ELECTRON-LATTICE INTERACTION
AND THE
GENERALIZED BORN-OPPENHEIMER APPROXIMATION

By

ALLAN DALE PIERCE

B. Sc., New Mexico State University
(1957)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF
PHILOSOPHY
at the
MASSACHUSETTS INSTITUTE OF
TECHNOLOGY
January, 1962

Signature of Author...
Department of Physics, Oct. 16, 1961

Certified by...
Thesis Supervisor

Accepted by..
Chairman, Departmental Committee
on Graduate Students

ELECTRON-LATTICE INTERACTION AND THE GENERALIZED BORN-OPPENHEIMER APPROXIMATION

Allan D. Pierce

Submitted to the Department of Physics on October 16, 1961 in partial fulfillment of the requirement for the degree of Ph. D.

The justifications and limitations of the use of the Born-Oppenheimer approximation are discussed. A method first introduced by Renner of generalizing this approximation to include the case of electronic degeneracies is then described. This approximation is applied to a general model of the interaction of a doubly degenerate electronic level with a doubly degenerate vibration. The models of Renner effects and dynamic Jahn-Teller effects previously studied by Renner, Moffitt, Liehr, Longuet-Higgins, Thorson, and others then appear as special cases of this model and the distinction between these two classes of models is clarified. Symmetry considerations are used in finding the form of the equations resulting from the generalized Born-Oppenheimer approximation and the invariants of the resulting Hamiltonians are described. These symmetry considerations are then used in studying the properties of the approximate eigenfunctions and eigenvalues. Special consideration is given to those properties which may be of use in extending the method to solid state problems. Although this extension still requires further study, one possible method is described. This method is based on the conceptual view of a crystal as being an aggregate of coupled molecules in degenerate electronic states and leads to a phenomenological model of electron-lattice interaction in solids.

Thesis Supervisor: Laszlo Tisza
Title: Professor of Physics

ACKNOWLEDGEMENTS

The author would like to express his sincere appreciation to Professor L. Tisza, who has supervised the writing of this dissertation. He is particularly grateful to Professor Tisza for suggesting the problem considered in this thesis and for many helpful suggestions which were used in the writing of the thesis.

He would like to thank Dr. A. D. Liehr for sending copies of his unpublished work and lectures. He is also grateful to Prof. W. Thorson for reviewing an early draft of the thesis and for helpful criticisms.

Special thanks are due to his colleagues, John Kenney and Robert Futrelle, with whom many stimulating conversations were held pertaining to the subject matter of this thesis.

BIOGRAPHICAL NOTE

Allan Dale Pierce was born on December 18, 1936 in Clarinda, Iowa. He attended public schools in Augusta, Kansas, Wichita, Kansas, Artesia, New Mexico, and Las Cruces, New Mexico. He entered New Mexico State University in June, 1954, and subsequently received his B.S. in Physics with highest honours in January, 1957.

He entered the Graduate School at the Massachusetts Institute of Technology in September, 1957. His graduate education was financed by two N. S. F. Predoctoral Fellowships (1957-58, 1958-59), an N. S. F. Cooperative Fellowship (1959-60), and a ShellOil Company Fellowship (1960-61). He received a research assistantship from the Research Laboratory of Electronics in July, 1961. While a graduate student he also gained teaching experience as a recitation instructor and as a laboratory teaching assistant.

Contents

Chapter 1. INTRODUCTION.

Chapter 2. THE BORN-RENNER OR GENERALIZED BORN-OPPENHEIMER APPROXIMATION.

Chapter 3. THEORY OF THE INTERACTION OF A DOUBLY DEGENERATE VIBRATION WITH A DOUBLY DEGENERATE ELECTRONIC LEVEL.

Chapter 4. A STUDY OF THE SOLUTIONS TO THE VIBRONIC EQUATIONS

Chapter 6. A NEW HAMILTONIAN FOR ELECTRON-LATTICE INTERACTIONS IN SOLIDS.

Chapter 7. SOME CONCLUDING REMARKS.

Appendices.

Chapter 1

INTRODUCTION

1. Statement of the Problem.

Problems in molecular and solid state physics often start with a Schroedinger equation in terms of nuclear coordinates in which electrons enter only indirectly by determining the potential energy for the nuclear motion. We shall call the Hamiltonian underlying this Schroedinger equation a quasi-Hamiltonian in order to distinguish it from the exact particle Hamiltonian that contains the interactions of all electrons and nuclei.

The transition from the particle Hamiltonian to the quasi-Hamiltonian is conveyed by the adiabatic, or Born-Oppenheimer approximation. In its original form, however, this approximation has its limitations . Its application is justified only if the relevant electronic state is non-degenerate and if the spacing between electronic levels is not small. These conditions are not fulfilled in metals. Nevertheless, the standard method works well enough for most problems in the theory of metals and the effects due to electron-lattice interaction may be explained by familiar perturbation methods.

Only the phenomenon of superconductivity gives indication of a break-down of the Born-Oppenheimer method. Indeed, the Bardeen, Cooper, Schrieffer, and equivalent theories [Ref. 1, 2] can be considered as implicit corrections to the Born-Oppenheimer approximation in which the unusually strong coupling of the lattice and electrons is taken into account. The inapplicability of the standard method to superconductivity does not necessarily mean a major break-down of the Born-Oppenheimer approximation and that the method must be discarded completely - but it does indicate that a modification of the Born-Oppenheimer approximation is needed.

This train of thought leads one to consider possible methods of modifying or generalizing the Born-Oppenheimer approximation. Such a method was devised some time ago by Renner [Ref. 3] who generalized the Born-Oppenheimer approximation to treat the case of a linear molecule in a degenerate electronic state. This theory, however, applied only to linear molecules. Because of the implications of the Jahn-Teller theorem [Ref. 4] - that a non-linear molecule cannot exist in a degenerate electronic state and that it tends to distort in order to remove any electronic degeneracies (static Jahn-Teller effect) - it was felt that the necessity for generalizing the Born-Oppenheimer approximation appeared only in linear molecules.

Recently, however, the Jahn-Teller problem was reopened by Moffitt and Liehr [Ref. 5] and by Longuet-Higgins et al. [Ref. 6]. They discovered that a consideration of the electron-vibration coupling was necessary if the Jahn-Teller distortion of the molecule was sufficiently small. The effects, e. g., anomalous vibrational spectra, associated with this occurrence were called dynamic Jahn-Teller effects. Indeed, as pointed out by Liehr [Ref. 7] these should be considered the true Jahn-Teller effects and one should avoid ascribing the appearance of every non-symmetrical molecule to being a consequence of the Jahn-Teller theorem.

