A Theoretical Perspective on the Bond Length Rule of Grochala, Albrecht, and Hoffmann

Paul W. Ayers and Robert G. Parr*

Department of Chemistry, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599

Grochala, Albrecht, and Hoffmann have recently observed that, for many organic and inorganic molecules, the bond lengths in the cation (R_+), neutral system (R_{gs}), first triplet excited state (R_{es}), and anion (R_-) approximately satisfy $R_+ + R_- - R_{gs} - R_{es} \approx 0$. This paper shows how the Fukui function, a concept from the density functional theory of chemical reactivity, can be used to rationalize the Grochala−Albrecht−Hoffmann rule. The treatment also suggests which types of chemical systems will be best described by the Grochala−Albrecht−Hoffmann rule.

I. Introduction

On the basis of results from a series of calculations, Grochala, Albrecht, and Hoffmann have proposed a new tool for understanding the geometry of molecules, observing that

$$R_+ + R_- - R_{gs} - R_{es} \approx 0 \qquad (1)$$

where R_+, R_-, R_{gs}, and R_{es} are the lengths of some bond for the cation, anion, singlet ground state, and first triplet excited state of a molecule, respectively.[1] Summarizing some of Grochala, Albrecht, and Hoffmann's results, eq 1 is most accurate when (1) the ground state of the molecule is nondegenerate and (2) the geometries of the cation, anion, singlet ground state, and first triplet excited state are similar.[1]

The density functional theory of chemical reactivity is useful for justifying existing chemical principles.[2] Density functional theory's principal tool for understanding the addition and subtraction of electrons is the Fukui function,[2−5] and in the present paper we use the Fukui function to derive the Grochala−Albrecht−Hoffmann (GAH) rule. Accordingly, our arguments involve formally exact theory rather than simple MO theory.

Our treatment reveals why the rule holds only for nondegenerate ground states in which the geometries of the cation, anion, ground state, and first excited state are similar and suggests other conditions which are necessary for the accurate fulfillment of the GAH rule. Finally, our derivation reveals how one may extend the GAH rule to degenerate ground states.

II. Theoretical Background

Within the Born−Oppenheimer approximation, the way the energy of a molecule changes when the nuclear positions change along some coordinate, λ (*keeping all other internuclear coordinates fixed*), is given by the Hellmann−Feynman theorem:

$$\frac{\partial E(\lambda)}{\partial \lambda} = \frac{\partial V_{nn}}{\partial \lambda} + \int \rho(\vec{r}_1;\lambda)\left(\frac{\partial \nu_0(\vec{r}_1;\lambda)}{\partial \lambda}\right)d\vec{r}_1 \qquad (2)$$

wherein we have introduced the electron density, $\rho(\vec{r}_1; \lambda)$, and the external potential, $\nu_0(\vec{r}_1; \lambda)$. In terms of the N-electron molecular wave function, the electron density is given by

$$\rho(\vec{r}_1;\lambda) \equiv N \int \int \int \Psi^*(\lambda)\,\Psi(\lambda)\,ds_1\,dx_2\,...\,dx_N \qquad (3)$$

where we have summed over the spin coordinates of all electrons and integrated over the spatial coordinates of the last $N - 1$ electrons. In atomic units, the external potential of a molecule is defined through

$$\nu_0(\vec{r}_1;\lambda) \equiv \sum_a -\frac{Z_\alpha}{|\vec{r}_1 - \vec{R}_\alpha(\lambda)|} \qquad (4)$$

where Z_α is the atomic number and $\vec{R}_\alpha(\lambda)$ is the position of the αth nucleus.

Equation 2 indicates that if the electron density is known for all λ, then $\partial E/\partial \lambda$ is also known for all λ. If we know $\partial E/\partial \lambda$ for all λ, then we can determine the potential energy surface, $E(\lambda)$, to within an additive constant; the constant depends on our choice of energy zero. In particular, knowledge of the electron density for every value of the coordinate λ allows us to determine the equilibrium value of λ, λ_0, since at λ_0

$$\left(\frac{\partial E}{\partial \lambda}\right)|_{\lambda = \lambda_0} = 0$$
$$\left(\frac{\partial^2 E}{\partial \lambda^2}\right)|_{\lambda = \lambda_0} > 0 \qquad (5)$$

Differentiating eq 2 gives an expression for $\partial^2 E/\partial \lambda^2$ in terms of the electron density, namely,

$$\frac{\partial^2 E}{\partial \lambda^2} = \frac{\partial^2 V_{nn}}{\partial \lambda^2} + \int \frac{\partial \rho(\vec{r}_1;\lambda)}{\partial \lambda}\left(\frac{\partial \nu_0(\vec{r}_1;\lambda)}{\partial \lambda}\right)dr_1 +$$
$$\int \rho(\vec{r}_1;\lambda)\left(\frac{\partial^2 \nu_0(\vec{r}_1;\lambda)}{\partial \lambda^2}\right)dr_1 \qquad (6)$$

The foregoing analysis indicates that if we are given the electron densities of the cation ($\rho_+(\vec{r};\lambda)$), anion ($\rho_-(\vec{r};\lambda)$), ground state ($\rho_{gs}(\vec{r};\lambda)$), and first excited state of a molecule ($\rho_{es}(\vec{r};\lambda)$) along some coordinate λ, then we can compute the potential energy curves, $E_+(\lambda)$, $E_-(\lambda)$, $E_{gs}(\lambda)$, and $E_{es}(\lambda)$, for these systems to within an additive constant. Moreover, given the appropriate densities, we can use eqs 2, 5, and 6 to find the equilibrium values of the coordinate λ for the cation (λ_+), anion (λ_-), ground state (λ_{gs}), and first excited state (λ_{es}). Hence, a relationship between the densities of the cation, anion, ground state, and

first excited state implies a relationship between their potential energy surfaces, which in turn implies a relationship between λ_+, λ_-, λ_{gs}, and λ_{es}.

This argument suggests a method for deriving the GAH rule. First derive a relationship between the densities of the cation, anion, ground state, and first excited state; we do this in section III.A. Use this result to derive relationships between the systems' potential energy surfaces (IIIB). Finally, use the relationships between the potential energy surfaces to derive a relationship between λ_+, λ_-, λ_{gs}, and λ_{es} (IIIC) and explore under what conditions this relationship becomes identical to the GAH rule (IIID). This method for deriving the GAH rule suggests an extension of the rule to degenerate ground states (IIIE).

III. Theoretical Development

A. Relationships Between the Densities. *Proposition 1.* Suppose that M is a molecule with a nondegenerate (and hence a singlet) ground state. Then, to a good approximation, the electron densities of the cation ($\rho_+(\vec{r};\lambda)$), anion ($\rho_-(\vec{r};\lambda)$), ground state ($\rho_{gs}(\vec{r};\lambda)$), and first excited state of a molecule ($\rho_{es}(\vec{r};\lambda)$) satisfy

$$\int [\rho_{gs}(\vec{r};\lambda) + \rho_{es}(\vec{r};\lambda) - \rho_+(\vec{r};\lambda) - \rho_-(\vec{r};\lambda)]\left(\frac{\partial v_0(\vec{r};\lambda)}{\partial \lambda}\right) d\vec{r} \approx 0 \quad (7)$$

for all λ.

Justification. If eq 7 holds for any particular value of λ, then one of the following three statements is true:

(i) $\dfrac{\partial v_0(\vec{r};\lambda)}{\partial \lambda = 0}$ for all $\vec{r}$

(ii) $\{\rho_{gs}(\vec{r};\lambda) + \rho_{es}(\vec{r};\lambda) - \rho_+(\vec{r};\lambda) - \rho_-(\vec{r};\lambda)\}$ is orthogonal to $\dfrac{\partial v_0(\vec{r};\lambda)}{\partial \lambda = 0}$

(iii) $\rho_{gs}(\vec{r};\lambda) + \rho_{es}(\vec{r};\lambda) - \rho_+(\vec{r};\lambda) - \rho_-(\vec{r};\lambda) = 0$ for all $\vec{r}$

Case (i) only occurs when the external potential, $v_0(\lambda)$, does not depend on the coordinate λ. Accordingly, case (i) never happens when λ is an internuclear coordinate of some type (see eq 4). While case (ii) may, perchance, occur for some exceptional value of λ, it certainly cannot cause eq 7 to be valid over a wide range of λ. Therefore, if the proposition is true for all λ, it must be that case (iii) is true for almost all λ (case (iii) may fail for that exceptional value of λ where case (ii) holds).

