

Hellmann-Feynman and Virial Theorems in the $X \alpha$ Method

John C. Slater

Citation: *The Journal of Chemical Physics* **57**, 2389 (1972); doi: 10.1063/1.1678599

View online: <http://dx.doi.org/10.1063/1.1678599>

View Table of Contents: <http://scitation.aip.org/content/aip/journal/jcp/57/6?ver=pdfcov>

Published by the AIP Publishing

Articles you may be interested in

[Hellmann-Feynman and virial theorems and the total energy of molecules](#)

J. Chem. Phys. **65**, 196 (1976); 10.1063/1.432795

[Generalized Hellmann-Feynman theorems and hypervirial theorems for \$X \alpha\$ -like methods](#)

J. Chem. Phys. **60**, 3328 (1974); 10.1063/1.1681525

[Integral Virial Theorem and Generalized Differential Hellmann-Feynman Formula](#)

J. Chem. Phys. **48**, 1425 (1968); 10.1063/1.1668837

[Integral Hellmann-Feynman Theorem](#)

J. Chem. Phys. **41**, 2892 (1964); 10.1063/1.1726371

[Derivation of the Virial Theorem from the Generalized Hellmann-Feynman Theorem](#)

J. Chem. Phys. **25**, 1299 (1956); 10.1063/1.1743226

How can you **REACH 100%**
of researchers at the Top 100
Physical Sciences Universities?
(TIMES HIGHER EDUCATION RANKINGS, 2014)

With *The Journal of Chemical Physics*.

AIP | The Journal of
Chemical Physics

THERE'S POWER IN NUMBERS. Reach the world with AIP Publishing.



Hellmann-Feynman and Virial Theorems in the $X\alpha$ Method*

JOHN C. SLATER

Quantum Theory Project, Department of Physics and Astronomy, University of Florida, Gainesville, Florida 32601

(Received 7 June 1972)

It is proved that the $X\alpha$ method leads rigorously to the Hellmann-Feynman and virial theorems. The proof of the virial theorem is general, not limited to the scaling argument used in earlier discussions. These proofs hold only if α is a constant, independent of nuclear positions. If α has one value inside one atomic sphere, another in another sphere, there are small correction terms, so that if one wishes the theorems to be rigorously applicable, it is best to use the same α everywhere. Discussion of the van der Waals force indicates that the $X\alpha$ method should give a satisfactory treatment of this force.

I. THE IMPORTANCE OF THE HELLMANN-FEYNMAN AND VIRIAL THEOREMS

The Hellmann-Feynman and virial theorems are two of the most powerful theorems applicable to molecules and solids, provided they can be used. The Hellmann-Feynman theorem¹ considers a set of fixed nuclei with electrons moving among them according to the principles of wave mechanics, (the Born-Oppenheimer approximation). It states that the force component $-\partial\langle E\rangle/\partial X_p$ acting on a nuclear coordinate X_p , where $\langle E\rangle$ is the total energy of the system as calculated wave-mechanically, is the force which would be calculated by pure classical electrostatics acting on this nuclear coordinate, resulting from the electronic charge cloud and all other nuclei. From this theorem we can then in principle find the total energy as a function of nuclear positions, by integrating the force components, provided we know the charge density exactly for arbitrary nuclear positions. From the virial theorem,²

kinetic energy = $-\frac{1}{2}$ (potential energy)

$$-\frac{1}{2}\sum (\rho) X_p [\partial\langle E\rangle/\partial X_p], \quad (1)$$

we can proceed as in Ref. 2 to find the kinetic and potential energies separately as functions of the nuclear coordinates. This means that both kinetic and potential energies of the whole system can be rigorously derived from the electronic charge density, or one-electron density matrix.

The proofs of these theorems for the exact wave-mechanical problem^{1,2} are well known. Unfortunately the theorems have been rather unimportant in practice, on account of the fact that ordinarily we are dealing with rather crude approximate wavefunctions for which they do not hold. One can appreciate the required accuracy from the work of Hirschfelder and Eliason,³ who have studied the applicability of the theorems for two hydrogen atoms at a large distance R from each other. If all quantities are expanded in powers of $1/R$, it is well known that the leading term in the van der Waals attractive energy is in $1/R^6$. Hirschfelder and Eliason show that the wavefunction has a constant term, and then terms in R^{-3} , \dots , R^{-7} , \dots , and that terms up to that in R^{-7} must be known in order to derive the R^{-6} term in the van der Waals energy from the Hellmann-

Feynman theorem. References to other applications of the theorems are given in Ref. 3.

Recently the author, K. H. Johnson, F. C. Smith, Jr., J. W. D. Connolly, and others⁴ have discussed the applicability of the so-called $X\alpha$ method to molecules and solids. Here one sets up a self-consistent-field treatment of an atom, molecule, or solid, in terms of spin-orbitals which are determined by one-electron equations derived by the variation method from a many-electron Hamiltonian. This Hamiltonian differs from the rigorous wave-mechanical Hamiltonian only in one respect: It replaces the true exchange-correlation term in the energy by an approximation proportional to the $1/3$ power of the local charge density. This term is taken to be proportional to that suggested by the author⁵ in an earlier discussion of the method, but multiplied by a factor α different from unity. Kohn and Sham⁶ and Gaspar⁷ have given reasons for thinking that α should be $2/3$, but in Ref. 4 reasons are given for thinking that α might better be chosen to depend on atomic number. Lindgren and Schwarz⁸ have found that the optimum value of α varies from almost unity for a one-electron atom, about 0.77 for a two-electron atom, going down to slightly higher than $2/3$ for the heavier atoms. It appears that the inner shells of the atoms are better represented by using the higher α 's, the outer shells by lower values, and the values of Lindgren and Schwarz represent a sort of compromise, handling the atom as a whole as well as possible from the point of view of total energy. The one-electron orbitals, and total charge density, given by this method come out in close agreement with those found by the Hartree-Fock method, in cases in which that method is applicable.

