PHYSICAL REVIEW A VOLUME 14, NUMBER 4 OCTOBER 1976

Adiabatic approximation and the Jahn-Teller theorem

S. Aronowitz
Department of Chemistry, University of Washington, Seattle, Washington 98195

(Received 5 January 1976)

An algebraic derivation of the Jahn-Teller theorem is presented within the framework of the adiabatic approximation. It is shown that two electronic states with nonzero energies cannot be both degenerate and exhibit vanishing forces for the same set of nuclear coordinates. This, in essence, is the general Jahn-Teller theorem. The adiabatic approximation is shown to fail completely in regions where electronic states are degenerate, if at least one of the states possesses nonvanishing forces. The significance of this result is that it demonstrates that there are regions in the nuclear configuration space which cannot be traversed adiabatically.

I. INTRODUCTION

An *algebraic* derivation of the general Jahn-Teller theorem is herein reported. The theorem obtained by Jahn and Teller,[1] extended by Jahn to include spin,[2] and reformulated more recently by Clinton and Rice,[3] asserts that the ground state of a polyatomic molecule will be free of degeneracy in orbit and spin to the greatest extent possible. It also has been assumed that application of the group-theoretic arguments of Jahn and Teller to excited electronic states of the molecule would lead to conclusions similar to those arrived at in the ground-state case. We will show that this supposition is indeed true. In fact, the theorem appears to be an inherent feature of the adiabatic approximation. Since the adiabatic approximation is invoked, tacitly or explicitly, for most calculations on molecular systems, it is important to understand any structure it might impose. It is in this context that the investigation of the general behavior of the adiabatic approximation will be conducted. The consequences of electronic degeneracy will be analyzed in terms of the forces acting on the nuclei.

The electronic energies of a molecule, within the adiabatic approximation, are parametrized with respect to an independent set of variables, i.e., the nuclear coordinates. The ensuing analysis relies on the behavior of invariants of a system, e.g., the electronic energies, when they are implicit functions of a set of independent variables, for example, the nuclear coordinates. Section II is devoted to the algebraic apparatus needed for the subsequent analysis.

II. ALGEBRAIC RELATIONSHIP

This section will sketch the derivation of an algebraic equality useful to our approach. A system of functions $\lambda_j(a_1, a_2, \ldots, a_k)$, $j = 1, \ldots, N$, which depend upon the variables $a_1, a_2, \ldots, a_k$, can be treated in the context of that system's invariants. The sum of the products, i at a time without repetitions, of the λ_j are formed, and we define

$$A_1 = \sum_{j=1}^{N} \lambda_j, \quad A_2 = \sum_{k<j} \sum \lambda_j \lambda_k, \ldots, \qquad A_N = \prod_{j=1}^{N} \lambda_j . \tag{2.1}$$

Only N such sums occur. Now the λ_j may be the roots of the polynomial

$$\prod_{j=1}^{N} (E - \lambda_j)$$

for a molecular system. The various A_l are identified as the coefficients of the powers of E in the expanded version of the product above, i.e.,

$$\sum_{l=0}^{N} (-1)^l A_l (E)^{N-l} = \prod_{j=1}^{N} (E - \lambda_j) .$$

Here, $A_0 \equiv 1$. Henceforth $\lambda_j(a)$ will be used with the understanding that a represents the set of variables $a_1, a_2, \ldots, a_k$. Differentiation of the quantities in Eq. (2.1) with respect to a_i, while the remaining a_m, $m \neq i$, are held constant, followed by evaluation at $a_i = c_i$, yields

$$\sum_{j=1}^{N} \left. \frac{\partial \lambda_j}{\partial a_i} \right|_{a=c} = \left. \frac{\partial A_1}{\partial a_i} \right|_{a=c},$$
$$\sum_{j=1}^{N} \left(\sum_{l \neq j} \lambda_l\{c\} \right) \left. \frac{\partial \lambda_j}{\partial a_i} \right|_{a=c} = \left. \frac{\partial A_2}{\partial a_i} \right|_{a=c}, \ldots, \tag{2.2}$$
$$\sum_{j=1}^{N} \left(\prod_{l \neq j} \lambda_l\{c\} \right) \left. \frac{\partial \lambda_j}{\partial a_i} \right|_{a=c} = \left. \frac{\partial A_N}{\partial a_i} \right|_{a=c} .$$