Within density functional theory,[2] the ground state electron density of a system is related to the ground state densities of its cation and its anion by the Fukui function from below, $f^-(\vec{r})$, and the Fukui function from above, $f^+(\vec{r})$, respectively:[2,3,5]

$$\rho_+(\vec{r};\lambda) = \rho_{gs}(\vec{r};\lambda) - f^-(\vec{r};\lambda) \quad (8)$$

$$\rho_-(\vec{r};\lambda) = \rho_{gs}(\vec{r};\lambda) + f^+(\vec{r};\lambda) \quad (9)$$

Substituting eqs 8 and 9 into eq 7 reveals that proposition 1 is true if and only if

$$\rho_{es}(\vec{r};\lambda) \approx \rho_{gs}(\vec{r};\lambda) - f^-(\vec{r};\lambda) + f^+(\vec{r};\lambda) \quad (10)$$

We now argue that eq 10 represents a good approximation to the excited state density. The excited state density of a system may be approximated by moving an electron from the HOMO orbital of the ground state to the LUMO orbital of the ground state. Neglecting the subsequent orbital relaxation, we have

$$\rho_{es}(\vec{r}) \approx \rho_{gs}(\vec{r}) - \rho_{HOMO}(\vec{r}) + \rho_{LUMO}(\vec{r}) \quad (11)$$

Orbital relaxation is not always negligible, however, and so we demonstrate that eq 10 incorporates the effects of orbital relaxation in an approximate way.

The Fukui functions may be expressed in terms of the Kohn–Sham[7] orbitals, $\{\phi_i(\vec{r})\}$:[4,8]

$$f^-(\vec{r};\lambda) \equiv \left(\frac{\partial \rho(\vec{r};\lambda)}{\partial N'}\right)^-_{v_0(\vec{r})}\bigg|_{N'=N} = |\phi_N(\vec{r};\lambda)|^2 +$$
$$\sum_{i=1}^{N-1}\left(\frac{\partial |\phi_i(\vec{r};\lambda)|^2}{\partial N'}\right)^-_{v_0(\vec{r})}\bigg|_{N'=N}$$

$$f^+(\vec{r};\lambda) \equiv \left(\frac{\partial \rho(\vec{r};\lambda)}{\partial N'}\right)^+_{v_0(\vec{r})}\bigg|_{N'=N} = |\phi_{N+1}(\vec{r};\lambda)|^2 +$$
$$\sum_{i=1}^{N}\left(\frac{\partial |\phi_i(\vec{r};\lambda)|^2}{\partial N'}\right)^+_{v_0(\vec{r})}\bigg|_{N'=N} \quad (12)$$

where the superscript $-$ and $+$ indicate that the derivatives are taken from below and above, respectively. Within the context of eq 12, Kohn–Sham orbitals which are doubly occupied occur twice in the summation. We recast eq 12 by introducing the highest occupied Kohn–Sham molecular orbital (HOMO; $\phi_N(\vec{r})$), the lowest unoccupied Kohn–Sham molecular orbital (LUMO; $\phi_{N+1}(\vec{r})$), and the Kohn–Sham orbital densities ($\rho_i(\vec{r};\lambda) \equiv |\phi_i(\vec{r};\lambda)|^2 = (\phi_i(\vec{r};\lambda))^*\phi_i(\vec{r};\lambda)$). These substitutions yield

$$f^-(\vec{r};\lambda) = \rho_{HOMO}(\vec{r};\lambda) + \sum_{i=1}^{HOMO-1}\left(\frac{\partial \rho_i(\vec{r};\lambda)}{\partial N'}\right)^-_{v_0(\vec{r})}\bigg|_{N'=N}$$

$$f^+(\vec{r};\lambda) = \rho_{LUMO}(\vec{r};\lambda) + \sum_{i=1}^{HOMO}\left(\frac{\partial \rho_i(\vec{r};\lambda)}{\partial N'}\right)^+_{v_0(\vec{r})}\bigg|_{N'=N} \quad (13)$$

Equation 13 reveals that that the Fukui functions are identical to the densities of the frontier molecular orbitals when orbital relaxation effects are neglected. Inserting eq 13, eq 10 becomes

$$\rho_{es}(\vec{r};\lambda) \approx \rho_{gs}(\vec{r};\lambda) + \rho_{LUMO}(\vec{r};\lambda) - \rho_{HOMO}(\vec{r};\lambda) +$$
$$\sum_{i=1}^{HOMO-1}\left\{\left(\frac{\partial \rho_i(\vec{r};\lambda)}{\partial N'}\right)^+_{v_0(\vec{r})}\bigg|_{N'=N} - \left(\frac{\partial \rho_i(\vec{r};\lambda)}{\partial N'}\right)^-_{v_0(\vec{r})}\bigg|_{N'=N}\right\} +$$
$$\left(\frac{\partial \rho_{HOMO}(\vec{r};\lambda)}{\partial N'}\right)^+_{v_0(\vec{r})}\bigg|_{N'=N} \quad (14)$$

Comparing eq 14 and eq 11, we see that eq 10 is accurate whenever

$$\left[\sum_{i=1}^{HOMO-1}\left\{\left(\frac{\partial \rho_i(\vec{r};\lambda)}{\partial N'}\right)^+_{v_0(\vec{r})}\bigg|_{N'=N} - \left(\frac{\partial \rho_i(\vec{r};\lambda)}{\partial N'}\right)^-_{v_0(\vec{r})}\bigg|_{N'=N}\right\}\right] +$$
$$\left[\left(\frac{\partial \rho_{HOMO}(\vec{r};\lambda)}{\partial N'}\right)^+_{v_0(\vec{r})}\bigg|_{N'=N}\right] \quad (15)$$

accurately approximates the effect of orbital relaxation upon the density of the excited state.

Bond Length Rule of Grochala, Albrecht, and Hoffmann

J. Phys. Chem. A, Vol. 104, No. 11, 2000 **2213**

Consider that $(-\sum_{i=1}^{HOMO-1}(\partial\rho_i(\vec{r};\lambda)/\partial N)_{v_0(\vec{r})}^{-})$ represents how the core orbital densities change when one removes an electron from the HOMO orbital (as when forming the cation from the ground state). Likewise, $(\sum_{i=1}^{HOMO-1}(\partial\rho_i(\vec{r};\lambda)/\partial N)_{v_0(\vec{r})}^{+})$ represents how the core orbital densities change when one adds an electron to the LUMO orbital (as when forming the anion from the ground state). Since the excitation process consists of removing an electron from the HOMO orbital and then adding it to the LUMO orbital, the first term in eq 15 approximately models the relaxation of the core orbitals. We expect the main remaining error in the equation to be due to the fact that while $(-\sum_{i=1}^{HOMO-1}(\partial\rho_i(\vec{r};\lambda)/\partial N)_{v_0(\vec{r})}^{-})$ tends to shift orbital density to where $\rho_{HOMO}(\vec{r};\lambda)$ is large (since these regions of the molecule now have a deficit of electron density, this is energetically favorable), it does not avoid shifting orbital density to regions where $\rho_{LUMO}(\vec{r};\lambda)$ is large (Coulombic repulsions with the LUMO electron render this energetically unfavorable). Likewise, while $(\sum_{i=1}^{HOMO-1}(\partial\rho_i(\vec{r};\lambda)/\partial N)_{v_0(\vec{r})}^{+})$ successfully reduces the electron density in regions where $\rho_{LUMO}(\vec{r};\lambda)$ is large, it does not concentrate electron density in regions where $\rho_{HOMO}(\vec{r};\lambda)$ is large. Accordingly, we expect that the first term in eq 15 needs to be corrected with a term which puts density in regions where $\rho_{HOMO}(\vec{r};\lambda)$ is large but $\rho_{LUMO}(\vec{r};\lambda)$ is small.

The final term in eq 15 represents how the HOMO orbital density changes when an electron is added to the LUMO orbital to form the anion. When forming the anion, the HOMO orbital density decreases where $\rho_{LUMO}(\vec{r};\lambda)$ is large (reducing the repulsive interactions between the HOMO and LUMO electrons) and increases where $\rho_{HOMO}(\vec{r};\lambda)$ is large but $\rho_{LUMO}(\vec{r};\lambda)$ is small (these are energetically favorable positions for the HOMO electron). Therefore, the last term in eq 15 corrects the deficiencies of the first term in a qualitatively correct way.

Since eq 10 consists of the frozen orbital approximation (eq 11) with approximate corrections for effects of the relaxation of the core orbitals (eq 15), we expect that eq 10 is an accurate approximation to the excited state density for most (if not all) values of λ. Since eq 10 is accurate, case (iii) holds and accordingly eq 7 is true.

This argument is somewhat heuristic in nature. An alternative derivation is presented in the appendix, which clarifies the assertion that eq 15 approximately accounts for the effects of orbital relaxation.

B. Relationships Among the Potential Energy Surfaces. Equation 7 relates the densities of the cation, anion, ground state, and the first excited state to one another and hence implies relationships between the potential energy surfaces of these systems. Specifically, we have the following corollary to proposition 1.

Corollary 1. For a system with a nondegenerate ground state,

$$\frac{\partial^n E_{gs}(\lambda)}{\partial\lambda^n} + \frac{\partial^n E_{es}(\lambda)}{\partial\lambda^n} - \frac{\partial^n E_+(\lambda)}{\partial\lambda^n} - \frac{\partial^n E_-(\lambda)}{\partial\lambda^n} \approx 0; \quad n \geq 1 \quad (16)$$

for all λ.

Proof. Use eq 7 and the Hellmann−Feynman theorem (eq 2) to establish the result for $n = 1$. The result for $n > 1$ is obtained by successive differentiation of the $n = 1$ case.