In the present paper, it is proved that the $X\alpha$ method, for a case where α is constant everywhere, rigorously obeys the Hellmann-Feynman and virial theorems. The proof of the Hellmann-Feynman theorem for this approximation, which is very simple, has not, as far as the author knows, been given before. As for the virial theorem, discussions have been given by several authors⁹ using the scaling argument dating back to the early work of Fock.¹⁰ This argument leads only to those components of force concerned when the molecule or crystal is homogeneously expanded or contracted, so that it is not perfectly general. In the present treat-

ment, in contrast, we give the general type of discussion used by the author in Ref. 2, in which no limitations are placed on the types of displacements of the nuclei.

This means that if the calculations are rigorously made by the $X\alpha$ method, we should have the same advantages in the way of getting forces and total, kinetic, and potential energies from the charge density which one has in the exact wavefunction of the many-electron problem. As is mentioned in Sec. IV, this leads to the result that the van der Waals forces should be given with fair accuracy by the $X\alpha$ method. Recent computational results of Trickey and Averill¹¹ on solid argon and xenon indicate that we are justified in hoping that calculations of van der Waals forces will be possible by the method.

Existing calculations by the $X\alpha$ method make approximations in addition to those inherent in the $X\alpha$ approximation. In particular, they assume the so-called muffin-tin potential, and corresponding charge densities. The muffin-tin potential corresponds to the assumption that the potential is spherically symmetrical within a sphere surrounding each nucleus, and constant in the region between the spheres, an assumption which is obviously not justified. The calculated charge

density arising from the method is not of this type, but shows the nonspherical charge concentrations characteristic of covalent bonds.⁴ However, in carrying through the successive stages of iteration of the self-consistent-field calculation, and in the resulting calculation of total energy, these nonspherical charge distributions are disregarded, and an assumption corresponding to the muffin-tin approximation is applied to the charge distribution as well as to the potential. Thus the calculations of total energy as a function of nuclear position as they have been carried out up to the present do not represent the true results of the $X\alpha$ method, and it is of interest that even in their present crude form they still give remarkably good agreement with experiment. These points will be discussed more in detail in Sec. IV.

II. THE HELLMANN-FEYNMAN THEOREM

We shall now go on to the proofs of the two theorems under discussion for the $X\alpha$ method. First we take up the Hellmann-Feynman theorem. We start with the $X\alpha$ expression for total energy, in terms of electronic coordinates $x_1y_1z_1\cdots$ and nuclear coordinates $X_1Y_1Z_1\cdots$. This energy, as in Ref. 4, is

$$\begin{aligned} \langle EX\alpha \rangle = & \sum (i)n_i \int u_i^*(1) [-\nabla_1^2 - \sum (p) (2Z_p/|\mathbf{r}_1 - \mathbf{R}_p|)] u_i(1) dv_1 \\ & + \frac{1}{2} \sum (i,j)n_i n_j \int u_i^*(1) u_j^*(2) u_i(1) u_j(2) (2/|\mathbf{r}_1 - \mathbf{r}_2|) dv_1 dv_2 \\ & - (9\alpha/2) (3/4\pi)^{1/3} \int \{ [\sum (i\uparrow) n_i u_i^*(1) u_i(1)]^{4/3} + [\sum (i\downarrow) n_i u_i^*(1) u_i(1)]^{4/3} \} dv_1 \\ & + \sum (\text{pairs } p, q) (2Z_p Z_q / |\mathbf{R}_p - \mathbf{R}_q|). \quad (2) \end{aligned}$$

Here the spin-orbitals are u_i , each associated either with spin up or spin down. The occupation numbers are n_i . Rydbergs are used as units of energy, atomic units as units of distance. The first term is the kinetic energy, plus the potential energy of the electrons in the field of the nuclei. The p th nucleus has charge Z_p , and $|\mathbf{r}_1 - \mathbf{R}_p|$ is the distance from this nucleus to electron 1. The second term is the Coulomb energy of interaction between the electronic charge distribution and itself, including the self-interaction terms $i=j$.

The third term is the exchange-correlation term, divided into two summations, first for those spin-orbitals with spin up, then for those with spin down. The last term is the Coulomb energy of interaction between pairs of nuclei, where the case $p=q$ is excluded.

We determine the spin-orbitals u_i by the variation principle, varying each spin-orbital separately and demanding that the energy $\langle EX\alpha \rangle$ be stationary (or for the ground state, a minimum) with respect to the variation. This leads to the one-electron equations

$$\begin{aligned} \{ -\nabla_1^2 - \sum (p) (2Z_p/|\mathbf{r}_1 - \mathbf{R}_p|) + \sum (j) n_j \int u_j^*(2) u_j(2) (2/|\mathbf{r}_1 - \mathbf{r}_2|) dv_2 \\ - 6\alpha (3/4\pi)^{1/3} [\sum (j\uparrow) n_j u_j^*(1) u_j(1)]^{1/3} \} u_i(1) = \epsilon_i u_i(1), \quad (3) \end{aligned}$$

where we have written the equation for a spin-orbital of spin up. The one-electron energy ϵ_i has the property⁴ that

$$\epsilon_i = \partial \langle EX\alpha \rangle / \partial n_i, \quad (4)$$

which leads to the result that the Fermi statistics apply rigorously to the method. In particular, if we fill all spin-orbitals for which ϵ_i is less than the Fermi energy ϵ_F , so that each of these has $n_i=1$, and empty all those

whose ϵ_i is greater than ϵ_F , we can prove that the energy of Eq. (2) has the lowest value which it can attain.

