sigma'/dx')\Phi(\sigma'). \tag{B.5}$$

For a point transformation in more than one variable, the $d\sigma'/dx'$ would be replaced by the Jacobian of the transformation.

The only remaining problem is to determine $\sigma'(x')$. With $W=\frac{1}{2}(pf+fp)$, we find

$$dX_U/dx=-f(X_U). \tag{B.6}$$

Thus,

$$-a=\int_x^{X_U}dz/f(z)=S(X_U)-S(x). \tag{B.7}$$

As a result, $S(\sigma')=S(x')+a$, and we have rederived Eq. (16). For a point transformation in more than one variable, Eq. (B.6) would be replaced by a set of first-order partial differential equations.

APPENDIX III. TOPOLOGICAL STRUCTURE OF THE TRANSFORMATION

Let us make some observations regarding the topological structure of the $x\rightarrow\sigma(x)$ transformations which can result from the defining equations for σ. Depending upon the generating function $f(x)$, the σ can have a rather wide range of behavior. For present purposes, let us focus our attention on $S(x)$.

1. *If $S(x)$ is a continuous monotonically increasing function of x*, then, for those values of x for which $\sigma(x)$ is defined, $\sigma(x)$ is also a continuous monotonically in-

creasing function of x. For this case, $f(x)$ cannot be negative. At points where $f(x)=0$, $S(x)$ has an inflection point with $dS/dx=\infty$. At points where $f(x)=\infty$, $S(x)$ has an inflection point with $dS/dx=0$. Furthermore:

(a) If $S(x)$ is bounded, and if $|a|>S(\infty)-S(-\infty)$, $\sigma(x)$ is not defined for any value of x. However, if a is positive and less than $S(\infty)-S(-\infty)$, there is a maximum value of x above which no value of σ is defined and there is a finite lower limit for σ corresponding to $x=-\infty$. For a negative and $-a<S(\infty)-S(-\infty)$, the σ is not defined for x less than a lower limit, and $\sigma(\infty)$ is finite.

(b) If $S(\infty)=\infty$ and $S(-\infty)$ is bounded, for positive values of a, there exists a value of σ corresponding to each value of x, but $\sigma(-\infty)$ is finite; for negative values of a, there is a smallest value of x for which σ is defined.

(c) If $S(\infty)=\infty$ and $S(-\infty)=-\infty$, for all values of a, σ is defined for all values of x. Furthermore, $\sigma(\infty)=\infty$ and $\sigma(-\infty)=-\infty$.

2. *If $S(x)$ has a finite maximum at x_0,* then since $dS/dx=1/f(x)$, the function f must have an infinite discontinuity at x_0 such that $x^-\rightarrow x_0$, $f(x)\rightarrow\infty$ and $x^+\rightarrow x_0$, $f(x)\rightarrow-\infty$. The situation is quite dependent upon the sign of a. If:

(a) a is positive, there may be a gap in the values of x for which σ is defined. Also, there may be a range of x for which σ is double-valued. *The multivalued nature of σ can generally be resolved by introducing the requirement that $\sigma(x)\rightarrow x$ as $a\rightarrow 0$.* In this case, the branch with $\sigma(x)<x_0$ would be associated with $x<x_0$ and the branch with $\sigma(x)>x_0$ would be associated with $x>x_0$.

(b) a is negative, σ may be double-valued for a range of x in the vicinity of x_0. Introducing the requirement that $\sigma(x)\rightarrow x$ as $a\rightarrow 0$, leads to the resolution: For $x<x_0$ it is the lower branch of σ that we should use, whereas for $x>x_0$ it is the higher branch. Thus, at x_0 there is a discontinuity in the value of σ.

3. *If $S(x)$ has a singular point at x_0,* then[25] $f(x_0)=0$. In this case, there may be no connection between the integral defining $S(x)$ for $x<x_0$ and the corresponding integral for $x>x_0$. Thus, $\sigma(x)$ may have different functional forms on the two sides of the singular point. However, $\sigma(x_0)=x_0$, so that the presence of singular points in $S(x)$ need not introduce discontinuities or singularities in $\sigma(x)$.

[25] If $f(x_0)=0$, $S(x_0)$ need not be singular.