We note that corollary 1 follows directly from Grochala, Albrecht, and Hoffmann's eq 5:[1]

$$E_+(\lambda) + E_-(\lambda) - E_{gs}(\lambda) - E_{es}(\lambda) \approx 0 \quad (17)$$

Since eq 17 implies eq 16 while the converse is not true, eq 16 is accurate under less restrictive conditions than eq 17. This is reflected by the need to invoke the frozen orbital approximation for the cation, anion, and excited state in the proof of eq 17.[9] On the other hand, we were able to derive eq 16 from a theory which is exact for the ground state, cation, and anion and which approximately includes orbital relaxation effects in the approximation of the first excited state's density.

C. Relationships Among λ_+, λ_-, λ_{gs}, and λ_{es}. Assume that the potential energy surfaces for the cation, anion, ground state, and first excited state are well approximated by quadratic functions. Introducing the generalized force constants

$$k_x = \left(\frac{\partial^2 E_x}{\partial\lambda^2}\right)\Bigg|_{\lambda=\lambda_{y\neq x}}; \quad x, y = +, -, gs, es \quad (18)$$

we have

$$E_+(\lambda) \approx E_+(\lambda_+) + \frac{1}{2}k_+(\lambda - \lambda_+)^2$$

$$E_-(\lambda) \approx E_-(\lambda_-) + \frac{1}{2}k_-(\lambda - \lambda_-)^2$$

$$E_{gs}(\lambda) \approx E_{gs}(\lambda_{gs}) + \frac{1}{2}k_{gs}(\lambda - \lambda_{gs})^2 \quad (19)$$

$$E_{es}(\lambda) \approx E_{es}(\lambda_{es}) + \frac{1}{2}k_{es}(\lambda - \lambda_{es})^2$$

where λ_+, λ_-, λ_{gs}, and λ_{es} are the equilibrium values of λ for the cation, anion, ground state, and first excited state, respectively.

Applying corollary 1 to eq 19, we find

$$0 \approx k_+(\lambda - \lambda_+) + k_-(\lambda - \lambda_-) - k_{gs}(\lambda - \lambda_{gs}) - k_{es}(\lambda - \lambda_{es}), \text{ if } n = 1 \quad (21)$$

$$0 \approx k_+ + k_- - k_{gs} - k_{es}, \text{ if } n = 2 \quad (22)$$

Equation 22 allows us to simplify eq 21:

$$0 \approx k_{gs}\lambda_{gs} + k_{es}\lambda_{es} - k_+\lambda_+ - k_-\lambda_- \quad (23)$$

Grochala, Albrecht, and Hoffmann have also derived eqs 22 and 23 (eqs 10 and 11 of ref 1).[1] Section II asserts a relationship between the potential energy surfaces of the cation, anion, ground state, and first excited-state implies a relationship between the equilibrium bond lengths of the systems. Indeed, manipulating eqs 22 and 23 gives

$$0 \approx k_{gs}(\lambda_{gs} + \lambda_{es} - \lambda_+ - \lambda_-) + (k_{es} - k_{gs})\lambda_{es} - (k_+ - k_{gs})\lambda_+ - (k_- - k_{gs})\lambda_-$$

$$\lambda_{gs} + \lambda_{es} - \lambda_+ - \lambda_- \approx$$
$$\frac{(k_{gs} - k_{es})\lambda_{es} - (k_{gs} - k_+)\lambda_+ - (k_{gs} - k_-)\lambda_-}{k_{gs}} \quad (24)$$

Rearranging eq 24,

$$\lambda_{gs} + \lambda_{es} - \lambda_+ - \lambda_- \approx$$
$$\frac{\{((k_{gs} - k_{es}) - (k_{gs} - k_+) - (k_{gs} - k_-))\lambda_{es}\}}{k_{gs}} -$$
$$\frac{\{(k_{gs} - k_+)(\lambda_+ - \lambda_{es}) - (k_{gs} - k_-)(\lambda_- - \lambda_{es})\}}{k_{gs}} \quad (25)$$

and substituting eq 22 into the first term of eq 25 yields our

main result,

$$\lambda_{gs} + \lambda_{es} - \lambda_+ - \lambda_- \approx$$
$$\frac{\{(k_+ - k_{gs})(\lambda_+ - \lambda_{es})\} + \{(k_- - k_{gs})(\lambda_- - \lambda_{es})\}}{k_{gs}} \quad (26)$$

Since eq 26 is a key result, a summary of its derivation is warranted. From corollary 1 (eq 16) and the harmonic potential energy surface approximation (eq 19) we derive the approximate relations (21) and (22). Through manipulation of eqs 21 and 22 we obtain eq 26. Hence, eq 26 is a good approximation whenever eqs 21 and 22 are accurate approximations. In turn, eqs 21 and 22 are accurate whenever eq 19 is accurate and the approximations involved in derivation of corollary 1 (notably, eq 11) are valid.

D. Analyzing the Grochala−Albrecht−Hoffmann Rule. Equation 26 reveals the following.

Theorem 1. Within the present treatment, the Grochala−Albrecht−Hoffmann rule holds if and only if

$$0 \approx \frac{\{(k_+ - k_{gs})(\lambda_+ - \lambda_{es})\} + \{(k_- - k_{gs})(\lambda_- - \lambda_{es})\}}{k_{gs}} \quad (27)$$

There are two cases where eq 27 is indeed approximately zero:

$$(1) \ \{(k_+ - k_{gs})(\lambda_+ - \lambda_{es})\} = -\{(k_- - k_{gs})(\lambda_- - \lambda_{es})\} \neq 0 \quad (28)$$

$$(2) \ \{(k_+ - k_{gs})(\lambda_+ - \lambda_{es})\} = \{(k_- - k_{gs})(\lambda_- - \lambda_{es})\} = 0 \quad (29)$$

The case where $k_{gs} = \infty$ is nonphysical and will not be considered. Moreover, whenever the internuclear coordinate, λ, corresponds to a bond length, case 1 never occurs. To show this, we start with the following.

Theorem 2. The only orderings of the generalized force constants which are consistent with eq 22 are

$$(a) \ k_+ \geq k_{gs} \geq k_{es} \geq k_-$$
$$(b) \ k_+ \geq k_{es} \geq k_{gs} \geq k_-$$
$$(c) \ k_- \geq k_{gs} \geq k_{es} \geq k_+$$
$$(d) \ k_- \geq k_{es} \geq k_{gs} \geq k_+ \quad (30)$$
$$(e) \ k_{gs} \geq k_+ \geq k_- \geq k_{es}$$
$$(f) \ k_{gs} \geq k_- \geq k_+ \geq k_{es}$$
$$(g) \ k_{es} \geq k_+ \geq k_- \geq k_{gs}$$
$$(h) \ k_{es} \geq k_- \geq k_+ \geq k_{gs}$$

Proof. We show that all other orderings are inconsistent with eq 22. Because eq 22 treats the ground state and the excited state on equal footing and the cation and the anion on equal footing, we need only consider the possibilities

$$(a) \ k_{gs} > k_{es} > k_+ > k_-$$
$$(b) \ k_{gs} > k_+ > k_{es} > k_- \quad (31)$$

All other forbidden orderings may be treated by modifying the following proof.

We rewrite eq 22 as

$$(k_{gs} - k_+) \approx (k_- - k_{es}) \quad (32)$$

Both eqs 31a and 31b imply that

$$(k_{gs} - k_+) > 0 > (k_- - k_{es}) \quad (33)$$

which contradicts eq 32.

Up until this point, our theoretical treatment has been accurate for any sort of internuclear coordinate λ, including coordinates corresponding to internal rotation, angle bending, bond lengths, and normal modes of the molecule. The following proposition and its associated corollary narrows our treatment to the case where λ represents a bond length.

Proposition 2. Suppose that the internuclear coordinate λ corresponds to a bond length. We may assume that as the force constant for a system increases the bond length decreases.

Justification. Inasmuch as stronger bonds have shorter bond lengths and larger force constants, this proposition is intuitively reasonable. Moreover, this observation agrees with empirical formulas that relate the bond length to the force constant.

Corollary 2. Let k_i; $i = 1, 2, 3, 4$ denote the bond stretching force constant for four different electronic states of the same molecule. If $k_1 \geq k_2 \geq k_3 \geq k_4$, then $\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \lambda_4$.

Theorem 1 and corollary 2 allow us to prove that the first and second terms in eq 27 always have the same sign. As a representative example, consider the case where $k_+ \geq k_{gs} \geq k_{es} \geq k_-$ (eq 30a). Corollary 2 implies that $\lambda_+ \leq \lambda_{gs} \leq \lambda_{es} \leq \lambda_-$. Hence, $(k_+ - k_{gs})$ and $(\lambda_- - \lambda_{es})$ are both greater than zero, while $(k_- - k_{gs})$ and $(\lambda_+ - \lambda_{es})$ are both less than zero. Substitution into eq 27 reveals that the first and second terms are *both* negative, and hence case 1 does not occur. Performing similar arguments for eqs 30b−h in theorem 1 reveals the following.

Theorem 2. When λ corresponds to a bond length, the first and second terms in eq 27 always have the same sign. That is, case 1 (eq 28) never occurs.