We now wish to compute the quantity $-\partial \langle EX\alpha \rangle / \partial X_p$, which appears in Eq. (1), and which is the x component of the force exerted by the atomic system on the p th nuclear coordinate, to be balanced by an external force $\partial \langle EX\alpha \rangle / \partial X_p$ if the nuclei are to be held fixed. If we forget that the u_i 's implicitly depend on the nuclear positions, we should differentiate in Eq. (2) under the integral sign with respect to X_p where it appears ex-

plicitly in $|\mathbf{r}_1 - \mathbf{R}_p|$ and in the last term. That is, we should conclude that

$$-\partial\langle EX\alpha\rangle/\partial X_p = \sum (i)n_i \int u_i^*(1)u_i(1)(\partial/\partial X_p)(2Z_p/|\mathbf{r}_1 - \mathbf{R}_p|)dv_1 - \sum (q \neq p)(\partial/\partial X_p)(2Z_p Z_q/|\mathbf{R}_p - \mathbf{R}_q|). \quad (5)$$

The first term represents the forces, computed by classical electrostatics, exerted by the total electronic charge distribution on the p th nucleus, while the second term is the force exerted by all nuclei except the p th, on the p th nucleus. The Hellmann-Feynman theorem, which holds rigorously if we use the true many-electron Hamiltonian instead of Eq. (2), states that Eq. (5) is true, for the reason that the additional terms arising from the implicit dependence of the u_i 's on X_p exactly cancel. We shall now show that except for one small point, the same theorem holds with the Hamiltonian of Eq. (1).

The spin-orbitals u_i really depend on the X_p 's since obviously they are tied to the nuclear positions. Hence there must be terms inside the integral involving

$$\partial u_i^*(x_1 \cdots X_p \cdots)/\partial X_p \quad (6)$$

and similarly in terms of the derivatives of the u_i 's. The quantity multiplying the expression of Eq. (6) inside the integral will be given by $-n_i$ times the left side of Eq. (3). On account of Eq. (3), we can replace this by $-n_i$ times the right side of Eq. (3). Thus we shall have

$$-\partial\langle EX\alpha\rangle/\partial X_p = -\sum (i)n_i \left\{ \int [\partial u_i^*(1)/\partial X_p] \epsilon_i u_i(1) dv_1 + \int u_i^*(1) \epsilon_i [\partial u_i(1)/\partial X_p] dv_1 \right\} + \text{terms of Eq. (5)}. \quad (7)$$

On account of the normalization of the u_i 's, we must have

$$\begin{aligned} (\partial/\partial X_p) \int u_i^*(1)u_i(1)dv_1 &= 0 = \int [\partial u_i^*(1)/\partial X_p] u_i(1) dv_1 \\ &+ \int u_i^*(1) [\partial u_i(1)/\partial X_p] dv_1. \end{aligned} \quad (8)$$

Thus we see that these additional terms vanish, and Eq. (5) is satisfied, which we wished to prove.

The small point which we have mentioned earlier, in which this derivation is incomplete, appears if α is not a constant holding for all positions of the nuclei,

but instead has one value in the sphere centered on one nucleus, another in the sphere centered on another nucleus, and so on, as we actually assume in the application of the $X\alpha$ method to compounds. This would introduce small correction terms depending on $\partial\alpha/\partial X_p$, which would have to be considered in a rigorous calculation of the force on the p th nucleus, though they have not been investigated up to the present time. They would not enter into the study of a molecule or crystal formed from a single type of atoms. They can be evaluated in terms of surface integrals over the surfaces of the various atomic spheres.

III. THE VIRIAL THEOREM

We have mentioned that several authors, cited in Ref. 9, have proved the virial theorem from the scaling argument used by Fock in Ref. 10. However, we shall now give a more general proof along the lines of that given by the author in Ref. 2. We start with the one-electron equation of the $X\alpha$ method, Eq. (3). We take the operator $x_1(\partial/\partial x_1) + y_1(\partial/\partial y_1) + z_1(\partial/\partial z_1)$ applied to this equation, multiply on the left by $u_i^*(1)$, and integrate over the coordinates of the electron. The result is

$$\begin{aligned} \int u_i^*(1) \left(x_1 \frac{\partial}{\partial x_1} + y_1 \frac{\partial}{\partial y_1} + z_1 \frac{\partial}{\partial z_1} \right) [-\nabla^2 u_i(1)] dv_1 &+ \int u_i^*(1) \left(x_1 \frac{\partial}{\partial x_1} + y_1 \frac{\partial}{\partial y_1} + z_1 \frac{\partial}{\partial z_1} \right) \left[-\sum (p) \frac{2Z_p}{|\mathbf{r}_1 - \mathbf{R}_p|} \right. \\ &\left. + \sum (j)n_j \int \frac{u_j^*(2)u_j(2)}{|\mathbf{r}_1 - \mathbf{r}_2|} dv_2 - 6\alpha \left(\frac{3}{4\pi} \right)^{1/3} [\sum (j \uparrow) n_j u_j^*(1)u_j(1)]^{1/3} - \epsilon_i \right] u_i(1) dv_1 = 0. \end{aligned} \quad (9)$$

In the second term of Eq. (9), the operator $x_1(\partial/\partial x_1) + y_1(\partial/\partial y_1) + z_1(\partial/\partial z_1)$ operates both on the quantity in brackets, and on $u_i(1)$. Thus Eq. (9) becomes transformed into