Renner's theory and the dynamical Jahn-Teller theories were developed independently and deal with different models but each may be considered as particular instances of a general method of extending the Born-Oppenheimer approximation beyond its traditional limitations. We call this extension the Born-Renner or generalized Born-Oppenheimer approximation. The problem this dissertation undertakes is that of giving a systematic exposition of this method of generalizing the Born-Oppenheimer approximation. In so doing we shall endeavor to explain the relationship of the two models and to show that they are both special cases of a more general model. We shall seek to determine the possible forms of the quasi-Hamiltonians resulting from the generalization and shall examine some of the properties of their eigenvalues and eigenfunctions. During the course of the development we shall seek to emphasize those features of the method which may be of relevance in application to solid state problems. Thus we shall be concerned with the general aspects of the method and not with a detailed consideration of any one given model. Later in the thesis we shall examine one possible method of constructing a phenomenological model of a solid based on this generalization of the Born-Oppenheimer approximation.

B. Reasons for the Study.

The study of this generalization of the Born-Oppenheimer approximation was motivated by the desire to provide a quantum mechanical foundation of the macroscopic thermodynamics of equilibrium, briefly denoted as MTE. This theory was recently formulated by Tisza [Ref. 8] and presents a developement of thermodynamics beyond its usual scope from a new postulational basis. In this theory the λ-type phase transitions occur when two modifications of a phase which differ in symmetry become identical. The nature of this symmetry is well understood for configurational order-disorder transitions and may be explained within the context of the Born-Oppenheimer approximation. Other λ-type phase transitions cannot be explained solely in terms of order-disorder of configuration, and the nature of the relevant symmetries is still not completely understood.

It appears that this latter class of phase transitions of the λ-type may be described by a formalism in which the Born-Oppenheimer approximation is valid only in a weakened form. Thus a study of the ways of generalizing the Born-Oppenheimer approximation seems to be a possible first step in determining the nature of these non-classical symmetries.

Although an attempt is made in this dissertation to apply the generalized Born-Oppenheimer approximation to solids, our remarks may be considered to be only a preparation for future investigations.

C. Plan of the Thesis.

In the second chapter of this thesis we have endeavored to give a careful discussion of the Born-Oppenheimer approximation - reviewing its justifications and its limitations. In the second part of the chapter we introdude the generalization of the Born-Oppenheimer approximation which will be studied in the thesis. This approximation we have called the Born-Renner approximation after its two principal innovators.

The third chapter deals with the nature of the quasi-Hamiltonians obtained by an application of the Born-Renner approximation to molecules. For simplicity we have limited our discussion for the most part to the interaction of a two-fold degenerate electronic state with a two-fold degenerate vibration. Symmetry considerations play a major role in the exposition and it is shown - just as was the case for the Born-Oppenheimer quasi-Hamiltonians - that the form of the quasi-Hamiltonian may be determined without a detailed derivation of the quasi-Hamiltonian from the exact particle Hamiltonian.

Our discussion, within the limits mentioned above, is very general and therefore of necessity incorporates some of the results obtained by Renner, Moffitt, Liehr, Thorson, and others who have written on this subject. We have endeavoured to give as complete a discussion as possible and have therefore reviewed some of the previous works. This has been done because we felt that there was a definite need for an exposition on the subject which emphasized those features of the method of interest to physicists. Hitherto, it has been studied only by workers interested in quantum chemical applications.

There are several features in our presentation which are not found in the previous papers. We do not restrict ourselves to any one particular type of molecule; and we derive our quasi-Hamiltonian solely from certain symmetry considerations common to all molecular groups having irreducble degenerate representations. For the most part, we have avoided the explicit use of group theoretical theorems. This is possible because the pertinent elements of the molecular groups are generated by only two operations: a rotation about the symmetry axis; and a reflection through a plane passing through the symmetry axis. Use is made also of the time-reversal symmetry of the exact particle-Hamiltonian and of a little-known theorem by Moffitt and Liehr, to obtain the quasi-Hamiltonian. We have added a second theorem which enables one to determine the symmetry group of the quasi-Hamiltonian. It turns out that there is a symmetry operation corresponding to each operation of the symmetry group of the ~~quasi~~-

equilibrium configuration. We have explicitly written these symmetry operations and have, in addition, determined the operator which corresponds to time-reversal symmetry. A knowledge of the symmetry operations and of certain analyticity requirements enables one to find the quasi-Hamiltonian. Furthermore, we have given the first proof of the validity of replacing the exact particle-Hamiltonian by the Born-Renner quasi-Hamiltonian. This proof, which appears in appendix C., represents the logical extension of the original derivation by Born and Oppenheimer.

The third chapter establishes the relationship of the "dynamic Jahn-Teller" problem to the "Renner" problem. The models devised previously to deal with linear and non-linear molecules, respectively, had several different features and the relationship between these models has never, in our opinion, been adequately discussed in the literature. Our general treatment enables us to simultaneously treat both cases and offers new insight into the distinction of the two models.

In Chapter 4 we classify the solutions of the quasi-Hamiltonians obtained in Chapter 3 and show that the ground state of the quasi-Hamiltonian can often be two-fold degenerate. Methods of solving the quasi-Hamiltonians are outlined for two limiting cases. The latter is largely review but is included for completeness as the results are needed in the next chapter.

In Chapter 5 we discuss the expectation value of the electric current for eigenstates of the quasi-Hamiltonian. The familiar electric current operator for the exact particle Hamiltonian is shown to be equivalent to a matrix operator depending only on the nuclear coordinates and it is shown that the expectation value of this operator is in general non-zero when the eigenstates of the quasi-Hamiltonian are non-degenerate. Such a feature is not found within the context of the Born-Oppenheimer approximation.

In Chapter 6 we attempt to determine the counterpart of the Born-Renner quasi-Hamiltonians for solids. The view is advanced there that such a quasi-Hamiltonian may be expressed as a function of the nuclear coordinates and momenta and of N spin vectors - where N is an integer proportional to the volume of the solid. Each spin vector is considered as being associated with a specific part of the crystal and we adopt the view that the crystal may be considered as a system of coupled molecules - each molecule being described by the classical coordinates of its constituent nuclei and by a non-classical entity which we call vibronic spin. By utilizing the analogy with the single molecule case we construct the symmetry operations of the crystal quasi-Hamiltonian. While these are still not uniquely determined certain plausible assumptions are made and a set of symmetry operations corresponding to these assumptions are exhibited. As an example a quasi-Hamiltonian for two coupled super cells is constructed which commutes with these symmetry operations. The extension of this method to an arbitrary number of cells is then considered and we find that additional assumptions are necessary in order that our Hamiltonian exhibit interactions between nearest neighbors only. A Hamiltonian satisfying our conditions is then exhibited.

We conclude with an appraisal of the theory and some suggestions for further study.