Here, the abbreviated notation represents the expanded forms $(\partial \lambda_j / \partial a_i)|_{a_i = c_i, \{a_m (m \neq i) = c_m\}}$ and $\lambda_l \{a_1 = c_1, \ldots, a_k = c_k\}$. One has N equations in the N unknowns $(\partial \lambda_j / \partial a_i)|_{a=c}$ to solve in Eq. (2.2). If $\vec{\mathrm{B}}$ denotes the $N \times N$ matrix of the coefficients of $(\partial \lambda_j / \partial a_i)|_{a=c}$, it can be shown that an element

$$b_{il} = \sum_{k=0}^{i-1} (-\lambda_l)^k A_{i-k-1} .$$

The determinant of $\vec{B}$ is relatively simple to evaluate. It is readily shown that $|\vec{B}| = |\vec{V}|$, where $|\vec{V}|$ is the order N Vandermonde determinant whose elements

$$v_{ij} = (-\lambda_j)^{i-1} .$$

Now if $|\vec{V}|$ is nonvanishing, then no pair of λ_j can be identical.[4] When $|\vec{V}| \neq 0$, and thus $|\vec{B}| \neq 0$, the set of equations in Eq. (2.2) have the unique solution

$$\left.\frac{\partial\lambda_j}{\partial a_i}\right|_{a=c} = \frac{|\vec{B}|_j}{|\vec{B}|}, \quad j = 1, \ldots, N. \tag{2.3}$$

Here $|\vec{B}|_j$ is the determinant formed upon replacing the jth column of $|\vec{B}|$ with the set of $(\partial A_l/\partial a_i)|_{a=c}$, $l = 1, \ldots, N$, from the right-hand side of Eq. (2.2). Equation (2.3) results directly from the application of Cramer's rule. The determinant $|\vec{B}|_j$ also may be recognized as related to the minor formed by the deletion of the jth column of the $N \times (N+1)$ augmented matrix.[5] In turn, $|\vec{B}|_j$ is found to equal the determinant $|\vec{V}(d_k)|_j$. This latter determinant is formed by replacing the jth column of the $N \times N$ Vandermonde determinant by new elements d_k which are defined through the recursive relation

$$d_k = \left.\frac{\partial A_k}{\partial a_i}\right|_{a=c} - \sum_{j=1}^{k-1} d_j A_{k-j}, \quad k > 1,$$

$$d_1 = \left.\frac{\partial A_1}{\partial a_i}\right|_{a=c} .$$

One knows from the theory of alternants[6] that the order $N-1$ determinant $|\vec{V}|_{ij}$, formed by deletion of the ith row and jth column of the order N Vandermonde determinant, is related to the order $N-1$ Vandermonde determinant in the following fashion:

$$|\vec{V}|_{ij} = p_{N-i,j} |\vec{V}(-\lambda_1, \ldots, -\lambda_{j-1}, -\lambda_{j+1}, \ldots, -\lambda_N)| .$$

Here $|\vec{V}(-\lambda_1, \ldots, -\lambda_{j-1}, -\lambda_{j+1}, \ldots, -\lambda_N)|$ is the order $N-1$ Vandermonde determinant comprised of the powers of the remaining $-\lambda_l$, $l \neq j$, and $p_{N-i,j}$ is the sum of the products $N-i$ at a time, without repetition, of the $-\lambda_l$, $l \neq j$; the coefficient $p_0 \equiv 1$. After fairly extensive manipulation one finds that

$$|\vec{V}(d_k)|_j = \left((-1)^{N+j} \sum_{s=0}^{N-1} (-\lambda_j)^s \left.\frac{\partial A_{N-s}}{\partial a_i}\right|_{a=c} \right) \times |\vec{V}(-\lambda_1, \ldots, -\lambda_{j-1}, -\lambda_{j+1}, \ldots, -\lambda_N)| . \tag{2.4}$$

Equation (2.3), with the use of Eq. (2.4) and the relation $|\vec{B}| = |\vec{V}|$, yields

$$\left.\frac{\partial\lambda_j}{\partial a_i}\right|_{a=c} = \left((-1)^{N+j} \sum_{s=0}^{N-1} (-\lambda_j)^s \left.\frac{\partial A_{N-s}}{\partial a_i}\right|_{a=c} \right) \times \frac{|\vec{V}(-\lambda_1, \ldots, -\lambda_{j-1}, -\lambda_{j+1}, \ldots, -\lambda_N)|}{|\vec{V}|} . \tag{2.5}$$

The determinant in the numerator (right-hand side) of Eq. (2.5) can be formed from $|\vec{V}|$ by deletion of its Nth row and jth column and thus will be denoted $|\vec{V}|_{Nj}$. The determinant of the $N \times N$ minor formed by deletion of the column comprised of powers of $-\lambda_j$ in the augmented matrix $\vec{M}$ is denoted $|\vec{M}|_j$. As mentioned earlier in this section, one then finds that

$$\left((-1)^{N+j} \sum_{s=0}^{N-1} (-\lambda_j)^s \left.\frac{\partial A_{N-s}}{\partial a_i}\right|_{a=c} \right) |\vec{V}|_{Nj} = \pm |\vec{M}|_j , \tag{2.6}$$

i.e., the numerator of the expression on the right-hand side of Eq. (2.5) is equal to $|\vec{M}|_j$ in magnitude.