$$\begin{aligned} \int u_i^*(1) \left(x_1 \frac{\partial}{\partial x_1} + y_1 \frac{\partial}{\partial y_1} + z_1 \frac{\partial}{\partial z_1} \right) [-\nabla^2 u_i(1)] dv_1 &+ \int u_i^*(1)u_i(1) \left(x_1 \frac{\partial}{\partial x_1} + y_1 \frac{\partial}{\partial y_1} + z_1 \frac{\partial}{\partial z_1} \right) \left[-\sum (p) \frac{2Z_p}{|\mathbf{r}_1 - \mathbf{R}_p|} \right. \\ &+ \sum (j)n_j \int \frac{u_j^*(2)u_j(2)}{|\mathbf{r}_1 - \mathbf{r}_2|} dv_2 - 6\alpha \left(\frac{3}{4\pi} \right)^{1/3} [\sum (j \uparrow) n_j u_j^*(1)u_j(1)]^{1/3} - \epsilon_i \Big] dv_1 \\ &+ \int \left[-\sum (p) \frac{2Z_p}{|\mathbf{r}_1 - \mathbf{R}_p|} + \sum (j)n_j \int \frac{u_j^*(2)u_j(2)}{|\mathbf{r}_1 - \mathbf{r}_2|} dv_2 \right. \\ &\left. - 6\alpha \left(\frac{3}{4\pi} \right)^{1/3} [\sum (j \uparrow) n_j u_j^*(1)u_j(1)]^{1/3} - \epsilon_i \right] u_i^*(1) [x_1(\partial/\partial x_1) + y_1(\partial/\partial y_1) + z_1(\partial/\partial z_1)] u_i(1) dv_1 = 0. \end{aligned} \quad (10)$$

In the last term of Eq. (10) we can use the complex conjugate of Eq. (3) to transform the integral into

$$\int [\nabla_1^2 u_i^*(1)] [x_1(\partial/\partial x_1) + y_1(\partial/\partial y_1) + z_1(\partial/\partial z_1)] u_i(1) dv_1. \quad (11)$$

Thus in place of Eq. (9) we have

$$\begin{aligned} & \int u_i^*(1) \left(x_1 \frac{\partial}{\partial x_1} + y_1 \frac{\partial}{\partial y_1} + z_1 \frac{\partial}{\partial z_1} \right) [-\nabla_1^2 u_i(1)] dv_1 + \int [\nabla_1^2 u_i^*(1)] \left(x_1 \frac{\partial}{\partial x_1} + y_1 \frac{\partial}{\partial y_1} + z_1 \frac{\partial}{\partial z_1} \right) u_i(1) dv_1 \\ & + \int u_i^*(1) u_i(1) \left(x_1 \frac{\partial}{\partial x_1} + y_1 \frac{\partial}{\partial y_1} + z_1 \frac{\partial}{\partial z_1} \right) \left[-\sum (p) \frac{2Z_p}{|\mathbf{r}_1 - \mathbf{R}_p|} + \sum (j) n_j \int u_j^*(2) u_j(2) \frac{2}{|\mathbf{r}_1 - \mathbf{r}_2|} dv_2 \right. \\ & \left. - 6\alpha \left(\frac{3}{4\pi} \right)^{1/3} [\sum (j \uparrow) n_j u_j^*(1) u_j(1)]^{1/3} - \epsilon_i \right] dv_1 = 0. \quad (12) \end{aligned}$$

We can now use an identity proved in Ref. 2. This identity can be written in the form

$$\sum (j) x_j \left(u^* \frac{\partial^3 u}{\partial x_i^2 \partial x_j} - \frac{\partial^2 u^*}{\partial x_i^2} \frac{\partial u}{\partial x_j} \right) = -2u^* \frac{\partial^2 u}{\partial x_i^2} + \frac{\partial}{\partial x_i} \left[u^* \frac{\partial}{\partial x_i} \left(\frac{\sum (j) x_j \partial u / \partial x_j}{u^*} \right) \right], \quad (13)$$

where u is any function. This can be proved, after a little manipulation, by performing the differentiations indicated in the last term on the right. If we integrate the right side of Eq. (13) over the coordinates, the derivative integrates to zero if we take the limits of integration at infinity, and assume that $u=0$ there, as it will be if u is the wavefunction of a bound electron in a finite (but arbitrarily large) cluster of atoms. When we use this result, we find that the sum of the first two integrals of Eq. (12) is $2 \int u_i^*(1) \nabla_1^2 u_i(1) dv_1$. We multiply by $-n_i$, sum over all spin-orbitals, and find from Eq. (12) that

$$\begin{aligned} 2 \sum (i) n_i \int u_i^*(1) [-\nabla_1^2 u_i(1)] dv_1 &= 2KE = \sum (i) n_i \int u_i^*(1) u_i(1) \left(x_1 \frac{\partial}{\partial x_1} + y_1 \frac{\partial}{\partial y_1} + z_1 \frac{\partial}{\partial z_1} \right) \\ & \times \left[-\sum (p) \frac{2Z_p}{|\mathbf{r}_1 - \mathbf{R}_p|} + \sum (j) n_j \int u_j^*(2) u_j(2) \frac{2}{|\mathbf{r}_1 - \mathbf{r}_2|} dv_2 \right. \\ & \left. - 6\alpha \left(\frac{3}{4\pi} \right)^{1/3} [\sum (j \uparrow) n_j u_j^*(1) u_j(1)]^{1/3} - \epsilon_i \right] dv_1, \quad (14) \end{aligned}$$

where KE is the kinetic energy. We must now work out the expressions on the right, to transform the resulting equation into the familiar formulation of the virial theorem.

We first have

$$\begin{aligned} \left(x_1 \frac{\partial}{\partial x_1} + y_1 \frac{\partial}{\partial y_1} + z_1 \frac{\partial}{\partial z_1} \right) \left(-\sum (p) \frac{2Z_p}{|\mathbf{r}_1 - \mathbf{R}_p|} \right) &= \sum (p) 2Z_p \frac{x_1(x_1 - R_{px}) + y_1(y_1 - R_{py}) + z_1(z_1 - R_{pz})}{|\mathbf{r}_1 - \mathbf{R}_p|^3} \\ &= \sum (p) \frac{2Z_p}{|\mathbf{r}_1 - \mathbf{R}_p|} - \left(R_{px} \frac{\partial}{\partial R_{px}} + R_{py} \frac{\partial}{\partial R_{py}} + R_{pz} \frac{\partial}{\partial R_{pz}} \right) \left(-\sum (q) \frac{2Z_q}{|\mathbf{r}_1 - \mathbf{R}_q|} \right). \quad (15) \end{aligned}$$