Most of the material which may be omitted at first reading and which may detract from the main argument of the thesis has been relegated to the appendices. These consist mainly of proofs of theorems utilized in the dissertation. With the exception of that of the Moffitt-Liehr theorem these proofs are of our own invention. Appendix D is somewhat disassociated from the central idea of the thesis but we have included it as the Jahn-Teller theorem is intimately associated with the history of the Born-Renner approximation and it has never previously been pointed out in the literature that Frohlich's one-dimensional model of a superconductor [Ref. 9] is an instance of the dynamic Jahn-Teller effect.

References are listed in the order of their appearance in the text and are tabulated at the end of the thesis. We have striven to give a complete list of references pertaining to the Born-Renner approximation although it is possible that we may have overlooked some of the very recent papers.

D. Notation.

In a work of this length it is nearly impossible to maintain an unambiguous notation throughout. Certain symbols have of necessity been utilized in different contexts but we have striven to use the same symbol for the same quantity throughout. Certain notational conventions which are apparently standard in the literature are adhered to. Vectors are indicated by arrows over the appropriate symbol and matrices are denoted by double letters, i. e., $\mathbb{A}$, $\mathbb{B}$, $\mathbb{\Gamma}$, $\mathbb{\Theta}$, $\mathbb{T}$, etc. The symbol [A, B) is reserved for the commutator of the operators A and B. Unless indicated otherwise by the use of parenthesis an operator is understood to operate on all functions appearing to its right. Sums over electron spin coordinates which should be included with integrations over electron coordinates have been omitted for simplicity - the integral sign being understood to also imply a sum over spin coordinates.

Chapter 2

THE BORN-RENNER OR GENERALIZED BORN-OPPENHEIMER APPROXIMATION

In this chapter our purpose is two-fold. First, we wish to demonstrate that there exist situations in molecular and solid state physics for which the Born-Oppenheimer approximation and its related concept of a quasi-potential is not applicable - even as a first approximation. Secondly, we wish to introduce a method of generalizing the Born-Oppenheimer approximation which will enable us to study such situations more fully.

A. The Born-Oppenheimer Approximation.

The Born-Oppenheimer, or adiabatic, approximation is an approximate method for simplifying the problem of solving the Schroedinger equation of a molecule. Its application enables one to perform in a certain sense an uncoupling of the electronic and nuclear motions. It is used widely - often unsuspectingly - in many physical theories. Thus, for example, when one uses terminology such as "the forces between atoms" he is implicitly making use of the Born-Oppenheimer approximation. The existence of the concept of "vibrational spectra" in molecules and solids as well as the theories of the specific heats of insulators is based on the notion of a quasi-potential and therefore in some sense on the assumption of the validity of the Born-Oppenheimer approximation.

Its physical meaning is evident from its stating. It assumes that the state of the molecule, in so far as the electrons are concerned, depends only on the instantaneous positions of the nuclei, i. e., the electrons move so "fast" compared to the nuclei that they adjust almost instaneously to any change in the nuclear positions. The nuclei in turn move in a quasi-potential which is simply the energy of the molecule as a function of the nuclear coordinates in the neglect of the nuclear kinetic energy. For this reason the Born-Oppenheimer approximation is often called the adiabatic approximation.

The formal stating of the Born-Oppenheimer approximation is as follows. If one writes the particle Hamiltonian for a molecule in the form

$$H = H_X + K \tag{2.1}$$

where K is the kinetic energy operator for the nuclei and H_X contains the remaining terms of the Hamiltonian, the approximation is based on the ansatz that the eigenfunctions of H are in the form

$$\Phi(x,X)\ \Psi_{B.O.}(X) \tag{2.2}$$

where $\Phi(x,X)$ is an eigenfunction of H_X:

$$H_X\, \Phi(x,X) = \mathcal{E}(X)\ \Phi(x,X) \tag{2.3}$$

The equation for $\Psi_{B.O}(X)$ may be obtained by substituting the ansatz (2.2) into the Schroedinger equation whose Hamiltonian operator is that given by (2.1) and subsequently multiplying both sides of the equation by $\Phi(x,X)^*$ and integrating over the electronic coordinates. Terms involving the derivatives of $\Phi(x,X)$ with respect to the nuclear coordinates are discarded and the following equation for $\Psi_{B.O.}(X)$ is obtained:

$$\left(K + \mathcal{E}(X)\right)\ \Psi_{B.O.}(X) = E_{B.O.}\Psi_{B.O.}(X) \tag{2.4}$$

Thus $\Psi_{B.O.}(X)$ is an eigenfunction of the <u>quasi-Hamiltonian</u> $K + \mathcal{E}(X)$ with the <u>quasi-potential</u> $\mathcal{E}(X)$. In the above we have used x to represent the totality of the electronic coordinates, both space and spin. The symbol X represents the totality of the nuclear coordinates.

B. Justifications of the Born-Oppenheimer Approximation.

As we are primarily interested in this thesis in problems for which the Born-Oppenheimer approximation is <u>not</u> applicable it is instructive to examine some of the arguments that have been given for the validity of the method.

The first justification from quantum mechanical considerations was made by Born and Oppenheimer [Ref. 10] in 1927 and it is from this fact that the approximation derives its name. Their considerations apply only to semi-rigid molecules for which there exists an <u>equilibrium configuration</u> about which the nuclei vibrate with small amplitudes compared to their relative positions. The fact that the masses of the nuclei are much greater than the electron mass led Born and Oppenheimer to consider (in a certain sense) the nuclear kinetic energy as a perturbation. To successfully do this within the framework of conventional perturbation theory, however, required some deliberation. As it was believed on intuitive grounds that the amplitudes of nuclear vibrations would go to zero as m/M goes to zero (here m is the electronic mass and M is some average nuclear mass) they argued that the amplitudes of vibrations should vary with m/M as $(m/M)^{\alpha}$ for some positive α. They chose α to be $1/4$ and let $\gamma = (m/M)^{1/4}$ be an ordering parameter for a step-wise solution of the equation

$$H \Psi = E \Psi \tag{2.5}$$

along the lines of conventional perturbation theory.

The operator H_X was expanded in a power series in the displacements of the nuclei from their equilibrium configuration, the first order term being treated as a first order perturbation, the second order term being treated as a second order perturbation, etc.. The nuclear kinetic energy K appears as being of second order[1] in this ordering scheme, the reason being that the individual terms appearing in K, i. e.,

$$- \frac{\hbar^2}{2M_k} \sum \frac{d^2}{dX_k^2}$$

can be expressed in the form

$$- \left(\frac{m}{M_k}\right)^{1/2} \frac{\hbar^2}{2m} \frac{d^2}{du_k^2} \circeq - \gamma^2 \frac{\hbar^2}{2m} \frac{d^2}{du_k^2}$$

where

$$\vec{u}_k = (\vec{X}_k - \vec{X}_{k,o}) \left(\frac{Mk}{m}\right)^{1/4}$$

is the part of the displacement of the k-th nucleus which is supposedly independent of m/M.