Case 2 (eq 29) only occurs when one of the following is true:

$$(2a) \ k_+ - k_{gs} = k_- - k_{gs} = 0$$
$$(2b) \ \lambda_+ - \lambda_{es} = \lambda_- - \lambda_{es} = 0$$
$$(2c) \ k_+ - k_{gs} = \lambda_- - \lambda_{es} = 0$$
$$(2d) \ \lambda_+ - \lambda_{es} = k_- - k_{gs} = 0$$

We now explore each of these cases in detail. Case 2a may be simplified by using eq 22, which reveals that case 2a occurs if and only if all of the force constants are equal; that is,

$$(2a) \Leftrightarrow k_+ = k_- = k_{gs} = k_{es} \quad (34)$$

It is easily seen that when eq 34 holds, eq 23 reduces to the GAH identity.

In case 2b, $\lambda_+ = \lambda_- = \lambda_{es}$. We may thus substitute λ_{es} for λ_+ and λ_- in eq 23, obtaining

$$0 \approx k_{gs}\lambda_{gs} + \lambda_{es}(k_{es} - k_+ - k_-) \quad (35)$$

Adding and subtracting $(k_{gs}\lambda_{gs})$

$$0 \approx k_{gs}(\lambda_{gs} - \lambda_{es}) + \lambda_{es}(k_{gs} + k_{es} - k_+ - k_-)$$

and then applying eq 22, we find that

Bond Length Rule of Grochala, Albrecht, and Hoffmann

J. Phys. Chem. A, Vol. 104, No. 11, 2000 **2215**

$$\lambda_{gs} = \lambda_{es} \quad (36)$$

Therefore, case 2b corresponds to the trivial case where the GAH rule holds because

$$\lambda_+ = \lambda_- = \lambda_{es} = \lambda_{gs} \quad (37)$$

In case 2c, $k_{gs} = k_+$. Substituting for k_+ in eq 22 reveals that $k_{es} = k_-$. Substitution into eq 23 yields

$$k_{gs}(\lambda_{gs} - \lambda_+) + k_{es}(\lambda_{es} - \lambda_-) = 0 \quad (38)$$

which, when combined with the relation $\lambda_- = \lambda_{es}$, reveals that $\lambda_+ = \lambda_{gs}$. Accordingly, case 2c holds if and only if

$$\begin{aligned} k_+ = k_{gs} \quad k_- = k_{es} \\ \lambda_+ = \lambda_{gs} \quad \lambda_- = \lambda_{es} \end{aligned} \quad (39)$$

The analysis for case 2d is similar to that for case 2c. Case 2d holds if and only if

$$\begin{aligned} k_+ = k_{es} \quad k_- = k_{gs} \\ \lambda_+ = \lambda_{es} \quad \lambda_- = \lambda_{gs} \end{aligned} \quad (40)$$

We expect that the GAH rule will be approximately valid when any of the cases 2a–2d are approximately valid. Unfortunately, it is not realistic to assume that any of these four cases approximately hold. Case 2b renders the GAH rule trivial. As was noted in ref 1, "approximate equality of force constants for the four species under discussion is evidently not true" because the bond orders of these species differ substantially. Accordingly, we do not expect case 2a to be even approximately valid. Inspection of the tabulated data of Grochala, Albrecht, and Hoffmann reveals that it is rare for either the cation or the anion to have a bond length very close to that of either the ground or the excited state, hence cases 2c and 2d are similarly unrealistic.

If none of the conditions 2a–2d hold, why does the GAH rule work as well as it does? Reference 1 tabulates the error in the GAH rule relative to the spread in the equilibrium bond lengths:

$$\epsilon \equiv \frac{\lambda_{gs} + \lambda_{es} - \lambda_+ - \lambda_-}{\max\{\lambda_{gs};\lambda_{es};\lambda_+;\lambda_-\} - \min\{\lambda_{gs};\lambda_{es};\lambda_+;\lambda_-\}} \quad (41)$$

Substituting eq 27, the error expression becomes

$$\epsilon \approx \frac{\{(k_+ - k_{gs})(\lambda_+ - \lambda_{es})\} + \{(k_- - k_{gs})(\lambda_- - \lambda_{es})\}}{k_{gs} \text{ spread}(\lambda)} \quad (42)$$

To try to explain why this error is small, we compute the magnitude of eq 42. It turns out that eq 22, theorem 2, and corollary 2 significantly restrict the force constants and bond lengths, leading to the worst case error

$$\epsilon_{worst} \approx \frac{\text{spread}(k)\,\text{spread}(\lambda)}{k_{gs}\text{spread}(\lambda)} = \frac{\text{spread}(k)}{k_{gs}} \quad (43)$$

For instance, this case occurs when eq 30a holds in the form $k_+ \approx k_{gs} > k_{es} \approx k_-$ and $\lambda_- > \lambda_{es} \approx \lambda_{gs} \approx \lambda_+$. While this ordering of bond lengths is in technical agreement with corollary 2, it is not in the spirit of proposition 2, which would indicate

that, since $k_{es} \approx k_-$, then $\lambda_{es} \approx \lambda_-$. This then reduces the situation to a case (case 2c; eq 39) in which the GAH error is zero.

Inspection of the tables of results in ref 1 reveals that the four systems under consideration generally have bond lengths which are more or less equally spaced within the interval $(\min\{\lambda_{gs};\lambda_{es};\lambda_+;\lambda_-\}, \max\{\lambda_{gs};\lambda_{es};\lambda_+;\lambda_-\})$. In keeping with this observation, we now estimate the error for a more typical case. Choosing our force constants to be ordered by eq 30a, we now distribute the force constants and bond lengths evenly throughout their range, i.e.,[11]

$$k_- + \text{spread}(k) = k_{es} + \frac{2}{3}\text{spread}(k) = k_{gs} + \frac{1}{3}\text{spread}(k) = k_+$$

$$\lambda_+ + \text{spread}(\lambda) = \lambda_{gs} + \frac{2}{3}\text{spread}(\lambda) =$$

$$\lambda_{es} + \frac{1}{3}\text{spread}(\lambda) = \lambda_- \quad (44)$$

In this case, which we term the "typical" case, eq 42 becomes

$$\epsilon_{typical} \approx -\frac{4}{9}\frac{\text{spread}(k)}{k_{gs}} \quad (45)$$

Accordingly, the typical error in the GAH rule is only half as bad as the worst case scenario. Since the variation in the force constants among the systems under consideration is frequently of the order of 10^2 N/m while the magnitude of k_{gs} is frequently about 10^3 N/m, we predict that the typical error in the GAH rule is about 5%, perhaps 2% in the best case scenario (where the force constants vary negligibly among these systems) to 25% in the worst scenario (in which the force constants vary substantially). This prediction is in general agreement with the computations of Grochala, Albrecht, and Hoffmann.

E. Degenerate Ground States. We now sketch a generalization of the GAH rule to degenerate ground states. Analogous to proposition 1, the first step is to link the densities of the cation, anion, and ground state.

Proposition 3. Suppose that M is a molecule which has a degenerate ground state and suppose, furthermore, that the order of this degeneracy is unchanged as we change λ. Then, the electron densities of the cation ($\rho_+(\vec{r};\lambda)$), anion ($\rho_-(\vec{r};\lambda)$), and ground state ($\rho_{gs}(\vec{r};\lambda)$) approximately satisfy

$$\int [2\rho_{gs}(\vec{r};\lambda) - \rho_+(\vec{r};\lambda) - \rho_-(\vec{r};\lambda)]\left(\frac{\partial v_0(\vec{r};\lambda)}{\partial \lambda}\right) d\vec{r} \approx 0 \quad (46)$$

for all λ. In eq 46 $\rho_{gs}(\vec{r})$ is the density of any one of the degenerate ground states.

Justification. A degenerate ground state occurs when the HOMO orbital(s) is not filled. In the general case, there are $N_{valence}$ electrons in g spatial orbitals ($2g$ spin–orbitals) and $0 < N_{valence} < 2g$. In this case, there may be many different ground state densities with the same energy but different symmetry. However, if the ground state's degeneracy is unchanged by a change in λ, then all of the different possible ground states, characterized by a range of different densities, have the same energy for all λ. Practically, this restriction confines our treatment to coordinates (as uniform stretching of all bonds in the molecule) which preserve the symmetry of the molecule. Stating the condition symbolically, we require that for all λ

$$E_{gs}^i(\lambda) = E_{gs}^j(\lambda) \quad (47)$$

where i and j index the degenerate ground states. Differentiating

eq 47 and applying the Hellmann/Feynman theorem (eq 2), we find that, for all λ,

$$\frac{\partial E_{gs}^i(\lambda)}{\partial \lambda} = \frac{\partial E_{gs}^j(\lambda)}{\partial \lambda}$$

$$\frac{\partial V_{nn}}{\partial \lambda} + \int \rho_{gs}^i(\vec{r}; \lambda)\left(\frac{\partial v_0(\vec{r}; \lambda)}{\partial \lambda}\right) d\vec{r} =$$
$$\frac{\partial V_{nn}}{\partial \lambda} + \int \rho_{gs}^j(\vec{r}; \lambda)\left(\frac{\partial v_0(\vec{r}; \lambda)}{\partial \lambda}\right) d\vec{r} \quad (48)$$

$$\int \rho_{gs}^i(\vec{r}; \lambda)\left(\frac{\partial v_0(\vec{r}; \lambda)}{\partial \lambda}\right) d\vec{r} = \int \rho_{gs}^j(\vec{r}; \lambda)\left(\frac{\partial v_0(\vec{r}; \lambda)}{\partial \lambda}\right) d\vec{r}$$

where we have denoted the densities of the degenerate ground states as $\rho_i(\vec{r})$ and $\rho_j(\vec{r})$. This implies that we may choose any of the degenerate ground state densities in eq 46.