Thus the first term on the right side of Eq. (14) becomes

$$\begin{aligned} \sum (i) n_i \int u_i^*(1) u_i(1) \sum (p) \frac{2Z_p}{|\mathbf{r}_1 - \mathbf{R}_p|} dv_1 &- \sum (p) \left(R_{px} \frac{\partial}{\partial R_{px}} + R_{py} \frac{\partial}{\partial R_{py}} + R_{pz} \frac{\partial}{\partial R_{pz}} \right) \\ & \times \sum (i) n_i \int u_i^*(1) u_i(1) \left(-\sum (q) \frac{2Z_q}{|\mathbf{r}_1 - \mathbf{R}_q|} \right) dv_1. \quad (16) \end{aligned}$$

Next we take the second term on the right of Eq. (14). This may be rewritten in the form

$$\sum (i, j) n_i n_j \int u_i^*(1) u_j^*(2) u_i(1) u_j(2) [x_1(\partial/\partial x_1) + \dots] (2/|\mathbf{r}_1 - \mathbf{r}_2|) dv_1 dv_2 \quad (17)$$

since $u_j^*(2) u_j(2)$ does not depend on the coordinates x_1 , etc. We have

$$\left(x_1 \frac{\partial}{\partial x_1} + y_1 \frac{\partial}{\partial y_1} + z_1 \frac{\partial}{\partial z_1} \right) \frac{2}{|\mathbf{r}_1 - \mathbf{r}_2|} = -2 \frac{x_1(x_1 - x_2) + y_1(y_1 - y_2) + z_1(z_1 - z_2)}{|\mathbf{r}_1 - \mathbf{r}_2|^3}. \quad (18)$$

In Eq. (17), we can rewrite the expression as

$$\begin{aligned} \frac{1}{2} \sum (i, j) n_i n_j \int [u_i^*(1) u_i(1) u_j^*(2) u_j(2) + u_j^*(1) u_j(1) u_i^*(2) u_i(2)] & \left(x_1 \frac{\partial}{\partial x_1} + y_1 \frac{\partial}{\partial y_1} + z_1 \frac{\partial}{\partial z_1} \right) \frac{2}{|\mathbf{r}_1 - \mathbf{r}_2|} dv_1 dv_2 \\ &= \frac{1}{2} \sum (i, j) n_i n_j \int u_i^*(1) u_i(1) u_j^*(2) u_j(2) \left(-2 \frac{x_1(x_1 - x_2) + \dots}{|\mathbf{r}_1 - \mathbf{r}_2|^3} - 2 \frac{x_2(x_2 - x_1) + \dots}{|\mathbf{r}_1 - \mathbf{r}_2|^3} \right) dv_1 dv_2 \\ &= \frac{1}{2} \sum (i, j) n_i n_j \int u_i^*(1) u_i(1) u_j^*(2) u_j(2) \left(-\frac{2}{|\mathbf{r}_1 - \mathbf{r}_2|} \right) dv_1 dv_2. \end{aligned} \quad (19)$$

The third term on the right of Eq. (14) is the one involving the $1/3$ power of the charge density. In setting up this term, we must specify more precisely than previously how the summation over spin-orbitals is to be carried out. For those spin-orbitals with spin up, we are to use the expression of Eq. (14), but for those with spin down, we must take the $1/3$ power of the charge density with spin down. In either case, we are faced with an expression like

$$\int \rho(1) [x_1(\partial/\partial x_1) + \dots] [\rho(1)]^{1/3} dv_1, \quad (20)$$

where ρ is of the nature of $\sum (i) n_i u_i^*(1) u_i(1)$. We can integrate each term of this by parts. Thus,

$$\begin{aligned} \int \rho(x) x(\partial/\partial x) [\rho(x)]^{1/3} dx &= x [\rho(x)]^{4/3} - \int [\rho(x)]^{1/3} d[\rho(x)x] \\ &= x [\rho(x)]^{4/3} - \int [\rho(x)]^{4/3} dx - \int [\rho(x)]^{1/3} x [\partial\rho(x)/\partial x] dx. \end{aligned} \quad (21)$$

We shall carry the integrations over dx out far enough so that the integrated terms will vanish. We note that

$$(\partial/\partial x) [\rho(x)]^{1/3} = 1/3 [\rho(x)]^{-2/3} [\partial\rho(x)/\partial x]. \quad (22)$$

As a result of this, we have

$$\int \rho(x) x(\partial/\partial x) [\rho(x)]^{1/3} dx = \frac{1}{3} \int [\rho(x)]^{1/3} x [\partial\rho(x)/\partial x] dx \quad (23)$$

from which, again disregarding the integrated terms,

$$\int \rho(x) x(\partial/\partial x) [\rho(x)]^{1/3} dx = -\frac{1}{4} \int [\rho(x)]^{4/3} dx. \quad (24)$$

We can then integrate this over dy and dz , to get the volume integrals and can add the terms arising from y and z . We then have

$$\int \rho(1) [x_1(\partial/\partial x_1) + y_1(\partial/\partial y_1) + z_1(\partial/\partial z_1)] [\rho(1)]^{1/3} dv_1 = -\frac{3}{4} \int [\rho(1)]^{4/3} dv_1. \quad (25)$$

When we use this result, the third term on the right of Eq. (14) becomes

$$\frac{3}{2} \alpha (3/4\pi)^{1/3} \int \{ [\rho \uparrow (1)]^{4/3} + [\rho \downarrow (1)]^{4/3} \} dv_1. \quad (26)$$

The last term on the right of Eq. (14), involving ϵ_i , gives zero, since ϵ_i is a constant and its derivatives are zero.