---

[1] Actually only the part representing the kinetic energy of internal vibrations appears in second order. The rotational energy appears in fourth order and operators representing an interaction between vibration and rotation appears in third order. The discussion we are giving in reality assumes at the outset that the contribution to the wavefunction from rotations and translations of the molecule as a whole has been factored out and that we may neglect any coupling of the rotational motions with the internal motions (vibrational and electronic) of the

This ordering may appear at first sight somewhat ad hoc. It would appear that there is no a-priori reason for both $\frac{d^2}{dx_i^2}$ and $\frac{d^2}{du_k^2}$ acting on $\bar{\Psi}$ to be of the same order of magnitude. (Here x_i is a typical electronic coordinate.) Haug and Sauermann [Ref. 13] have criticized Born and Oppenheimer's theory for this reason. Nevertheless, it can be shown [Ref. 14] that the ordering assumed by Born and Oppenheimer is the only possible choice if one wishes to work within the framework of conventional perturbation theory and yet have physically meaningful solutions. Also, Sponer and Teller [Ref. 15]

---

molecule. The operator H in (2.5) is then an operator which only gives the internal energy of the molecule, i. e., the total energy minus the rotational and translational energies. We shall call (2.5) the Schroedinger equation for the internal motions of the molecule. The nuclear and electronic coordinates are then measured relative to a coordinate system fixed in the molecule (see Nielsen, H. H. [Ref. 11]). The nuclear coordinates are furthermore subject to five or six linearly independent constraints depending on whether or not the molecule is linear.

We are presenting Born and Oppenheimer's arguments here in this form as we feel that it achieves a considerable amount of simplification and yet retains the essence of their considerations. Properly speaking, our discussion is closer to the revision of Born and Oppenheimer's justification given by Born and Huang [Ref. 12].

and El'yashevich [Ref. 16] have given rough order of magnitude arguments which apparently justify Born and Oppenheimer's ordering.

Born and Oppenheimer then compared (formally) the solutions of the equation (2.5) obtained by successive approximations in this perturbation scheme with solutions obtained by using an identical perturbation expansion of the equation (2.3). The wavefunction

$$\Psi_{B.O.}(X)\ \Phi(x, X)$$

turned out to be the solution of (2.5) correct to second order in Born and Oppenheimer's ordering scheme. However, in third order an additional term F(x, X) appeared which represented the non-adiabatic part of Ψ. This function F was the particular solution of

$$(H_{X_o} - \mathcal{E}(X_o))F = - (K, \Phi^{(1)})\ \Psi_{B.O.}(X) \qquad (2.6)$$

which was orthogonal to $\Phi(x, X_o)$ and $\Phi^{(1)}(x,X)$ are the zeroth and first order terms in a power series expansion of $\Phi(x, X)$ in the displacements of the nuclei from the equilibrium configuration. The eigenvalue E similarly turned out to be $E_{B.O.}$ correct to third order in Born and Oppenheimer's perturbation scheme. Thus Born and Oppenheimer's justification shows that the approximation correctly gives the wavefunction to second order and the energy to third order in their perturbation scheme.[1]

The fact that Born and Oppenheimer's justification applies only to semi-rigid molecules and the apparent arbitrariness of their perturbation ordering scheme has led some recent writers (Haug and Sauermann [Ref. 13] and Markham [Ref. 17]) to argue that the Born-Oppenheimer approximation is in reality no better an approximation than the <u>static approximation</u> in which the

$$\Phi(x, X)\ \Psi_{B.O.}(X)$$

of (2.2) is replaced by

$$\Phi(x, X_o)\ \Psi_{B.O.}(X)$$

where X_o denotes the nuclear equilibrium positions. Indeed, Markham has given a simple example for which in certain extreme cases the static approximation is a better approximation. Nevertheless, if one examines the Markham's results carefully, he will find that the improvement obtained by using the static approximation is at best so small as to be negligible compared to the error introduced in the exact solution by utilizing either approximation. There is, furthermore, certainly a wide range of problems in which the static approximation has no meaning. We mention among others that of

---

[1] In fourth order the influence of the term (2.8) is felt. Born and Oppenheimer did not neglect this term in their original paper. In their scheme, the rotational motion was not factored out at the outset and thus their fourth order term was the rotational energy plus a constant which included the effect of (2.8). (See footnote, pages

the molecule NH_3 which has two equivalent equilibrium configurations due to the fact that the molecule has no center of inversion [Ref. 18]. A resonance between these these two configurations is indicated by the vibrational spectra. The static approximation clearly is not applicable to such a molecule for its application has no meaning if there exists more than one equilibrium configuration. The Born-Oppenheimer approximation, however, is applicable even though the justification of its applicability lies outside the realm of justification originally given by Born and Oppenheimer.

A discussion of the Born-Oppenheimer approximation which is probably the one with which the majority of physicists are most familiar is that given by Seitz [Ref. 19]. He considered the ansatz

$$\Psi(X)\ \Phi(x,\ X)$$

as a solution of (2.1) where $\Phi(x,\ X)$ satisfied equation (2.3) and $\Psi(X)$ was as yet undetermined. The following equation was then derived for Ψ:

$$(K + \mathcal{E}(X) + \int \Phi^{*}(K, \Phi)\, d\tau_e)\Psi = E\Psi \qquad (2.7)$$

He then justified the neglect of the term

$$\int \Phi^{*}(K,\ \Phi)\ d\tau_e\ \Psi(X) \qquad (2.8)$$

by considering various possible forms of Φ (extremes of free electrons and tightly bound electrons). His considerations indicate that this term can be expected to be negligible due to the smallness of the ratio of m/M.

This discussion, however, assumes at the outset that the system is in a unique electronic state - and this assumption is an integral part of the Born-Oppenheimer approximation. Nevertheless, Seitz's arguments for the smallness of (2.8) can also be used to demonstrate the smallness of matrix elements of H between two states of the form

$$\Psi_{B.O.}(X)\ \Phi(x, X)$$

which contain different eigenstates of H_X as factors. The possibility of the neglect of these matrix elements is the true essence of the Born-Oppenheimer approximation.

Pelzer and Wigner [Ref. 20] and El'yashevich [Ref. 16] have given arguments for the validity of the Born-Oppenheimer approximation along these lines. In particular, the discussion by El'yashevich gives order of magnitude arguments for expecting (1) the spacing between vibrational levels to be $(m/M)^{1/2}$ times the spacing between electronic levels, (2) the magnitude of the matrix elements of H between different states (obtained by the Born-Oppenheimer approximation) to be $(m/M)^{3/4}$ times the spacing between electronic levels.