Unless $g = 1$, Hund's maximum multiplicity rule indicates the orbital from which one removes an electron to form the cation ($\phi_+(\vec{r})$) and the orbital into which one puts an electron to form the anion ($\phi_-(\vec{r})$) will not be identical (although the two orbitals will have identical energies).[12] Moreover, the orbitals $\phi_+(\vec{r})$ and $\phi_-(\vec{r})$ generally depend on which of the degenerate ground states is under consideration. Therefore, within the frozen orbital approximation, the densities of the cation and anion are given by

$$\rho_+^i(\vec{r}) = \rho_{gs}^i(\vec{r}) - |\phi_+^i(\vec{r})|^2$$
$$\rho_-^i(\vec{r}) = \rho_{gs}^i(\vec{r}) + |\phi_-^i(\vec{r})|^{2'} \quad (49)$$

where $\rho_{gs}^i(\vec{r})$ is the electron density of the ground state under consideration.

Since two different degenerate ground state densities are

$$\rho_{gs}^+(\vec{r}) = \rho_{core}(\vec{r}) + |\phi_+^i(\vec{r})|^2$$
$$\rho_{gs}^-(\vec{r}) = \rho_{core}(\vec{r}) + |\phi_-^i(\vec{r})|^{2'} \quad (50)$$

eq 48 implies that

$$\int |\phi_+^i(\vec{r})|^2\left(\frac{\partial v_0(\vec{r}; \lambda)}{\partial \lambda}\right) d\vec{r} = \int |\phi_-^i(\vec{r})|^2\left(\frac{\partial v_0(\vec{r}; \lambda)}{\partial \lambda}\right) d\vec{r} \quad (51)$$

Adding the two eqs 49 and integrating against $(\partial v_0(\vec{r}; \lambda)/\partial \lambda)$ gives

$$\int (\rho_+^i(\vec{r}) + \rho_-^i(\vec{r}))\left(\frac{\partial v_0(\vec{r}; \lambda)}{\partial \lambda}\right) d\vec{r} = \int (2\rho_{gs}^i(\vec{r}) -$$
$$|\phi_+^i(\vec{r})|^2 + |\phi_-^i(\vec{r})|^2)\left(\frac{\partial v_0(\vec{r}; \lambda)}{\partial \lambda}\right) d\vec{r} \quad (52)$$

Because of eqs 48 and 51, eq 52 is equivalent to the desired result (eq 46).

Analogous to corollary 1 in section III.B is the following.

Corollary 3. For a system with a degenerate ground state described by proposition 3,

$$2\frac{\partial^n E_{gs}(\lambda)}{\partial \lambda^n} - \frac{\partial^n E_+(\lambda)}{\partial \lambda^n} - \frac{\partial^n E_-(\lambda)}{\partial \lambda^n} \approx 0; \quad n \geq 1 \quad (52)$$

for all λ.

Assuming that the potential energy surfaces for the cation, anion, and ground state are harmonic (eq 19) and performing analysis similar to that of section III.C reveals that

$$0 \approx k_+ + k_- - 2k_{gs} \quad (53)$$

$$0 \approx 2k_{gs}\lambda_{gs} - k_+\lambda_+ - k_-\lambda_- \quad (54)$$

Equations 53 and 54 are analogous to eqs 22 and 23. Manipulations similar to eqs 24−26 yield

$$\frac{(k_+ - k_{gs})(\lambda_+ - \lambda_-)}{k_{gs}} = \lambda_+ + \lambda_- - 2\lambda_{gs} \quad (55)$$

Equation 55 is the generalization of eq 26 to degenerate ground states and leads to a GAH rule for degenerate ground states:

$$\lambda_+ + \lambda_- - 2\lambda_{gs} \approx 0 \quad (56)$$

Equation 56 is exact if and only if one of the following equations holds:

$$(1) \; \lambda_+ = \lambda_- = \lambda_{gs}$$
$$(2) \; k_+ = k_- = k_{gs}$$

As in the nondegenerate case, neither (1) nor (2) is likely to be approximately true for many systems. Analysis similar to that in section III.D shows that

$$\epsilon_{degenerate} = \frac{1}{2}\frac{\text{spread}(k)}{k_{gs}} \quad (57)$$

The results for degenerate ground states could be obtained from the results for the nondegenerate ground states by assuming that the first excited state and the ground state are degenerate. Indeed, in many cases, a degenerate ground state is caused by an intersection between the potential energy surfaces of the ground state and the first excited state. This intersection generally occurs at a point or along a curve which corresponds to a molecular structure of high symmetry. It is to the internuclear coordinate associated with this curve that the results of this section apply.

IV. Discussion

Revisiting the assumptions which were made in our derivation of the GAH rule reveals under what conditions the GAH rule will hold. The first assumption is mentioned in the introduction of the Hellmann−Feynman theorem (eq 2); the derivation considers the way the energy of a molecular system changes when the internuclear coordinate λ changes *but all other internuclear coordinates remain fixed.* If any of the remaining $3K - 7$ coordinates (K is the number of nuclear centers in the molecule)[13] differs substantially among the cation, anion, ground state, and excited state, then holding all but one coordinate fixed as we change the state of the molecule is no longer a good approximation. Even when this approximation fails, the GAH rule is useful for computational work since it predicts the results of performing a geometry optimization with respect to the coordinate λ while keeping the other $3K - 7$ coordinates fixed.[13]

We now turn to propositions 1 and 3, which provide approximate relationships between the densities of the systems under consideration (eqs 7 and 46). Proposition 3 depends on the validity of the frozen orbital approximation, which likely provides a poor approximation to the density of the cation and the anion. Orbital mixing and Jahn−Teller effects frequently accompany the formation of the cation and the anion when the ground state is degenerate, and hence the validity of the frozen

Bond Length Rule of Grochala, Albrecht, and Hoffmann

J. Phys. Chem. A, Vol. 104, No. 11, 2000 **2217**

orbital assumption, upon which the justification of proposition 3 rests, is questionable. Since Jahn–Teller distortions do not occur when there is only one degenerate orbital, we expect our treatment to be most accurate when the HOMO orbital is nondegenerate but singly occupied so that the ground state is a doublet but both the anion and the cation are singlet states.

In contrast to our justification for proposition 3, we use the Fukui function to justify proposition 1; this allows the cation, anion, and ground state to be treated exactly and allows us to approximate the effects of orbital relaxation upon the excited state density. When one has a nondegenerate (necessarily singlet) ground state, then one can form the singlet excited state (by promoting the electron from the HOMO spin-up orbital to the LUMO spin-up orbital) or the triplet excited state (by promoting an electron from the HOMO spin-up orbital to the LUMO spin-down orbital). The orbital relaxation effects included in our treatment of the excited state density (eq 15) approximate the orbital relaxation of the core orbitals and neglect the relaxation of the LUMO orbital. Accordingly, our treatment will be most accurate when the LUMO orbital relaxes little. If the excited state is a triplet, the unpaired electron in the HOMO orbital and the unpaired electron in the LUMO orbital have the same spin, and hence the Pauli exclusion principle prevents them from coming close together, effectively preventing the electron in the LUMO orbital from "relaxing" into the region of space vacated by the removed HOMO electron. Thus, the ground state's LUMO changes less when we form the triplet excited state than when we form the singlet excited state. *Therefore, we expect that proposition 1, and hence the GAH rule, will be most accurate for triplet excited states.*

Since orbital relaxation may be modeled by mixing an orbital with other orbitals which are near to it in energy, we expect that the relaxation of the LUMO orbital will be smaller for systems with large band gaps. Large band gaps are associated with high values of the chemical hardness,[14] and hence *we expect that the proposition 1 will be most accurate for systems with high chemical hardness.* Alternatively, a high value of the aromaticity, a concept which correlates strongly with both the band gap and the chemical hardness,[15] typifies cases where proposition 1 will be accurate.

The next approximation is eq 19, which states that, at least in the vicinity of the equilibrium bond lengths for the various systems, the potential energy surfaces for the systems under consideration are harmonic. Looking back at section III.C, it becomes apparent that we could omit the harmonic assumption (and indeed, also the assumptions embodied in propositions 1 and 3), by stating that there exist constants, k_+, k_-, k_{gs}, and k_{es} such that eqs 22 and 23 hold. From this perspective, the purpose of section III is to uncover information about the values of k_+, k_-, k_{gs}, and k_{es} so that the error in the GAH rule can be estimated. We conclude that the constants which satisfy eqs 22 and 23 will resemble the physical force constants for the systems whenever each potential energy surface ($E_+(\lambda)$, $E_-(\lambda)$, $E_{gs}(\lambda)$, and $E_{es}(\lambda)$) has a small third derivative with respect to λ (hence a small anharmonicity) when evaluated at λ_+, λ_-, λ_{gs}, and λ_{es}.