III. GENERAL JAHN-TELLER THEOREM

The molecule is pictured as confined in an arbitrarily large impenetrable box in order to insure that the set of electronic energy eigenvalues is denumerable. The nth eigenvector, $\Psi_n(q, a)$, is a function of both the coordinates q of the electrons and a of the nuclei. Within the framework of the adiabatic approximation, $\Psi_n(q, a)$ is given as the product[7]

$$\Psi_n(q, a) = \Phi_\nu(q, a)\Theta_\mu(a) , \tag{3.1a}$$

where

$$\hat{H}_1(q, a)\Phi_\nu(q, a) = \lambda_\nu(a)\Phi_\nu(q, a) , \qquad \hat{H}_2(a)\Theta_\mu(a) = E_n \Theta_\mu(a) . \tag{3.1b}$$

Here, $\Phi_\nu(q, a)$ is the electronic wave function whose effective Hamiltonian is $\hat{H}_1(q, a)$, and $\Theta_\mu(a)$ is the nuclear wave function whose effective Hamiltonian is $\hat{H}_2(a)$. The electronic eigenvalue is parametrized with respect to the nuclear coordinates while the nth eigenvalue E_n for the system is not. Only the $\lambda_\nu(a)$ will be considered. Now, displacement of the nuclei within the adiabatic approximation deform the electronic state without causing transitions.[7] What is implied and will be assumed is that one may calculate the electronic energy as though the nuclei were fixed at each instantaneous nuclear configuration, but *afterward* one treats an electronic energy state as a continuous function of the nuclear coordinates. This is tacitly assumed when one writes $\lambda_\nu(a)$. In order that Eqs. (3.1) be consistent and solvable, there must exist a nuclear configuration $\{c(\nu)\}$ for which $\lambda_\nu(a)$ satisfies the relationship[8]

$$\frac{\partial \lambda_\nu}{\partial a_i}\bigg|_{a=c} = \frac{\partial E_n}{\partial a_i}\bigg|_{a=c} .$$

Since E_n is independent of the a_i, this condition reduces to

$$\frac{\partial \lambda_\nu}{\partial a_i}\bigg|_{a=c} = 0 .$$

The identification of a particular configuration as an equilibrium configuration is thus intimately connected, in the adiabatic approximation, with the criterion for the consistency and solvability of the energy equations for a molecular system. The ensuing analysis will consider, in part, the forces present in the various electronic states l for a given nuclear configuration. The quantities $-(\partial\lambda_l/\partial a_j)|_{a=c}$ are the components of those forces evaluated at the particular nuclear configuration $\{c\}$. In turn, this motivates an examination of the set of equations in Eq. (2.2), since these equations delineate a relationship among the $(\partial\lambda_l/\partial a_j)|_{a=c}$.

The elements of the appropriate order N Vandermonde determinant $|\vec{V}|$ are given by

$$v_{il} = (-\lambda_l\{c\})^{i-1} , \quad i, l = 1, \ldots, N .$$

Assume that $\lambda_k = \lambda_{k+1}$, with the remaining λ_l distinct. The rank of $|\vec{V}|$ is then $N-1$. In order that the set of equations in Eq. (2.2) be consistent, the rank of the augmented matrix cannot exceed that of the matrix of coefficients; the two ranks must be equal if the set is solvable.[9] The order $N-1$ Vandermonde determinant in the numerator of Eq. (2.5) vanishes when $l \neq k, k+1$, since it then is comprised of λ_k and λ_{k+1}. This is a property of the Vandermonde determinant, as was noted in Sec. II. Hence the polynomial

$$\sum_s (-\lambda_l)^s \frac{\partial A_{N-s}}{\partial a_i}\bigg|_{a=c}$$

must vanish when $l = k, k+1$. We are led to consider

(i) $\lambda_k = \lambda_{k+1} = 0$;

or, for $\lambda_k = \lambda_{k+1} \neq 0$, when

(ii) $\displaystyle \frac{\partial A_{N-s}}{\partial a_j}\bigg|_{a=c} = 0$,

for all s and j; or, with some $(\partial A_{N-s}/\partial a_j)|_{a=c} \neq 0$,

(iii) $\displaystyle \frac{\partial \lambda_k}{\partial a_j}\bigg|_{a=c} = \frac{\partial \lambda_{k+1}}{\partial a_j}\bigg|_{a=c} = 0$,

which will cause the polynomial to vanish. Aside from an identity relationship, these three cases exhaust the alternatives, since the λ_l are not functions of the $(\partial A_{N-s}/\partial a_j)|_{a=c}$.

Case (i) is trivial. It corresponds to states where the molecule is dissociated, and therefore will be excluded from further consideration. Case (ii) transforms the set of inhomogeneous linear equations in Eq. (2.2) into a homogeneous set. If the set of values $\{c\}$ of the nuclear coordinates truly represents an equilibrium configuration for the degenerate states k and $k+1$, then the forces must vanish in these respective states at $\{c\}$. This requirement directly results in

$$\frac{\partial \lambda_k}{\partial a_j}\bigg|_{a=c} = \frac{\partial \lambda_{k+1}}{\partial a_j}\bigg|_{a=c} = 0$$

for all j. The set of homogeneous equations can be solved upon eliminating the Nth equation and transferring the products with either $(\partial\lambda_k/\partial a_j)|_{a=c}$ or $(\partial\lambda_{k+1}/\partial a_j)|_{a=c}$ to the right-hand side. Since