We can now combine our results. We find from Eqs. (14), (16), (19), and (26), that

$$\begin{aligned} 2KE &= \int \rho(1) \sum (q) (2Z_q/|\mathbf{r}_1 - \mathbf{R}_q|) dv_1 - \frac{1}{2} \int \rho(1) \rho(2) (2/|\mathbf{r}_1 - \mathbf{r}_2|) dv_1 dv_2 + (9/2) \alpha (3/4\pi)^{1/3} \int \{ [\rho \uparrow (1)]^{4/3} \\ &+ [\rho \downarrow (1)]^{4/3} \} dv_1 - \sum (q) [R_{qx}(\partial/\partial R_{qx}) + R_{qy}(\partial/\partial R_{qy}) + R_{qz}(\partial/\partial R_{qz})] \int \rho(1) (2Z_q/|\mathbf{r}_1 - \mathbf{R}_q|) dv_1, \end{aligned} \quad (27)$$

where as before we write $\rho(1)$ instead of $\sum (i) n_i u_i^*(1) u_i(1)$. We recognize on the right side of Eq. (27) the negative of the potential energy of interaction between the electrons and nuclei, the negative of the potential energy of repulsion between the electrons, the negative of the exchange-correlation term in the potential energy, as well as the last term. Let us then add and subtract the negative of the potential energy of repulsion between the nuclei, which is $\sum (\text{pairs } p, q) 2Z_p Z_q / |\mathbf{R}_p - \mathbf{R}_q|$. We find

$$\begin{aligned} 2KE &= -PE + \sum (q) [R_{qx}(\partial/\partial R_{qx}) + R_{qy}(\partial/\partial R_{qy}) + R_{qz}(\partial/\partial R_{qz})] \int \rho(1) (2Z_q/|\mathbf{r}_1 - \mathbf{R}_q|) dv_1 \\ &+ \sum (\text{pairs } p, q) (2Z_p Z_q / |\mathbf{R}_p - \mathbf{R}_q|). \end{aligned} \quad (28)$$

But one can prove, by methods similar to those used above, that

$$\sum (\text{pairs } p, q) 2Z_p Z_q / |\mathbf{R}_p - \mathbf{R}_q| = - \sum (p, q \neq p) [R_{qx}(\partial/\partial R_{qx}) + R_{qy}(\partial/\partial R_{qy}) + R_{qz}(\partial/\partial R_{qz})] (2Z_p Z_q / |\mathbf{R}_p - \mathbf{R}_q|). \quad (29)$$

Hence one can rewrite Eq. (27) in the form

$$\begin{aligned} 2KE &= -PE - \sum (q) [R_{qx}(\partial/\partial R_{qx}) + R_{qy}(\partial/\partial R_{qy}) + R_{qz}(\partial/\partial R_{qz})] \\ &\times [\sum (p \neq q) (2Z_p Z_q / |\mathbf{R}_p - \mathbf{R}_q|) - \int \rho(1) (2Z_q/|\mathbf{r}_1 - \mathbf{R}_q|) dv_1]. \end{aligned} \quad (30)$$

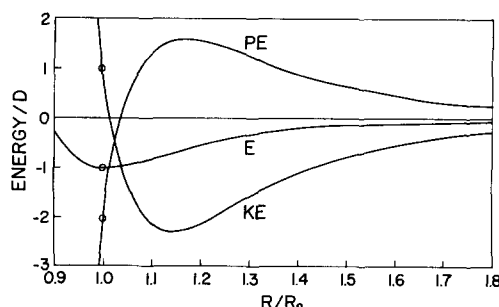


FIG. 1. Total energy E , potential energy PE , and kinetic energy KE , derived from the Lennard-Jones potential by use of the virial theorem, as a multiple of the dissociation energy.

We now use the Hellmann-Feynman theorem, Eq. (5), and divide by 2. We have

$$KE = -\frac{1}{2}PE - \frac{1}{2}\sum (p)X_p(\partial\langle EX\alpha\rangle/\partial X_p) \quad (31)$$

which is the virial theorem, as in Eq. (1), for the $X\alpha$ method.

IV. DISCUSSION

In Sec. I we have pointed out that if we know the charge density exactly as a function of nuclear positions, and if we have a single constant α , we can use the Hellmann-Feynman and virial theorems to find the total energy, kinetic energy, and potential energy as functions of nuclear positions. The calculations which have been made by the $X\alpha$ method give the impression of yielding very accurate charge densities, and this procedure then may be practical with the $X\alpha$ method. However, as yet it has not been tried. What we have at present is the calculation of total energy, kinetic energy, and potential energy of a number of solids as functions of volume,^{11,12,13} and of the same quantities for several molecules¹⁴ as a function of nuclear positions. These calculations have been made using the muffin-tin approximation, and yet it is encouraging to find that the virial theorem proves to be very accurately satisfied. This encourages us to think that the errors arising from the muffin-tin approximation may be rather small. There are indications, however, that for the molecules¹⁴ these errors are more serious than for crystals.

At the present stage of the work with the $X\alpha$ method, it is more practical to use the Hellmann-Feynman theorem in a reverse direction, to indicate the type of deviations from spherically symmetrical charge distributions which we must have in order to explain the calculated energy curves as functions of nuclear position. Such applications of the Hellmann-Feynman theorem have been made, without thinking of the $X\alpha$ method, by Berlin¹⁵ in a well-known paper, by Bader¹⁶ and associates in a series of papers, and by others. Let us see what sort of information can be obtained from the known behavior of the energy as a function of nuclear position, by use not only of the virial theorem but also of the Hellmann-Feynman theorem. Such information was discussed in Ref. 2 with respect to the

virial theorem. It is best presented in terms of a specific example, and we shall choose that of two rare gas atoms attracting at large distances by van der Waals attractions, repelling at short distances by the interactions of the closed shells. Our reason for choosing this case is that the work of Trickey and Averill¹ indicates that the $X\alpha$ method works well for the rare gas crystals, whereas various physicists, on being informed of this fact, have been very skeptical as to whether the $X\alpha$ method was even able in principle to explain the van der Waals attraction.