The fact that these order of magnitude arguments can also be used to justify Born and Oppenheimer's ordering scheme gives a firmer basis for the original justification of Born and Oppenheimer.

C. Limitations of the Born-Oppenheimer Approximation.

1. In Molecular Problems.

The justifications for the validity of the Born-Oppenheimer approximation which we have just mentioned have been subject to various assumptions. It is therefore not surprising that the Born-Oppenheimer approximation has limitations and that there exist situations for which it is not applicable. While general criteria for the validity or invalidity of the Born-Oppenheimer approximation have not yet been rigorously formulated it is safe to say that the success or failure of the approximation depends for the most part on the validity of the notion of a unique quasi-potential and therefore on the nature of the quasi-potential surfaces $\mathcal{E}(X)$ which form the eigenvalues of H_X. In each argument for the validity of the Born-Oppenheimer approximation it has always been assumed, either explicitly or implicitly, that the spacing, particularly in the region of the equilibrium configuration, between these surfaces is large in some sense.

The convergence of Born and Oppenheimer's expansion depends critically on this assumption and their method of justification would not be valid if this were not satisfied. It is difficult to say just how large this spacing should be to guarantee convergence but it is reasonable to suppose that it should be at least greater than $\hbar\omega$ where ω is some representative vibrational frequency. The order of magnitude arguments given by El'yashevich indicate that this will in general be satisfied but do not preclude the possibility that there will exist cases for which this is not true.

The possibility of a break-down of the Born-Oppenheimer approximation occuring when the spacing between two electronic levels becomes small is strongly demonstrated by a consideration of the non-adiabatic term F(x,X) which appeared in third order in Born and Oppenheimer's formulation. The correction $F_n(x,X)$ to a state

$$\Phi_n(x,X)\ \psi_{B.O.}(X)$$

can be found by solving equation (2.6) using an expansion in terms of the eigenfunctions $\Phi_n(x,X_o)$ of H_{X_o}

$$F_n(x,X) = \sum_{\substack{n \\ n \neq \bar{n}}} \frac{\int \varphi_n^*(x,X_o)\,(K,\ \Phi_{\bar{n}}^{(1)}(x,X))\,d\tau_e\ \psi_{B.O.}(X)}{\mathcal{E}_{\bar{n}}(X_o) - \mathcal{E}_n(X_o)}\ \Phi_n(x,X_o) \tag{2.9}$$

(In this equation X_0 refers to the configuration for which $\mathcal{E}_{\bar{n}}$ is minimum. The eigenfunctions $\varphi_n(x,X)$ and eigenvalues $\mathcal{E}_n(X)$ of H_X have been labeled by an index n. The function $\Psi_{B.O.}(X)$ is an eigenfunction of $K + \mathcal{E}(X)$ while the function $\varphi_{\bar{n}}^{(1)}(x,X)$ is the first order term in the expansion of $\varphi_{\bar{n}}(x,X)$ about the equilibrium configuration.)

The "smallness" of F and therefore the validity of Born and Oppenheimer's expansion is easily seen from expression (2.9) to depend on the absolute values of the spacings $\mathcal{E}_{\bar{n}}(X_0) - \mathcal{E}_{n}(X_0)$. The vanishing of one of these denominators (providing the corresponding numerator did not also vanish) would be a very clear indication of the invalidity of applying Born and Oppenheimer's method to such an example and would suggest that the Born-Oppenheimer approximation were not applicable.

The difficulties we mention we mention are similar to those in all problems in which a perturbation expansion is non-convergent because of a degeneracy or near degeneracy. In the cases we are considering, this degeneracy (or near degeneracy) occurs when two eigenstates of H_{X_0} are degenerate (or nearly degenerate), i. e., when two eigenvalues giving the electronic energy at the equilibrium configuration coincide (or very nearly coincide).

In general, such a degeneracy implies a symmetry in the operator H_{X_0}, although not necessarily. Of the possible types of symmetry which could be present the only type which will lead to a noticeable breakdown of the Born-Oppenheimer approximation is that symmetry due to a geometrically symmetrical configuration of the nuclei in their equilibrium configuration. A degeneracy due to time reversal symmetry cannot be removed by a distortion of the equilibrium configuration. Also, such a distortion cannot noticably split an approximate degeneracy which might arise if one neglects the coupling between the electron spin and the orbital motions of the electrons [Ref. 21]. If the degeneracy is split very little or not at all by a distortion of the nuclear configuration the concept of an unique quasi-potential is still retained and we would expect that any effects due to ignoring the interaction between nuclear and electronic motions to be small in such cases.

Such a case is that of the Λ-type doubling in diatomic molecules [Ref. 22]. Here the angular momentum of the electrons interacts with the nuclear angular momentum for the molecular rotation. This is a relatively small effect and is only noticeable in the spectra of diatomic molecules. For polyatomic molecules the equivalent effects due to rotation-electronic coupling are not noticable as they are obscured by the much greater effects due to electron-vibration interaction. The case of Λ-type doubling does not really require a drastic revision of the Born-Oppenheimer approximation as the notion of a quasi-potential is still retained and the coupling between rotation and electronic motion is so small that it need only be considered as a mechanism for a slight splitting of originally degenerate levels obtained by the Born-Oppenheimer approximation.

The concept of a unique quasi-potential, however, has no meaning if a degeneracy at the equilibrium configuration may be split by a distortion of this configuration, i. e., by motions of the nuclei. In this context the terminology"forces between atoms" becomes ambiguous for there is no unique potential for the nuclear motions. Such a situation will require a generalization of the Born-Oppenheimer approximation and willbbe considered in some detail in this thesis.

An interesting theorem concerning the possibility of an electronic degeneracy at the equilibrium configuration due to a geometrical symmetry is the Jahn-Teller theorem [Ref. 4]. This theorem states that a geometrical configuration at which the electronic state is degenerate cannot be stable, i. e., be an equilibrium configuration, unless the nuclear configuration is linear or the degeneracy is the two-fold Kramer's degeneracy due to time reversal symmetry.

The Jahn-Teller theorem might be taken as an argument that the situation we have been describing as leading to a breakdown of the Born-Oppenheimer approximation cannot occur in non-linear molecules. (Indeed, this tendency of non-linear molecules to have non-degenerate electronic states probably accounts for the reason why effects due to the coupling of electronic and vibrational motions in non-linear molecules were not studied until relatively recently.) However, we cannot exclude the possibility of a near degeneracy at the equilibrium configuration[1]. Among other possibilities, this circumstance can occur if the equilibrium configuration is very nearly symmetric and if the nearly degenerate electronic levels become degenerate when the nuclei assume

---

[1] Actually there will be in general no unique equilibrium configuration in such a case. By equilibrium in the above we simply mean any configuration for which the electronic energy is minimum. Since there are really two (or more) surfaces which we are considering we are referring to the minimum of the lowest surface.

the symmetric configuration. Such a possibility is known to occur in a large number of molecules [Ref. 23].