$$\frac{\partial \lambda_k}{\partial a_j}\bigg|_{a=c} = \frac{\partial \lambda_{k+1}}{\partial a_j}\bigg|_{a=c} = 0 ,$$

however, the resulting set of $N-1$ equations is still homogeneous. The matrix of coefficients of this set of equations has the rank of $N-1$, which follows upon noting that its determinant equals $|\vec{V}|_{Nk}$. Here, $|\vec{V}|_{Nk}$ is formed by deletion of the Nth row and kth column of $|\vec{V}|$; this order $N-1$ Vandermonde determinant cannot vanish, since it is comprised of only nondegenerate λ_l, unlike $|\vec{V}|$. Hence the set of equations has only the trivial solution

$$\frac{\partial \lambda_l}{\partial a_j}\bigg|_{a=c} = 0 ,$$

for all l. Thus all electronic states are required to display vanishing forces at the same set of values $\{c\}$ of the nuclear coordinates, i.e., the set $\{c\}$ represents an equilibrium point for the system in any λ_l. In real molecules, excited electronic states involve electron density redistribution at the very least. This consideration alone makes it physically unreasonable to expect all forces in every electronic state to vanish for the same nuclear configuration; however, we shall pursue the consequences arising from case (ii) irrespective of applicability to real systems. This is undertaken so that the analysis will be as complete as is possible.

Equation (2.2) is the first-order derivatives, evaluated at $\{c\}$, of the set of functions in Eq. (2.1). Since the set in Eq. (2.2) is identically zero for case (ii), as is clear from the discussion in the preceding paragraph, the next-higher-order derivative set will be considered. This set is obtained from the set of functions in Eq. (2.1) and is evaluated at $\{c\}$. One finds that the form taken by the set is the following:

$$\begin{aligned} &\sum_{l=1}^{N} \frac{\partial^2\lambda_l}{\partial a_j^2}\bigg|_{a=c} = \frac{\partial^2 A_1}{\partial a_j^2}\bigg|_{a=c}, \\ &\sum_{l=1}^{N}\Big(\sum_{k\neq l} \lambda_k\{c\}\Big)\frac{\partial^2\lambda_l}{\partial a_j^2}\bigg|_{a=c} = \frac{\partial^2 A_2}{\partial a_j^2}\bigg|_{a=c}, \ldots, \\ &\sum_{l=1}^{N}\Big(\prod_{k\neq l} \lambda_k\{c\}\Big)\frac{\partial^2\lambda_l}{\partial a_j^2}\bigg|_{a=c} = \frac{\partial^2 A_N}{\partial a_j^2}\bigg|_{a=c}. \end{aligned} \tag{3.2}$$

Here, all cross products $[(\partial\lambda_s/\partial a_j)(\partial\lambda_t/\partial a_j)]|_{a=c}$ vanish by virtue of the result that case (ii) leads to $(\partial\lambda_l/\partial a_j)|_{a=c} = 0$ for all j and l, given that

$$\frac{\partial\lambda_k}{\partial a_j}\bigg|_{a=c} = \frac{\partial\lambda_{k+1}}{\partial a_j}\bigg|_{a=c} = 0$$

for all j. The analysis of Eq. (3.2) separates into two parts. First, assume that

$$\text{(ii a)}\quad \frac{\partial^2 A_m}{\partial a_j^2}\bigg|_{a=c} = 0\,, \quad \text{for all } j, m\,.$$

The resulting set of homogeneous linear equations in Eq. (3.2) may be solved, using the techniques outlined in Sec. II, upon eliminating the Nth equation and transferring the appropriate products with either $(\partial^2\lambda_k/\partial a_j^2)|_{a=c}$ or $(\partial^2\lambda_{k+1}/\partial a_j^2)|_{a=c}$ to the right-hand side. One then has a set of $N-1$ inhomogeneous linear equations to solve. Say the products with $(\partial^2\lambda_k/\partial a_j^2)|_{a=c}$ have been transferred; then the determinant of the matrix of coefficients is readily shown to equal the order $N-1$ Vandermonde determinant $|\vec{V}|_{Nk}$. Here, $|\vec{V}|_{Nk}$ is formed from the order N Vandermonde determinant by deletion of the Nth row and kth column. Since the remaining λ_l $(l\neq k)$ are nondegenerate, $|\vec{V}|_{Nk}\neq 0$. Thus the set of $N-1$ inhomogeneous equations is solvable and possesses a unique solution. Application of Cramer's rule followed by fairly extensive manipulation results in

$$\frac{\partial^2\lambda_l}{\partial a_j^2}\bigg|_{a=c} = -\frac{\partial^2\lambda_k}{\partial a_j^2}\bigg|_{a=c} \frac{|\vec{V}(-\lambda_k)|_{(Nk)l}}{|\vec{V}|_{Nk}}\,. \tag{3.3}$$