Even when the anharmonicity is not small, there exist constants, k_i, such that eqs 22 and 23 hold. But now we cannot identify these constants with the force constants for the system and hence cannot estimate the error in terms of these constants. In this case, we can assume a potential energy surface of the form

$$E_i(\lambda) = \frac{1}{2} k_i (\lambda - \lambda_i)^2 + \frac{1}{3!} \alpha_i (\lambda - \lambda_i)^3 \quad (58)$$

and perform an analysis using the $n = 3$ case of corollary 1. We omit this analysis, as it is less illuminating that the preceding, but mention that the $n = 2$ case of corollary 1 indicates that

$$k_+ + k_- - k_{gs} - k_{es} = \alpha_+(\lambda - \lambda_+) + \alpha_-(\lambda - \lambda_-) - \alpha_{gs}(\lambda - \lambda_{gs}) - \alpha_{es}(\lambda - \lambda_{es}) \quad (59)$$

Equation 59 provides a criterion for how anharmonic the potential energy surface may be before eqs 22 and the subsequent analysis become inaccurate. Since the potential energy surface of any coordinate is harmonic in the immediate vicinity of the minimum, eq 59 gives an indication of how large spread(λ) may be before the harmonic approximation becomes unacceptably inaccurate. *We conclude that eqs 22 and 23 are most accurate when spread(λ) is small.*

At a qualitative level, the restriction of the potential energy to the form of eq 19 restricts the type of motion the coordinate λ may represent to one well-described by eq 19 for a moderate range of λ. This reduces the effectiveness of the GAH rule for coordinates, λ, which correspond to angle bendings and torsions. On the other hand, *harmonic potentials are good approximations for bond stretching modes.*

The only remaining question is, for what sorts of systems is the error in the GAH rule (eqs 42, 43, 45, and 57) expected to be small? Our formulas reveal that when spread(k) $\ll k_{gs}$ (equivalently, $0 \approx$ spread(k)/$k_{gs} \ll 1$), the GAH rules will be highly accurate. This is expected to be true for multiple bonds since removing electrons from the HOMO orbital and/or adding electrons to the LUMO orbital generally changes the bond order by at most 1. Therefore, when the total bond order for the bond described by the coordinate λ is greater than 1, we expect the range of force constants to be moderate. This is especially true when the HOMO and LUMO orbitals are both π orbitals (of either the bonding or the antibonding variety); because most of the bond strength (and hence most of the magnitude of k) is due to the unperturbed "skeleton" of σ bonds, changing the occupation of π-type orbitals changes k by a small amount (compared to the change in k which would result from changing the occupation of σ-type orbitals). *Accordingly, we expect that the GAH rule will be most accurate for molecules with multiple bonds. We expect systems where the HOMO and/or LUMO orbitals are π-type orbitals to be especially well modeled.* This conclusion agrees with the considerations of Grochala, Albrecht, and Hoffmann.[1]

The GAH rule provides insight into molecular geometry, but it also may be used for predictions. For the ground state, cation, and first triplet state of many molecules, accurate geometries may be determined by Hartree–Fock calculations[16] with small to moderate-sized basis sets. By contrast, calculating accurate geometries for anions is substantially more difficult, often requiring correlated computational methods and large basis sets augmented by diffuse functions. When only a semiquantitative geometry for the anion is needed, the GAH rule allows one to simply predict the anion's geometry. Such semiquantitative geometries might be useful as initial guesses for geometry optimizations.

V. Additional Results

Part II asserts that a relationship between the densities of two systems provides a link between their potential energy surfaces. Substitution of the relationship between the ground-state density and densities of the cation, anion, and first excited state (eqs 8–10) into the first (eq 2) and second-order Hellmann/Feynman

2218 *J. Phys. Chem. A, Vol. 104, No. 11, 2000*

Ayers and Parr

theorems (eq 6) gives these relationships explicitly. Namely,

$$E_+'(\lambda) = E_{gs}'(\lambda) - \int f^-(\vec{r};\lambda)\, v_0'(\vec{r};\lambda)\, d\vec{r}$$

$$E_+''(\lambda) = E_{gs}''(\lambda) - \int f^{-\prime}(\vec{r};\lambda)\, v_0'(\vec{r};\lambda)\, d\vec{r} - $$
$$\int f^-(\vec{r};\lambda)\, v_0''(\vec{r};\lambda)\, d\vec{r} \quad (60)$$

$$E_-'(\lambda) = E_{gs}'(\lambda) + \int f^+(\vec{r};\lambda)\, v_0'(\vec{r};\lambda)\, d\vec{r}$$

$$E_-''(\lambda) = E_{gs}''(\lambda) + \int f^{+\prime}(\vec{r};\lambda)\, v_0'(\vec{r};\lambda)\, d\vec{r} + $$
$$\int f^+(\vec{r};\lambda)\, v_0''(\vec{r};\lambda)\, d\vec{r} \quad (61)$$

$$E_{es}'(\lambda) \approx E_{gs}'(\lambda) + \int (f^+(\vec{r};\lambda) - f^-(\vec{r};\lambda))v_0'(\vec{r};\lambda)\, d\vec{r}$$

$$E_{es}''(\lambda) \approx E_{gs}''(\lambda) + \int (f^{+\prime}(\vec{r};\lambda) - f^{-\prime}(\vec{r};\lambda))v_0'(\vec{r};\lambda)\, d\vec{r} + $$
$$\int (f^+(\vec{r}) - f^-(\vec{r}))v_0''(\vec{r};\lambda)\, d\vec{r} \quad (62)$$

Equations 60 and 61 are exact since eqs 8 and 9 give the exact density for the cation and the anion, respectively. Moreover, since if we know the derivative of the function we can determine the value of that function to within an additive constant, eqs 60 and 61 allow one to determine the potential energy surfaces of the cation and the anion to within an additive constant (which amounts to a choice of the zero of energy). Unfortunately, even this approach is not yet practical; evaluating eqs 60 and 61 requires knowledge of the Fukui functions and efficient methods for the accurate calculation of the Fukui functions remain elusive. On an optimistic note, computing the Fukui function from a Kohn−Sham DFT calculation is a subject of active research, and progress is being made.[17]

We should also comment on the use of density functional theory concepts in this derivation. Inspection of the derivation reveals that one can formally derive the same results using frontier molecular orbital theory. For instance, one could assume that the wave function of the cation was obtained by removing the HOMO orbital from the ground state wave function, that the wave function of the anion was obtained by adding the LUMO orbital to the ground state wave function, and that the wave function of the excited state was obtained by adding the LUMO orbital to the cation's wave function. Such a treatment would give the same results as the treatment in the text, but the assumptions are much more severe. By using Fukui functions to obtain the exact densities of the cation and the anion, we were able to treat these systems exactly. Even our approximation to the excited state density is much less restrictive than the molecular orbital theory approximation just mentioned. Requiring the excited-state density to be given by eq 10 is not a true "frozen orbital approximation" since, for instance, the LUMO orbital density can change significantly as long as there is a compensating change in the densities of the core orbitals. Moreover, even if eq 10 is not accurate for all $\vec{r}$, eq 7 can be still be accurate. Similar comments apply to our generalization of the GAH rule to degenerate ground states; eq 46 may hold even when eq 52 is not accurate in a pointwise sense, and eq 52 can hold whenever the effects of orbital relaxation on the core orbitals' density and the appropriate frontier orbital's density cancel.

As has also been implicitly suggested by Grochala, Albrecht, and Hoffmann, the GAH rule may be generalized to states other than the cation, anion, ground state, and first triplet excited state. In the general case we have, corresponding to eqs 7 and 46,

$$\int \Big\{ \sum_{i=1}^{\text{no. of states}} \alpha_i \rho_i(\vec{r};\lambda) + \beta \rho_\beta(\vec{r};\lambda) \Big\} \left(\frac{\partial v_0(\vec{r};\lambda)}{\partial \lambda} \right) d\vec{r} \approx 0 \quad (63)$$

for every λ, where $\rho_i(\vec{r};\lambda)$ is the density of some state of the molecule, $\{\alpha_i\}_{i=1}^{\text{no. of states}}$ are real numbers and

$$\rho_\beta(\vec{r};\lambda) \approx \frac{-1}{\beta} \sum_{i=1}^{\text{no. of states}} \alpha_i \rho_i(\vec{r};\lambda) \quad (64)$$

When $\rho_\beta(\vec{r};\lambda)$ happens to be close to the density of some state of the molecule (as is the case for the special cases of eqs 63 and 64 we consider in this paper), generalization of the derivation in the text gives

$$\sum_{i=1}^{\text{no. of states}} \alpha_i \lambda_i + \beta \lambda_\beta \approx 0 \quad (65)$$

Given a specific case, the applicability of eq 65 may be ascertained through a detailed analysis similar to that contained in sections III and IV of the present paper.