2. In Solid State Problems. The uncoupling of the electronic motions and lattice vibrations is generally always assumed as a first approximation in solid state theories. The coupling between these two motions is then usually treated as a small perturbation. This is, for example, the method employed by Bloch [Ref. 24] in his theory of electrical conductivity. As been shown by several recent writers [Ref. 25, 26, 27], this general procedure is essentially equivalent to using the Born-Oppenheimer approximation as a first approximation and treating the interactions between the states obtained in this manner as a perturbation[1].

The wide successes of the resulting solid state theories is an indication that this procedure is in a certain sense justifiable. The remarks which we have made previously in our discussion of the applicability of the Born-Oppenheimer approximation to molecules

---

[1] In actuality, we should say that these writers have indicated that the use of the Born-Oppenheimer approximation as a first step would be no worse an approximation than those usually employed. The static approximation (see section B of this chapter) is generally employed in solid state theories to represent the wavefunctions. In addition, the electon-electron interaction is generally treated by the Hartree approximation and the vibrational frequencies are assumed to be independent of the electronic state in first approximation.

suggest, however, that at best this procedure is rigorously applicable only to insulators at sufficiently low temperatures that the system may be considered as being in the ground electronic state. This state is non-degenerate and the excitation energy for an excited electronic state is greater than the vibrational excitation energies - thus indicating, just as in the case of molecules (for a solid is a large molecule), that no break-down of the Born-Oppenheimer approximation should occur.

In metals, however, the low lying electronic states form a continuum and the applicability of the Born-Oppenheimer approximation becomes suspect. The question of the applicability of the standard treatment in which a Born-Oppenheimer-like approximation is taken as a first approximation is closely analogous to the question of whether ordinary perturbation theory can be applied to a system whose unperturbed energy levels form a continuum. In spite of the mathematical difficulties of such a procedure it appears that this can be successfully done in certain instances. In a recent paper, for example, Chester and Houghton [Ref. 28] state that "the (Born-Oppenheimer) expansion...appears to converge very rapidly...for temperatures considerably greater than the Debye temperature."

The success of this method is, in our opinion, due largely to the possibility that the gross distribution of energy levels over sufficiently large energy ranges may be relatively unaltered by the perturbation (although one may have to take into account a shifting of the ground state energy level). The thermodynamic properties of the system at sufficiently large temperatures can similarly be expected to be only slightly perturbed. At sufficiently low temperatures, however, the detailed structure of the distribution of energy levels becomes important and we can expect that effects might occur which might not be explainable using theories based on the conventional procedure.

It is generally felt, for example, that the phenomenon of superconductivity lies in this category. We need only note that the most successful explanations of superconductivity to date, namely those developed along the lines of Frohlich [Ref. 29], Bardeen, Cooper, and Schrieffer [Ref. 1], and Bogoliubov et al. [Ref. 2] have all been theories in which the electron-lattice interaction has not been treated by conventional perturbation theory and have therefore essentially been theories in which the Born-Oppenheimer approximation was not taken to be the initial step.

The problem of the formulation of a theory of the polaron [Ref. 30] furnishes another example of a situation in which an initial uncoupling of the electrons and lattice vibrations should not in general be assumed. Here again we encounter a situation in which the electronic levels form a continuum. Furthermore, we here interested in the detailed structure of the states and levels of the system as a whole (lattice plus electron) and cannot take refuge in a statistical averaging over states as in the case of the theory of the thermodynamic properties of metals at room temperatures.

3. Summary. In summarizing our discussion of the limitations of the Born-Oppenheimer approximation let us first note that the arguments we have given indicate that in general the only situations which will occur that could cause a really drastic breakdown of the Born-Oppenheimer approximation and a radically different analysis are those which are associated with the break-down of the concept of a unique quasi-potential. Otherwise, the electron-vibration interaction or electron-lattice interaction may be treated as a slight perturbation with a possible splitting of originally degenerate levels.

If the concept of a unique quasi-potential breaks down, however, we may expect in general that a new formalism may be necessary and that a radically different structure of the energy spectrum may occur, the associated states involving an intimate coupling of the electron and vibrational motions not describable by the Born-Oppenheimer approximation.

As we have indicated, this concept of a unique quasi-potential can be destroyed if the electronic states are degenerate or nearly degenerate at the equilibrium configuration, or if, as may occur in solid state problems, the electronic levels form a continuum.

In some cases we may be lucky due to a statistical averaging over states and the gross energy spectrum may be given to a good approximation by the Born-Oppenheimer approximation. They we may study the properties of the system by merely treating the interactions between approximate states by conventional perturbation theories. This procedure, however, may become invalid at low temperatures when a detailed knowledge of the energy spectrum becomes necessary.

D. Renner's Theory of the π-States of CO_2.

The earliest problem of the type described in the precious section to be studied was that of the π-states of CO_2. A discussion of this problem will give a concrete illustration of many of the general remarks made previously. Furthermore, the method devised by Renner [Ref. 3] for attacking this problem furnishes a classic example of the method of generalizing the Born-Oppenheimer approximation which will be discussed in this thesis. c

The molecule CO_2 is a linear molecule: the equilibrium configuration X_o is apparently one where the carbon nucleus lies midway between the two oxygen nuclei. Because this linear configuration is axially symmetric the Hamiltonian H_{X_o} will be invariant under the group $C_{\infty v}$, i. e., rotations about the axis and reflections through any plane passing through the axis.

When the linear configuration of the molecule is distorted to a configuration X such that the carbon nucleus no longer lies on the line passing through the two oxygen nuclei the resulting configuration will no longer be invariant under the operations of $C_{\infty v}$. It will instead be invariant only under the operator σ_v'

where σ'_v denotes a reflection through the plane passing through the three nuclei. The resulting group has only one-dimensional representations in contrast to the group $C_{\infty v}$ which has two-dimensional representations as well as one-dimensional representations. We can therefore expect in general that any degenerate level of H_{X_o} will be split by a configurational distortion of this type.

The concept of a unique quasi-potential remains nearly valid for Δ, Φ, etc. states, i.e., the splitting of the degenerate levels is very small. This, however, is not the case for π-states and it was therefore here that Renner discovered the greatest necessity for generalizing the Born-Oppenheimer approximation. We shall accordingly concentrate our attention on the π-states.