Here $|\vec{V}(-\lambda_k)|_{(Nk)l}$ is the order $N-1$ Vandermonde determinant formed from $|\vec{V}|_{Nk}$ upon substituting the appropriate powers of $(-\lambda_k)$ for the elements of the lth column. Specifically, $v_{il} = (-\lambda_k)^{i-1}$. Only when $l \equiv k+1$ will $|\vec{V}(-\lambda_k)|_{(Nk)l}$ be nonvanishing. All other l values result in $|\vec{V}(-\lambda_k)|_{(Nk)l}$ having two identical columns formed from powers of $-\lambda_k$ and $-\lambda_{k+1}$. Hence

$$\begin{aligned} &\frac{\partial^2\lambda_k}{\partial a_j^2}\bigg|_{a=c} = -\frac{\partial^2\lambda_{k+1}}{\partial a_j^2}\bigg|_{a=c}, \\ &\frac{\partial^2\lambda_l}{\partial a_j^2}\bigg|_{a=c} = 0\,, \quad l\neq k, k+1\,. \end{aligned} \tag{3.4}$$

Equation (3.4) reveals that the force constants must vanish for all electronic states but the two states characterized by λ_k and λ_{k+1}. Thus nuclei do not influence each other in those states where $l\neq k, k+1$; this eliminates vibrational motion, in effect. Clearly, the result is physically incorrect. The picture is not complete, however. If, as an example, $(\partial^2\lambda_k/\partial a_j^2)|_{a=c} > 0$, then $(\partial^2\lambda_{k+1}/\partial a_j^2)|_{a=c} < 0$, and one state ($k$) has a minimum at $\{c\}$ while the other state ($k+1$) has a maximum. The degeneracy, which now could exist only at $\{c\}$, is not stable. On the other hand, $(\partial^2\lambda_k/\partial a_j^2)|_{a=c} = 0$ is not only physically unrealistic but also is unstable, since any kinetic motion would cause the system to fly apart. Therefore it is concluded that $(\partial^2 A_m/\partial a_j^2)|_{a=c}$ cannot vanish for all m and j. We now consider the resulting case,

$$\text{(ii b)}\quad \frac{\partial^2 A_m}{\partial a_j^2} \neq 0\,, \quad \text{for some } m, j\,.$$

Equation (3.2) once more represents a set of N inhomogeneous linear equations. Now the determinant of the matrix of coefficients is identical to that in Eq. (2.5), i.e., $|\vec{V}|$. It is the order N Vandermonde determinant comprised of elements

$$v_{il} = (-\lambda_l\{c\})^{i-1}, \quad i, l = 1, \ldots, N\,.$$

Its rank is $N-1$, since $\lambda_k = \lambda_{k+1}$. Suppose that λ_k were not equal to λ_{k+1}; then one could apply Cramer's rule and follow the reasoning in Sec. II to obtain

$$\frac{\partial^2\lambda_l}{\partial a_j^2}\bigg|_{a=c} = (-1)^{N+l}\sum_{s=0}^{N-1}(-\lambda_l)^s \frac{\partial^2 A_{N-s}}{\partial a_j^2}\bigg|_{a=c} \frac{|\vec{V}|_{Nl}}{|\vec{V}|}\,. \tag{3.5}$$

Here, $|\vec{V}|_{Nl}$ is the order $N-1$ determinant formed from $|\vec{V}|$ by deletion of the Nth row and lth column. Suppose for some λ_l, $l\neq k, k+1$, that $\lambda_l(\infty) < \lambda_l(c)$. If $(\partial^2\lambda_l/\partial a_j^2)|_{a=c} > 0$, there is a relative *minimum* at $\{c\}$. There also must be at least one relative *maximum* for some configuration between $\{c\}$ and $\{\infty\}$, since the slope is increasing in the neighborhood of $\{c\}$ but must ultimately decrease if the minimum at infinity is to be realized. If $(\partial^2\lambda_l/\partial a_j^2)|_{a=c} < 0$, there is a relative *maximum* at $\{c\}$. Then there also must be at least one relative *minimum* for some configuration which lies between $\{c\}$ and a totally collapsed system experiencing infinite repulsion. Thus we are confronted by the nonphysical result that there must be regions in the configuration space of the nuclei where the nuclei experience greater repulsive forces as greater separation among the nuclei is effected. Finally, if $(\partial^2\lambda_l/\partial a_j^2)|_{a=c} = 0$, the nuclei do not exert influence while in $\{c\}$ but do when some nuclei are infinitely separated. Clearly, this alternative is also impossible. We can avoid such unacceptable choices only by stipulating that a system which must display vanishing forces in every possible state for a particular nuclear configuration cannot possess *any* repulsive states. This poses an impossible condition to be satisfied by physical systems; it would mean that under all conditions the molecule at $\{c\}$, in any state, is more stable than its appropriate components.

We are forced to conclude, therefore, that case (ii) is physically unreasonable in that it requires molecular systems either to display vanishing forces in all states but $k, k+1$ or to be completely devoid of repulsive states. Quite naturally then, we turn to the consideration of case (iii).