VI. Summary

If the ground state of a system is nondegenerate, the Grochala−Albrecht−Hoffmann rule, eq 1, is valid for bond lengths in the cation, anion, ground state, and first triplet excited state. On the other hand, when the ground state is degenerate, the generalized GAH rule, eq 56, is valid for changes along any internuclear coordinate λ that does not change the order of the degeneracy. We have also explained some of the restrictions on the validity of the GAH rule which were noted in the original paper. Specifically, we have observed that the GAH rule is not expected to hold for systems where the cation, anion, ground state, and excited state molecules have markedly different geometries. We have also derived some new conditions which are favorable (though perhaps not always absolutely necessary) for the accurate fulfillment of the Grochala−Albrecht−Hoffmann rule and its generalization to degenerate ground states.

(1) The geometric parameter under consideration corresponds to a bond stretching. Accordingly, the Grochala−Albrecht−Hoffmann rule is expected to be most accurate for bond lengths and the symmetric and antisymmetric stretching modes of a molecule.

(2) The geometries of the cation, anion, ground state, and first excited state are qualitatively similar except for the value of one internal coordinate. The harmonic approximation (eq 19) remains valid when the equilibrium value of this coordinate does not change dramatically among the systems under discussion.

(3) For systems with a nondegenerate ground state, it is advantageous for the band gap to be large. Equivalently, systems which are highly aromatic or have a high value of the chemical hardness are expected to follow the Grochala−Albrecht−Hoffmann rule more accurately.

(4) $0 \approx \text{spread}(k)/k_{gs} \ll 1$. (See eqs 42, 43, 45, and 57.) This condition is most readily satisfied for multiple bonds and is expected to be most accurate when the σ-bond "skeleton" of a molecule is unchanged upon formation of the cation, anion, and first triplet excited state. Accordingly, we expect that the error in the Grochala−Albrecht−Hoffmann rule will be smallest when the HOMO and/or LUMO orbitals are π, δ, ϕ, etc. type orbitals.

Bond Length Rule of Grochala, Albrecht, and Hoffmann

J. Phys. Chem. A, Vol. 104, No. 11, 2000 **2219**

Appendix

This appendix sketches an alternative approach to propositions 1 and 3. We start with the special case of proposition 3 where the HOMO orbital energy is nondegenerate but the HOMO orbital is only singly occupied in the ground state. In this case, the ground state is a doublet while both the cation and the anion are singlet states. One may then approximate the wave functions for the ground state by annihilating spin−orbitals from the anion:

$$\tilde{\Psi}^{\beta}_{\text{gs}}(x_1,x_2,...,x_N) \equiv \int \Psi_-(x_1,x_2,...,x_{N+1})\tilde{\phi}_\alpha(x_{N+1})\,\mathrm{d}x_{N+1}$$

$$\tilde{\Psi}^{\alpha}_{\text{gs}}(x_1,x_2,...,x_N) \equiv \int \Psi_-(x_1,x_2,...,x_{N+1})\tilde{\phi}_\beta(x_{N+1})\,\mathrm{d}x_{N+1}$$

(66)

where x_i denotes both the spatial and the spin coordinates associated with the ith electron, $\int \mathrm{d}x_i$ denotes integration over the spatial coordinates and summation over the spin coordinate of the ith electron, $\tilde{\phi}_{\alpha(\beta)}(x)$ is a spin−orbital with $\alpha(\beta)$ spin, Ψ_- is the *exact* wave function for the anion, and $\tilde{\Psi}^{\alpha}_{\text{gs}}$ and $\tilde{\Psi}^{\beta}_{\text{gs}}$ are approximate wave functions for the ground state which have spins α and β, respectively. The extended Koopmans' theorem[18−20] indicates that the wave function, Ψ^{β}_{gs}, obtained by minimizing the energy of $\tilde{\Psi}^{\beta}_{\text{gs}}$ with respect to all possible choices of $\tilde{\phi}_\alpha(x)$ is the *exact* wave function for the ground state.[20−22] Writing the optimal choice for $\tilde{\phi}_\alpha(x)$ as $\phi_\alpha(x)$, one finds that the density of the ground state is given by

$$\rho^{\beta}_{\text{gs}}(\vec{r}) \equiv \rho_-(\vec{r}) - |\phi_\alpha(\vec{r})|^2 \qquad (67)$$

A similar result holds for $\Psi^{\alpha}_{\text{gs}}$, and we note that, *in this special case, the spatial densities of the two degenerate ground states are identical*.

Likewise, the wave function of the cation can be approximated by annihilating either an α-spin orbital, $\tilde{\chi}_\alpha(x_N)$, from $\Psi^{\alpha}_{\text{gs}}$ or a β-spin−orbital, $\tilde{\chi}_\beta(x_N)$, from Ψ^{β}_{gs}. Using eq 66, we may then write

$$\tilde{\Psi}^{\beta}_+(x_1,x_2,...,x_{N-1}) \equiv$$
$$\int \Psi_-(x_1,x_2,...,x_{N+1})\,\phi_\alpha(x_{N+1})\,\tilde{\chi}_\beta(x_N)\,\mathrm{d}x_{N+1}$$

$$\tilde{\Psi}^{\alpha}_+(x_1,x_2,...,x_{N-1}) \equiv$$
$$\int \Psi_-(x_1,x_2,...,x_{N+1})\,\phi_\beta(x_{N+1})\,\tilde{\chi}_\alpha(x_N)\,\mathrm{d}x_{N+1} \quad (68)$$

We might anticipate that the extended Koopmans' theorem would indicate that minimizing the energy of $\tilde{\Psi}^{\beta}_+$ with respect to $\tilde{\chi}_\beta(x)$ yields the exact wave function for the cation, Ψ^{β}_+. This is not the case; minimizing the energy of $\tilde{\Psi}^{\beta}_+$ with respect to $\tilde{\chi}_\beta(x)$ yields a wave function for the cation, Φ^{β}_+, which is "spin-contaminated", that is, Φ^{β}_+ is not an eigenfunction of the squared-magnitude total spin operator, $\hat{S}^2$.[23,24] Since $\hat{S}^2$ and the Hamiltonian operator, $\hat{H}$, commute, it follows that Φ^{β}_+, is not the exact wave function for the cation. In practice, the spin contamination of Φ^{β}_+ is found to be relatively minor and the energy of Φ^{β}_+ is accurate.[23] Applying the same procedure to

$\Psi^{\alpha}_{\text{gs}}$, we obtain a *different* approximation to the wave function of the cation, Φ^{α}_+. Φ^{α}_+ is an equally accurate approximation to the exact wave function of the cation,[23] and hence we suspect that Φ^{β}_+ and Φ^{α}_+ are approximately equal. It is somewhat disconcerting that Φ^{β}_+ is not *exactly* equal to Φ^{α}_+, since this indicates that our approximation to the cation's wave function depends on whether we first optimize the α-spin orbital and then the β-spin orbital, or vice versa. Since Φ^{β}_+ and Φ^{α}_+ are close to one another anyway, we decide not to choose $\tilde{\chi}_\alpha(x)$ and $\tilde{\chi}_\beta(x)$ to minimize the cation's energy but instead choose $\tilde{\chi}_\alpha(x)$ and $\tilde{\chi}_\beta(x)$ so that[25]

$$\tilde{\Psi}^{\beta}_+(x_1,x_2,...,x_{N-1}) = \tilde{\Psi}^{\alpha}_+(x_1,x_2,...,x_{N-1}) \qquad (69)$$

One immediately observes that eq 69 is only satisfied when

$$\tilde{\chi}_\alpha(x) = \phi_\alpha(x)$$
$$\tilde{\chi}_\beta(x) = \phi_\beta(x) \qquad (70)$$

Since Φ^{β}_+ and Φ^{α}_+ are both close to the true wave function of the cation, and hence to one another, we expect that substitution of eq 70 into eq 68 will yield a satisfactory approximation to the wave function of the cation. Equation 70 indicates that eq 69 holds when the α- and β-spin electrons of the ground state are "decoupled"; that is, the best way to remove an α-spin electron from the system is unchanged by the removal of a β-spin electron from the system. Interactions between electrons of opposite spin are hence effectively ignored.

Within the approximation of eqs 68 and 70, the density is given by

$$\rho_+(\vec{r}) = \rho_-(\vec{r}) - |\phi_\alpha(\vec{r})|^2 - |\phi_\beta(\vec{r})|^2$$
$$= \rho_-(\vec{r}) - 2|\phi_\alpha(\vec{r})|^2 \qquad (71)$$

where the second equality follows because the extended Koopmans' theorem method of construction (eq 66) indicates that $\phi_\alpha(x)$ and $\phi_\beta(x)$ differ only with respect to spin, and not spatially. Equations 68 and 71 imply that $2\rho_{\text{gs}}(\vec{r};\lambda) - \rho_+(\vec{r};\lambda) - \rho_-(\vec{r};\lambda) = 0$, and hence proposition 3 (eq 46) is satisfied.