In this context it is instructive to examine the nature of the two energy surfaces which degenerate to a π-eigenvalue $\varepsilon_\pi(X_o)$ of H_{X_o} at the equilibrium configuration as a function of some parameter "r" which characterizes the magnitude of the non-linear distortion of the molecule. We choose this coordinate r to be the radial coordinate of the two-fold degenerate

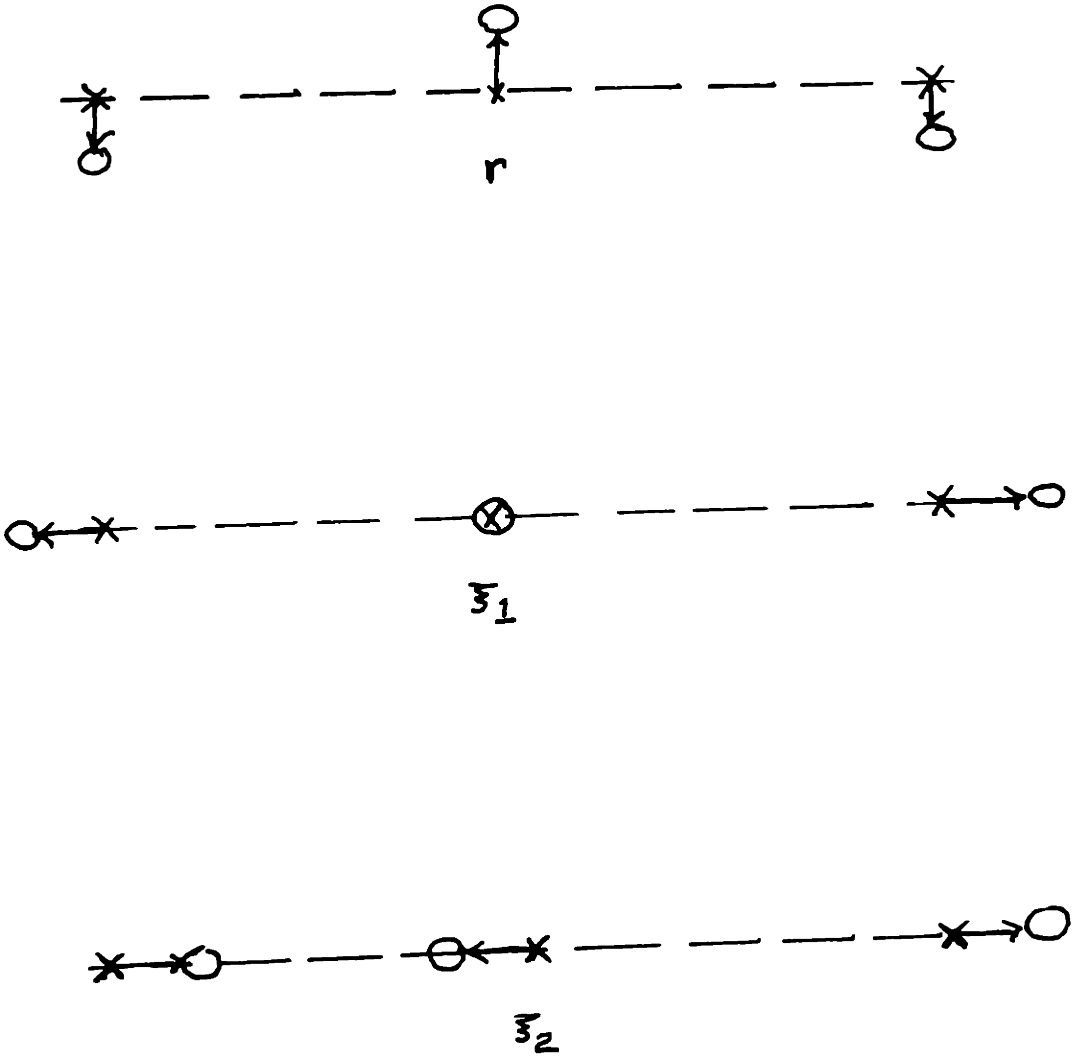


Figure 1.

Normal coordinates of the molecule CO_2.

normal vibrations of CO_2 (knick-vibrations) which disturb the linear configuration (Fig. 1). The coordinates ξ_1 and ξ_2, of the other normal vibrations do not disturb the axial symmetry and we may expect the splitting of the degenerate levels to be independent of these coordinates, at least to lowest order in r. If we let $\xi_1 = \xi_2 = 0$ the resulting energy surfaces $\mathcal{E}_\pi^{I}$ and $\mathcal{E}_\pi^{II}$, as found by Renner, are of the form (to lowest order in r)

$$\mathcal{E}_\pi^{I}(X(r)) = (\alpha + \beta)r^2 \tag{2.10a}$$

$$\mathcal{E}_\pi^{II}(X(r)) = (\alpha - \beta)r^2 \tag{2.10b}$$

apart from a constant (Fig. 2). For $\alpha \geqslant \beta \geqslant 0$ the linear configuration is the true equilibrium configuration for both surfaces and this was the case considered by Renner.

Fig. 2 illustrates how the concept of a unique quasi-potential may break down. It is impossible to say what is the restoring force on the displaced carbon nucleu tending to restore it to its equilibrium position for there is an ambiguity of potential surfaces. Furthermore, there is an ambiguity of vibrational frequencies for vibrations of the molecule away from the linear configuration.