Case (iii). The only stipulation in this case is that forces in the degenerate states must vanish at $\{c\}$. The forces in the other electronic states need not vanish. This allows for the existence of states which are repulsive as an example. Thus case (iii) represents a physically reasonable situation, unlike case (ii). It may not be obvious that case (iii) does cause the polynomial in the numerator of Eq. (2.5) to vanish; thus a derivation is presented in the Appendix.

It is well known that if the rank r of the matrix of coefficients of a solvable set of N inhomogeneous equations in N unknowns is less than N, then $N-r$ equations may be discarded, as long as the rank of the resulting matrix of coefficients of the remaining r equations is r. This set of equations can be solved by retaining r unknowns, on the left-hand side, whose $r\times r$ matrix of coefficients is of rank r, and transferring to the right-hand side the remaining $N-r$ terms. Explicitly, with reference to Eq. (2.2),

$$\begin{gathered}\sum_{l=1}^{r} \left.\frac{\partial\lambda_l}{\partial a_j}\right|_{a=c} = \left.\frac{\partial A_1}{\partial a_j}\right|_{a=c} - \sum_{l=r+1}^{N} \left.\frac{\partial\lambda_l}{\partial a_j}\right|_{a=c}, \ldots, \\ \sum_{l=1}^{r} f_l \left.\frac{\partial\lambda_l}{\partial a_j}\right|_{a=c} = \left.\frac{\partial A_r}{\partial a_j}\right|_{a=c} - \sum_{l=r+1}^{N} f_l \left.\frac{\partial\lambda_l}{\partial a_j}\right|_{a=c},\end{gathered} \tag{3.6}$$

where

$$f_l \equiv \sum_{m,\ldots,n\neq l} (\lambda_m\{c\}\cdots\lambda_n\{c\}).$$

The unknowns on the right-hand side of Eq. (3.6), i.e., $(\partial\lambda_l/\partial a_j)|_{a=c}$, $l>r$, are *completely arbitrary*. Assignment of specific values uniquely determine the unknowns on the left-hand side of Eq. (3.6), i.e., $(\partial\lambda_l/\partial a_j)|_{a=c}$, $l\leq r$. The rank of the matrix of coefficients for case (iii) is $N-1$, as was true for case (ii). This leaves one equation of the set in Eq. (2.2) superfluous. Furthermore, it can be shown that only by choosing either $(\partial\lambda_k/\partial a_j)|_{a=c}$ or $(\partial\lambda_{k+1}/\partial a_j)|_{a=c}$ to be arbitrary will the matrix of coefficients of the remaining $N-1$ unknowns be nonsingular. However, *neither* $(\partial\lambda_k/\partial a_j)|_{a=c}$ *nor* $(\partial\lambda_{k+1}/\partial a_j)|_{a=c}$ *can be arbitrary*; they both must be set equal to zero *initially* or the set of equations in Eq. (2.2) will be inconsistent and insolvable. In other words, in order to satisfy conditions ensuring the consistency and solvability of the set of equations in Eq. (2.2) we are forced to require that both $(\partial\lambda_k/\partial a_j)|_{a=c}$ and $(\partial\lambda_{k+1}/\partial a_j)|_{a=c}$ vanish; but the solution of this set of equations requires that either $(\partial\lambda_k/\partial a_j)|_{a=c}$ or $(\partial\lambda_{k+1}/\partial a_j)|_{a=c}$ be completely arbitrary. Thus one is confronted by the contradiction that if the set of equations is solvable then either $(\partial\lambda_k/\partial a_j)|_{a=c}$ or $(\partial\lambda_{k+1}/\partial a_j)|_{a=c}$ must be chosen to be arbitrary; however, neither $(\partial\lambda_k/\partial a_j)|_{a=c}$ nor $(\partial\lambda_{k+1}/\partial a_j)|_{a=c}$ can be arbitrary if the set of equations is to be solvable. This situation is *not* equivalent to choosing a physically acceptable solution from a continuum of mathematically allowed solutions; indeed, a continuum of solutions is impossible, since $(\partial\lambda_k/\partial a_j)|_{a=c}$ and $(\partial\lambda_{k+1}/\partial a_j)|_{a=c}$ can be assigned only one value if the set of equations is to remain consistent. This mathematical contradiction results directly from the premise that $\lambda_k=\lambda_{k+1}\neq 0$. Here the argument is that one would be examining a dissociated molecule if λ_k and λ_{k+1} both vanished. In turn, the set $(\partial\lambda_i/\partial a_j)|_{a=c}$ would have only the trivial solution $(\partial\lambda_l/\partial a_j)|_{a=c} = 0$, for all l. Hence one is forced to conclude that λ_k and λ_{k+1} cannot be both nonzero and degenerate, given the requirement that

$$\left.\frac{\partial\lambda_k}{\partial a_j}\right|_{a=c} = \left.\frac{\partial\lambda_{k+1}}{\partial a_j}\right|_{a=c} = 0.$$