It is well known that the van der Waals attraction at large distances goes as R^{-6} , as we have mentioned earlier. Furthermore, it is well known that by supplementing this R^{-6} attractive energy by a R^{-12} repulsive energy, giving the so-called Lennard-Jones potential, we can represent the energy curve of interaction of two rare gas atoms reasonably accurately. The calculations of Trickey and Averill¹¹ show that the $X\alpha$ results can be adequately approximated by the Lennard-Jones potential. Let us then assume that the interatomic energy of two atoms, as a function of R , is given by

$$E = D[-(2/x^6) + (1/x^{12})], \quad \text{where } x = R/R_0. \quad (32)$$

Here the two constants D and R_0 involved in the Lennard-Jones potential are chosen so that D is the dissociation energy, R_0 the equilibrium separation. From the virial theorem, we can show, as is indicated in Ref. 2, that

$$KE = -E - R dE/dR, \quad PE = 2E + R dE/dR \quad (33)$$

from which for the Lennard-Jones potential

$$\begin{aligned} KE &= D[-(10/x^6) + (11/x^{12})], \\ PE &= D[(8/x^6) - (10/x^{12})]. \end{aligned} \quad (34)$$

Curves for E/D , KE/D , and PE/D are shown in Fig. 1. It is to be noted that E , KE , and PE in these formulas represent the differences between the energies at distance R , and those for infinite internuclear distance.

The curves of Fig. 1, like similar ones shown in Ref. 2, lead to simple and unambiguous information as to the way in which the charge density is changing as the atoms are brought together from infinite distance. The most striking feature at large distances is the fact that the potential energy rises, the kinetic energy falls, with decreasing R , while the total energy has a much smaller variation with R . This means that there must be a shift of electronic charge from the inner parts of the atoms to the peripheral regions. The small change of total energy is a result of the variation principle: the type of shift of charge we have mentioned may produce quite large changes in the potential and kinetic energies separately, with relatively small change in the total energy. The virial theorem by itself cannot tell us where on the periphery of the atoms the added charge must reside. However, the Hellmann-Feynman theorem, as Berlin¹⁵ pointed out, can give us definite information.

It must correspond to a nonspherical charge distribution, with excess electronic density in the region between the nuclei, the region which Berlin calls the binding region. For then each nucleus will be attracted to this electronic charge in the binding region. The concentration of electronic charge in this binding region was found by Hirschfelder and Eliason³ in their study of the hydrogen molecule.

From this explanation of the binding part of the curve, we can see why it is that the muffin-tin correction is not very important in the application of the virial theorem. The increase of potential energy, decrease of kinetic energy, which we have noted, depends primarily on the radial displacement of electrons outward in the two atoms, and would be represented with good approximation even if we disregarded the fact that the charge was actually mostly on the side of the atom facing its neighbor. On the other hand, for the application of the Hellmann-Feynman theorem, it would be absolutely necessary to have a charge density which showed the proper angular behavior. As we have mentioned, calculations of local charge density reported in Ref. 4 indicate that the $X\alpha$ method is at least qualitatively representing the angular distribution properly, so that there is good hope that the existing calculations could be used for the application of the Hellmann-Feynman theorem.

Next we ask what happens at smaller values of interatomic distance. As we approach the equilibrium distance, the potential energy drops precipitously, and the kinetic energy rises. This is the beginning of the process of compression of the atoms, as they are pushed together. The electronic distribution is compressed to smaller volume, bringing the electrons closer to the nucleus, and decreasing the wavelengths of the wavefunctions, thereby increasing the kinetic energy. It is natural that this compression is most severe on the side of each atom facing its neighbor, and there is some tendency of the electron cloud to be pushed to the back of each atom, away from its neighbor. In this region, Berlin's antibinding region, the resulting force on each nucleus is an attraction toward this electronic charge, which amounts to a repulsion by the other atom.

The description which we have given of this process is precisely the same which was given in Ref. 2 to explain covalent binding. The writer believes that there is no very fundamental distinction between the van der Waals binding and covalent binding. If we refer them back to the electronic charge distribution, which as we have seen is solely responsible for the interatomic forces, we have the same type of behavior in both cases. And there is no reason to doubt that the $X\alpha$ method can give as good an account of the process in one case as in the other. Thus it is hardly surprising that Trickey and Averill¹¹ are getting reasonable values for the cohesive energy of solid argon and xenon. These cases are of course different from that of a diatomic molecule, in that each atom is equally surrounded by a number

of neighbors. At the larger distances, there will be an excess of electrons on the periphery of the atom facing each of the neighbors, and resulting in practically a spherically symmetrical distortion of the charge. Hence it is not at all surprising that the muffin-tin approximation is of small importance in such a case.

It is of course obvious that in the argon crystal, there is no net force on any nucleus, as the Hellmann-Feynman theorem would at once show. This is sometimes thought to be a difficulty with the interpretation of the theorem. It is not, however, as we can see if we consider a crystal with a boundary. The atoms in the interior will have no net force acting on them, but a surface atom will have a nonspherical charge distribution surrounding it, much as if it were in a diatomic molecule. If the interatomic distance is greater than that corresponding to equilibrium, there will be excess electronic charge on each of the surface atoms on the side facing the crystal. Thus each of the surface nuclei will have a force in toward the crystal, which must be balanced by an external outward tension. At a smaller interatomic distance than that of equilibrium, the excess electronic charge will be on the outer face of the crystal, and will pull each of the surface nuclei outwards. This must be balanced by an inward force on these surface atoms, and the vector sum of those forces on all surface atoms in unit area will be just the pressure. In other words, to get a proper understanding of the Hellmann-Feynman theorem we must consider a finite crystal or other body, and must make specific study of its surface atoms. This of course is not necessary if we use the virial theorem instead, and determine the pressure as usual from the change of energy with volume.