Proposition 1 may be justified in a similar way. Given the α-spin state of a doublet anion, annihilate an α-spin orbital to form an approximation to the wave function of the singlet ground state and a β-spin orbital to form an approximation to the wave function of the triplet excited state:

$$\tilde{\Psi}^{\alpha\beta}_{\text{gs}}(x_1,x_2,...,x_N) \equiv \int \Psi^{\alpha}_-(x_1,x_2,...,x_{N+1})\tilde{\phi}_\alpha(x_{N+1})\,\mathrm{d}x_{N+1}$$

$$\Psi^{\alpha\alpha}_{\text{es}}(x_1,x_2,...,x_N) \equiv \int \Psi^{\alpha}_-(x_1,x_2,...,x_{N+1})\tilde{\phi}_\beta(x_{N+1})\,\mathrm{d}x_{N+1}$$

(72)

Now, choose $\tilde{\phi}_\alpha(x)$ and $\tilde{\phi}_\beta(x)$ to minimize the energy of the ground state and excited state, respectively. Because the anion is a doublet, the wave functions for the ground state and first excited state so obtained are not exact. However, Morrison, Day, and Smith have noted that this construction gives accurate wave functions.[23] Approximating the densities of the first triplet excited state and the ground state through these wave functions, we find

$$\rho^{\alpha\beta}_{\text{gs}}(\vec{r}) = \rho^{\alpha}_-(\vec{r}) - |\phi_\alpha(\vec{r})|^2$$
$$\rho^{\alpha\alpha}_{\text{es}}(\vec{r}) = \rho^{\alpha}_-(\vec{r}) - |\phi_\beta(\vec{r})|^2 \qquad (73)$$

Like the anion, the cation is a doublet. Hence, the wave function of the cation may be approached by annihilating either a β-spin orbital, $\tilde{\chi}_\beta(x_N)$, from $\tilde{\Psi}_{gs}^{\alpha\beta}$ or an α-spin orbital, $\tilde{\chi}_\alpha(x_N)$, from $\tilde{\Psi}_{es}^{\alpha\alpha}$:

$$\tilde{\Psi}_{+,gs}^\alpha(x_1,x_2,...,x_{N-1}) \equiv$$
$$\int \Psi_-^\alpha(x_1,x_2,...,x_{N+1})\, \phi_\alpha(x_{N+1})\, \tilde{\chi}_\beta(x_N)\, dx_{N+1}$$
$$\tilde{\Psi}_{+,es}^\alpha(x_1,x_2,...,x_{N-1}) \equiv$$
$$\int \Psi_-^\alpha(x_1,x_2,...,x_{N+1})\, \phi_\beta(x_{N+1})\, \tilde{\chi}_\alpha(x_N)\, dx_{N+1} \quad (74)$$

If we choose $\tilde{\chi}_\beta(x_N)$ and $\tilde{\chi}_\alpha(x_N)$ so that the energies of $\tilde{\Psi}_{+,gs}^\alpha$ and $\tilde{\Psi}_{+,es}^\alpha$ are minimized, we expect that the two wave functions obtained through eq 74 will resemble one another, but not be identical. If instead of choosing $\tilde{\chi}_\beta(x_N)$ and $\tilde{\chi}_\alpha(x_N)$ so as to minimize the energy, we force

$$\tilde{\Psi}_{+,gs}^\alpha(x_1,x_2,...,x_{N-1}) = \tilde{\Psi}_{+,es}^\alpha(x_1,x_2,...,x_{N-1}) \quad (75)$$

we recover eq 70. Substituting eq 70 in eq 74 reveals that

$$\rho_+(\vec{r}) \equiv \rho_-(\vec{r}) - |\phi_\alpha(\vec{r})|^2 - |\phi_\beta(\vec{r})|^2 \quad (76)$$

Equations 73 and 76 reduce eq 11 to an identity, thereby establishing the validity of proposition 1. The present justification of proposition 1 may be clarified by establishing a link between it and the justification provided in the body of the paper. The essential link is the identification of $|\phi_\alpha(\vec{r})|^2$ as an accurate approximation to the Fukui function from above, $f^+(\vec{r})$ (eq 9), and $|\phi_\beta(\vec{r})|^2$ as an accurate approximation to the Fukui function from below, $f^-(\vec{r})$ (eq 8).

This appendix demonstrates that propositions 1 and 3 may be derived from two assumptions.

(A) The wave functions of the molecular ground state, triplet excited state (or the other degenerate ground state), and cation may be obtained by annihilating some combination of spin–orbitals from the wave function of the anion.

(B) When forming the wave function of the cation, it does not matter whether one first annihilates an optimal α-spin orbital from the anion wave function and then a β-spin orbital, or vice versa.

The validity of assumption A follows from the exactness of the extended Koopmans' theorem (eq 66), when the wave function is nondegenerate,[20–22] and from the accuracy of the simplified extended Koopmans' theorem (eqs 68 and 72) when the wave function is degenerate.[23,24] The validity of assumption B is questionable. Perhaps the most compelling argument for the accuracy of assumption A is intuitive—it seems that the result of a double-ionization process should be identical whether one

first annihilates an electron from an optimal α-spin orbital and then from an optimal β-spin orbital, or vice versa.

Assumptions A and B represent sufficient, but probably not necessary, conditions for the validity of propositions 1 and 3. Nonetheless, assumptions A and B represent this work's most refined set of sufficient conditions for propositions 1 and 3.[26]

References and Notes

(1) Grochala, W.; Albrecht, A. C.; Hoffmann, R. *J. Phys. Chem. A* **2000**, *104*, 2195.

(2) Parr, R. G.; Yang, W. T. *Density Functional Theory of Atoms and Molecules;* Oxford University Press: New York, 1989.

(3) Parr, R. G.; Yang, W. T. *J. Am. Chem. Soc.* **1984**, *106*, 4049.

(4) Yang, W. T.; Parr, R. G.; Pucci, R. *J. Chem. Phys.* **1984**, *81*, 2862.

(5) The Fukui functions are formally defined through the one-sided functional derivatives: $f^{+/-}(\vec{r}) \equiv (\partial\rho(\vec{r})/\partial N)_{v_0(\vec{r})}^{+/-}$. This definition is exactly equivalent to the one presented in the text when the zero-temperature grand canonical ensemble ansatz is used to define density functionals. This is explored in more detail in the following: Ayers, P. W.; Levy, M. *Theor. Chem. Acc.*, in press.

(6) Feynman, R. P. *Phys. Rev.* **1939**, *56*, 340.

(7) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133.

(8) Chermette, H. *J. Comput. Chem.* **1999**, *20*, 129.

(9) Equation 17 can be derived from DFT, but the derivation depends on the development of a universal energy functional which may be applied to both ground and excited states. Given such a functional[10] and appropriate approximations to the functional derivatives thereof, eq 17 results from truncating the functional Taylor series expansions of $E_+ - E_{gs}$, $E_- - E_{es}$ at second order.

(10) Ayers, P. W. Unpublished.

(11) This equation assumes that the change in bond length is inversely proportional to the change in force constant. To justify this assumption, assume that one can write the bond length as a function of the force constant. Then a Taylor series expansion gives $\lambda_1 - \lambda_0 = (k_1 - k_0)(\partial\lambda/\partial k)_{\lambda=\lambda_0} + (k_1 - k_0)^2/2!(\partial^2\lambda/\partial k^2)|_{\lambda=\lambda_0} + ...$ When the change in k is small, the change in bond length is approximately proportional to the change in force constant. The proportionality constant is given by $(\partial\lambda/\partial k)|_{\lambda=\lambda_0}$, which proposition 2 predicts is less than zero.

(12) When $g = N_{valence}$, all the orbitals are equally good choices for adding (or subtracting) an electron, and hence it is possible (but *not necessary*) that the cation and anion will be formed by modifying the occupation number of the same orbital.

(13) Of course, there are only $3K - 6$ frozen coordinates in a linear molecule.

(14) Parr, R. G.; Pearson, R. G. *J. Am. Chem. Soc.* **1983**, *105*, 7512.

(15) Zhou, Z.; Parr, R. G.; Garst, J. F. *Tetrahedron Lett.* **1988**, *29*, 4843.

(16) Levy, M.; Perdew, J. P. *J. Chem. Phys.* **1986**, *84*, 4519.

(17) A paper which discusses existing methods and the proposes a new method is: Michelak, A.; De Proft, F.; Geerlings, P.; Nalewajski, R. F. *J. Phys. Chem. A* **1999**, *103*, 762.

(18) Day, O. W.; Smith, D. W.; Garrod, C. *Int. J. Quantum Chem. Symp.* **1974**, *8*, 501.

(19) Smith, D. W.; Day, O. W. *J. Chem. Phys.* **1975**, *62*, 113.

(20) Morrell, M. M.; Parr, R. G.; Levy, M. *J. Chem. Phys.* **1975**, *62*, 549.

(21) Katriel, J.; Davidson, E. R. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 4403.

(22) Morrison, R. C.; Zhou, Z.; Parr, R. G. *Theor. Chim. Acta* **1993**, *86*, 3.

(23) Morrison, R. C.; Day, O. W.; Smith, D. W. *Int. J. Quantum Chem. Symp.* **1975**, *9*, 229.

(24) Morrison, R. C. *J. Chem. Phys.* **1993**, *99*, 6221.

(25) This choice eliminates spin contamination.

(26) For instance, assumptions A and B approach necessity more closely than eq 17, since eq 17, which is a sufficient condition for proposition 1, does not follow from approximations A and B.