The arguments developed in this section can be readily applied to higher-order degeneracies, as, for example, $\lambda_k=\lambda_{k+1}=\lambda_{k+2}$, without modification. The thrust of the arguments of this section also remain unchanged when $\lambda_k=\lambda_{k+1}$, $\lambda_m=\lambda_{m+1}$, $\lambda_k\neq\lambda_m$. The analysis of cases (ii) and (iii) thus reveals that the adiabatic approximation *does not permit both degeneracy and simultaneously vanishing forces* to occur among electronic states with nonzero energies in an equilibrium configuration of the nuclei. This relationship between degenerate electronic states with nonzero energies and vanishing forces is the essence of the Jahn-Teller theorem. The inescapable conclusion is that the Jahn-Teller theorem is an inherent feature of the adiabatic approximation. Extension of the analysis to include accidental degeneracies among electronic states in configurations where not all will exhibit vanishing forces is of immediate interest. This examination will be the thrust of Sec. IV.

IV. BREAKDOWN OF THE ADIABATIC APPROXIMATION

A. Accidental degeneracy, electronic states crossing

Here, the interest rests with cases where degeneracies exist but some of the involved electronic states are experiencing nonvanishing forces. Let $\lambda_l(a=d)$, $l=1,\ldots,N$, denote the set of electronic energy states when the nuclei are in the configuration $\{d\}$. It is assumed that $\lambda_k\{d\}=\lambda_{k+1}\{d\}\neq 0$, with the remaining $\lambda_l\{d\}$ nondegenerate; furthermore, let, $[\partial\lambda_k(a)/\partial a_j]|_{a=d}\neq 0$ for some j. It is clear from the discussion in Sec. III that some factors $(\partial A_m/\partial a_j)|_{a=d}$ on the right-hand side of Eq. (2.2) will not vanish, particularly since $\{d\}$ cannot refer to an equilibrium configuration. One is dealing with a set of N simultaneous linear inhomogeneous equations. Since two electronic states are degenerate, the determinant of the matrix of coefficients in Eq. (2.2), which equals the Vandermonde determinant $|\vec{V}|$, vanishes. The rank of $\vec{V}$ is $N-1$. As was true for the equilibrium case (iii), this leaves one equation of the set in Eq. (2.2) superfluous, and only by choosing either $(\partial\lambda_k/\partial a_j)|_{a=d}$ or $(\partial\lambda_{k+1}/\partial a_j)|_{a=d}$ to be arbitrary will the matrix of coefficients of the remaining $N-1$ unknowns be nonsingular. However, if the set of equations in Eq. (2.2) is to be consistent and solvable, $(\partial\lambda_k/\partial a_j)|_{a=d}$ *must be determinable* from the consistency condition mentioned in Sec. III, case (iii), i.e., $(\partial\lambda_k/\partial a_j)|_{a=e}=0$. Here $\{e\}$ denotes the equilibrium configuration for the kth electronic state. One finds formally that

$$\frac{\partial\lambda_k}{\partial a_j}\bigg|_{a=d}=\sum_{n=2}\sum_{n_i\neq n_j}\sum_{n_j=1}^{n}\frac{(d_j-e_j)^{n_j-1}}{(n_j-1)!}\times\prod_{i\neq j}\frac{(d_i-e_i)^{n_i}}{n_i!}\left(\frac{\partial^n\lambda_k}{\partial a_1^{n_1}\cdots\partial a_l^{n_l}}\bigg|_{a=e}\right),$$

$$\sum_{j=1}^{n}n_i=n\,.$$

Specifying $\{d\}$ in the adiabatic approximation scheme fixes the permissible value assignment for $(\partial\lambda_k/\partial a_j)|_{a=d}$. Thus $(\partial\lambda_k/\partial a_j)|_{a=d}$ cannot be arbitrary, or the set of equations in Eq. (2.2) will be inconsistent and insolvable. Similarly, appealing to arguments offered in the equilibrium case (iii), $(\partial\lambda_{k+1}/\partial a_j)|_{a=d}$ cannot be assigned arbitrary values. The mathematical contradiction which now arises is that the set of linear inhomogeneous equations has a matrix of coefficients whose rank is less than the number of unknowns, yet the set is required to possess an unique solution. We are forced to conclude that the adiabatic approximation scheme fails in this region. Furthermore, the breakdown of the adiabatic approximation encompasses the entire set of electronic states.

B. Application

The results obtained in Sec. IV A also demonstrate that there are regions in the configuration space of the nuclei which cannot be traversed adiabatically. These regions are delineated by the onset of degeneracy between electronic states, where forces are present in at least one of the states. Longuet-Higgins recently attempted to derive a criterion for the intersection of potential surfaces in polyatomic molecules.[10] The assumption that all pathways in the configuration space of the nuclei comprising the molecule can be traversed adiabatically is critical to his subsequent arguments. By an adiabatic traversal is meant that the electronic states can be deformed by all nuclear displacements without the incurring of transitions of any sort. The regions which must be treated adiabatically, if the approach adopted by Longuet-Higgins is valid, are those regions where degeneracy among electronic states occurs. It has been demonstrated in Sec. III, within the framework of the adiabatic approximation, that equilibrium configurations cannot be degenerate. The alternative, nonequilibrium configurations, has been dealt with in Sec. IV A. One must therefore conclude that the arguments which Longuet-Higgins employed are faulty. His assertion that electronic degeneracy is common in particular polyatomic systems is left unsupported.