Let us now turn our attention to a different problem, the necessity of using a constant value of α everywhere in order to have the virial and Hellmann-Feynman theorems holding rigorously. We have mentioned the work of Lindgren and Schwarz,⁸ indicating that the values of α which we have been using in the $X\alpha$ method form a sort of compromise between the lower α which is preferable for the outer orbitals of the atom, and the higher values required for the inner shells. Now the cohesive energy properties which we are discussing in this paper are largely arising from the outer orbitals, and it is natural to ask if we might not do all right by using a constant value of α , which might well be $2/3$, for all these outer orbitals. In the muffin-tin method this might amount to using $2/3$ outside the spheres, but different α 's inside the spheres. This by itself, however, would not be enough to assure the correctness of the virial and Hellmann-Feynman theorems. It is worthwhile asking whether we might expect to get reasonable values by using the Kohn-Sham⁶ value $2/3$ everywhere. As we shall show, this is not unlikely.

We have first some computational evidence that this may be a good procedure. Averill¹² in unpublished work has studied the cohesive energy of the alkali metals, first using the α of Lindgren and Schwarz,⁸ second using

$\alpha=2/3$. For the heaviest alkali he has studied, Cs, it makes very little difference which value is used. On the other hand, there are increasingly large differences between the two calculations as we go to lighter elements, and these differences are quite large, of the order of twenty percent, by the time we get to Li. The remarkable fact is that the cohesive energies found from $\alpha=2/3$ are in better agreement with experiment than those with the larger α 's.

We must realize, of course, that the energy of an atom, as computed by the $X\alpha$ method, is strongly dependent on the value of α used. Thus if we use $\alpha=2/3$ for Li, instead of the optimum value, which is approximately 0.77, the energy of the isolated Li atom will come out a good deal too high. But the energy of the crystal seems to have an almost identical error, so that in subtracting the one from the other, to get the cohesive energy, this error is largely canceled out. We can see by considering the Hellmann-Feynman theorem why this must be.

Let us ask what would be the effect on the charge density of using $\alpha=2/3$ for Li, instead of $\alpha=0.77$. There would be a numerically smaller exchange-correlation term, and therefore not such a deep potential well around the nucleus. It would be almost as if we had a nuclear charge which was too small. The result would be that the $1s$ orbital of Li would come out too large, or in a heavier atom, the orbitals of the inner shells would all be too large. This would not, however, make any major difference in the interatomic forces, as found by the Hellmann-Feynman theorem. The reason is that these inner shells would still be spherical, and by electrostatics a spherical shell exerts the same force on any charge as if it were concentrated at its center. In other words, the electrostatic force would not be modified by the slightly enlarged inner shells. The only effect which the enlargement would have would be the indirect one, that the slightly incorrect charge distribution of the inner shells would modify slightly the potential in which the outer electrons

moved, and hence the resulting molecular orbitals. This certainly seems like a small effect, and it may well be the explanation of the reason why Averill¹² is finding good values of cohesive energy for $\alpha=2/3$. On the other hand, of course we must remember that the use of a wrong α could have a considerable adverse effect on the calculation of x-ray or ESCA energy levels.

ACKNOWLEDGMENTS

The author is much indebted to Dr. F. W. Averill, Dr. J. W. D. Connolly, Dr. T. M. Hattox, Dr. K. H. Johnson, and Dr. S. B. Trickey, for valuable discussions, and for communicating the results of research before its publication.

* Supported in part by National Science Foundation Grant GH-32006.

¹ H. Hellmann, *Einführung in die Quantenchemie* (Deuticke, Leipzig, Germany, 1937); R. P. Feynman, *Phys. Rev.* **56**, 340 (1939).

² J. C. Slater, *J. Chem. Phys.* **1**, 687 (1933); *Quantum Theory of Molecules and Solids* (McGraw-Hill, New York, 1963), pp. 29-34, 57-59, 254-255.

³ J. O. Hirschfelder and M. A. Eliason, *J. Chem. Phys.* **47**, 1164 (1967).

⁴ J. C. Slater and K. H. Johnson, *Phys. Rev. B* **5**, 844 (1972); K. H. Johnson and F. C. Smith, Jr., *ibid.* **5**, 831 (1972); J. W. D. Connolly and K. H. Johnson, *Chem. Phys. Letters* **10**, 616 (1971), and many other references given in these papers.

⁵ J. C. Slater, *Phys. Rev.* **81**, 385 (1951); **82**, 538 (1951).

⁶ W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).

⁷ R. Gaspar, *Acta Phys. Akad. Sci. Hung.* **3**, 263 (1954).

⁸ I. Lindgren and K. Schwarz, *Phys. Rev. A* **5**, 542 (1972); K. Schwarz, *Phys. Rev. B* **5**, 2466 (1972).

⁹ M. Ross, *Phys. Rev.* **179**, 612 (1969); W. E. Rudge, *ibid.* **181**, 1033 (1969); L. J. Sham, *Phys. Rev. A* **1**, 169 (1970).

¹⁰ V. Fock, *Z. Physik* **63**, 855 (1950).

¹¹ S. B. Trickey and F. W. Averill, *Bull. Am. Phys. Soc. Ser. II*, **17**, 346 (1972), and unpublished work, with F. R. Green, Jr. ¹² F. W. Averill, *Phys. Rev. B* **4**, 3315 (1971), and unpublished work.

¹³ T. M. Hattox, Ph.D. thesis, University of Florida, Gainesville, Fla., 1972 (unpublished).

¹⁴ J. W. D. Connolly and K. H. Johnson (unpublished); J. W. D. Connolly and J. R. Sabin, *J. Chem. Phys.* **56**, 5529 (1972).

¹⁵ T. Berlin, *J. Chem. Phys.* **19**, 208 (1951).

¹⁶ R. F. W. Bader, *Can. J. Chem.* **38**, 2117 (1960), and later papers.