V. CONCLUSION

Jahn and Teller assumed for simplicity that molecules possessing electronic degenerate states do have an equilibrium position where forces vanish for those states. It was upon considering small displacements from that equilibrium position, where forces could no longer be expected to vanish, that they showed degeneracy was broken in first order for nonlinear polyatomic molecules. Renner[11] had shown that degeneracy was broken in second order for linear molecules. It would appear that this convenient assumption of an equilibrium configuration need not be made. The imposition of the adiabatic approximation leads directly to the general Jahn-Teller theorem. This result was obtained under the broadest conditions. Electronic wave functions must be orthogonal, normalizable, and denumerable; further specification is unnecessary. It also was demonstrated that there are regions in the configuration space of the nuclei comprising the molecule which cannot be traversed adiabatically. These regions represent nonequilibrium configurations where degeneracy is present amongst some electronic states. Therefore whenever the assumption is made that the adiabatic approximation is applicable for a particular region of the nuclear configuration space of a molecule, an accompanying condition must be the exclusion of degeneracy among electronic states; otherwise, hidden contradictions and indeterminacies are introduced.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge fruitful conversations with Professor B. E. Eichinger and Professor M. P. Gouterman. This work was supported by a grant from the National Science Foundation.

APPENDIX

It will be demonstrated that setting $(\partial\lambda_l\{a\}/\partial a_i)|_{a=c}=0$ causes

$$\sum_{s=0}^{N-1}(-\lambda_l\{c\})^s\,\frac{\partial A_{N-s}}{\partial a_i}\bigg|_{a=c}$$

to vanish.

Expansion of the polynomial $f(\lambda)=\prod_{l=1}^{N}(\lambda-\lambda_l)$ yields

$$f(\lambda)=\sum_{i=0}^{N}(-1)^i A_i\lambda^{N-i}\,, \tag{A1a}$$

where the A_i, $i>0$, are defined in Eq. (2.1), and $A_0\equiv 1$. The index is converted by setting $s=N-i$. This results in

$$f(\lambda)=(-1)^N\sum_{s=0}^{N}(-\lambda)^s A_{N-s}\,. \tag{A1b}$$

Now, $f(\lambda_l)=0$, $l=1,\ldots,N$. In turn, referring to Eq. (A1b) one has

$$\sum_{s=0}^{N}(-\lambda_l)^s A_{N-s}=0\,. \tag{A2}$$

Differentiating Eq. (A2) with respect to a_i and evaluating the resulting terms at $\{c\}$ yields

$$\frac{\partial\lambda_l\{a\}}{\partial a_i}\bigg|_{a=c}\sum_{s=1}^{N}s(-\lambda_l\{c\})^{s-1}A_{N-s}\{c\}$$
$$+\sum_{s=0}^{N-1}(-\lambda_l\{c\})^s\,\frac{\partial A_{N-s}\{a\}}{\partial a_i}\bigg|_{a=c}=0\,. \tag{A3}$$

If $(\partial\lambda_l\{a\}/\partial a_i)|_{a=c}=0$, then the first term on the left-hand side of Eq. (A3) vanishes. Thus Eq. (A3) reduces to

$$\sum_{s=0}^{N-1}(-\lambda_l\{c\})^s\,\frac{\partial A_{N-s}\{a\}}{\partial a_i}\bigg|_{a=c}=0 \tag{A4}$$

when $(\partial\lambda_l/\partial a_i)|_{a=c}=0$, which was to be shown.

---

[1] H. A. Jahn and E. Teller, Proc. R. Soc. A 161, 220 (1937).

[2] H. A. Jahn, Proc. R. Soc. A 164, 117 (1938).

[3] W. L. Clinton and B. Rice, J. Chem. Phys. 30, 542 (1959); W. L. Clinton, *ibid.* 32, 626 (1960).

[4] E. T. Browne, *Introduction to the Theory of Determinants and Matrices* (Univ. of North Carolina Press, Chapel Hill, 1958), Sec. 13.

[5] The determinant formed from the minor of the augmented matrix by deletion of the jth column equals $|\tilde{B}|_j$ to within a possible change in sign. See Ref. 4, Chap. VI.

[6] T. Muir, *A Treatise on the Theory of Determinants* (Dover, New York, 1960), Sec. 337.

[7] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford U.P., London, 1954), p. 171 and Appendixes VII and VIII.

[8] Reference 7, p. 169.

[9] Reference 4, Theorem 22.2.

[10] H. C. Longuet-Higgins, Proc. R. Soc. A 344, 147 (1975).

[11] E. Renner, Z. Phys. 92, 172 (1934).