Let us now present Renner's technique (in a

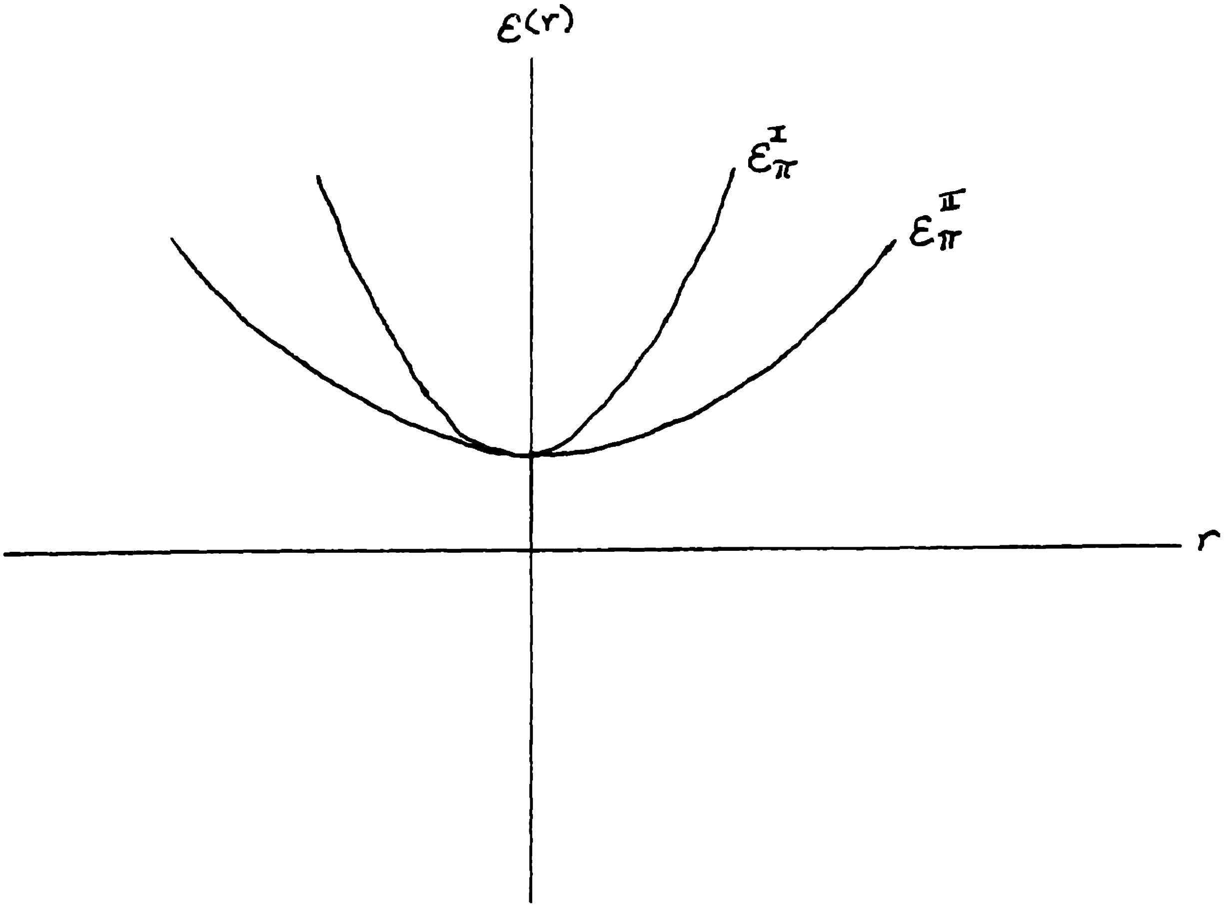


Figure 2

slightly altered form) for attacking this problem. We denote the eigenfunctions of H_X (X denoting an arbitrary configuration) associated with the eigenvalues $\mathcal{E}^{I}_{\pi}(X)$ and $\mathcal{E}^{II}_{\pi}(X)$ by $\varphi^{I}_{\pi}(x, X)$ and $\varphi^{II}_{\pi}(x, X)$, respectively. Since the Born-Oppenheimer approximation appeared inapplicable, Renner tried instead the ansatz

$$\psi^{I}(X)\, \varphi^{I}_{\pi}(x, X) + \psi^{II}(X)\, \varphi^{II}_{\pi}(x, X) \qquad (2.11)$$

with the functions $\psi^{I}(X)$ and $\psi^{II}(X)$ yet to be determined.

If this ansatz is substituted into the particle Schroedinger equation (or more precisely, the Schroedinger equation for the internal motions, as we shall assume that the rotational and translational motions have already been factored out)

$$(H_X + K)\, \Psi = E\, \Psi$$

we find

$$\varphi^{I}_{\pi}(K + \mathcal{E}^{I}_{\pi}(X) - E)\, \psi^{I} + \varphi^{II}_{\pi}(K + \mathcal{E}^{II}_{\pi}(X) - E)\, \psi^{II}$$
$$+ (K, \varphi^{I}_{\pi})\psi^{I} + (K, \varphi^{II}_{\pi})\psi^{II} = 0 \qquad (2.12)$$

If we now multiply this equation by $\varphi^{I*}_{\pi}(x,X)$ and integrate over the electronic coordinates only (also summing over electron spin coordinates) we obtain the equation

$$\left(K + \mathcal{E}^{I}_{\pi}(X)\right)\psi^{I} + \int \varphi^{I*}_{\pi}(K, \varphi^{I}_{\pi})\, d\tau_e\, \psi^{I}$$
$$+ \int \varphi^{I*}_{\pi}(K, \varphi^{II}_{\pi})\, d\tau_e\, \psi^{II} = E\, \psi^{II}$$
$$(2.13)$$

Here we have used

$$\int \varphi_\pi^{I}(x,\ X)^{*}\,\varphi_\pi^{II}(x,\ X)\ d\tau_e \ = 0$$

which follows from the fact that the two wave functions are both non-degenerate eigenfunctions of H_X[1]. We have also assumed that these wavefunctions are normalized.

A similar equation is obtained if we multiply by $\varphi_\pi^{II\,*}$ and integrate over electron coordinates:

$$(K + \mathcal{E}_\pi^{II}(X))\ \psi^{II} + \int \varphi_\pi^{II\,*}\,(K,\ \varphi_\pi^{I})\ d\tau_e\ \psi^{I} + \int \varphi_\pi^{II\,*}(K,\ \varphi_\pi^{II})\ d\tau_e\ \psi^{II} \ = E\ \psi^{II} \qquad (2.14)$$

Equations (2.13) and (2.14) constitute two coupled equations for $\psi^{I}(X)$ and $\psi^{II}(X)$. Their solutions give the eigenfunctions and eigenvalues of $K + H_X$ which are in accord with the ansatz (2.11).

The analogy of Renner's ansatz with the Born-Oppenheimer approximation becomes clearer if we write these two equations as a single equation in matrix form:

$$\left\{ K \quad + \begin{pmatrix} \mathcal{E}_\pi^{I}(X) & 0 \\ 0 & \mathcal{E}_\pi^{II}(X) \end{pmatrix} \begin{pmatrix} \int \varphi_\pi^{I\,*}(K,\ \varphi^{I})\ d\tau_e & \int \varphi_\pi^{I\,*}(K,\ \varphi^{II})\,d\tau_e \\ \int \varphi_\pi^{II\,*}(K,\ \varphi^{I})\ d\tau_e & \int \varphi_\pi^{II\,*}\ (K,\ \varphi^{II})\,d\tau_e \end{pmatrix} \right\} \begin{pmatrix} \psi^{I} \\ \psi^{II} \end{pmatrix} = E \begin{pmatrix} \psi^{I} \\ \psi^{II} \end{pmatrix} \qquad (2.15)$$

---

[1] The functions $\varphi_\pi^{I}(x,\ X_o')$ and $\varphi_\pi^{II}(x,\ X_o')$ where X_o' denotes any linear configuration are assumed to be chosen to be orthogonal.

If we in addition define

$$\vec{\psi} = \begin{pmatrix} \psi^{I} \\ \psi^{II} \end{pmatrix} \tag{2.16}$$

and

$$\boldsymbol{\mathcal{E}}(X) = \begin{pmatrix} \mathcal{E}_{\pi}^{I}(X) & 0 \\ 0 & \mathcal{E}_{\pi}^{II}(X) \end{pmatrix} + \begin{pmatrix} \int \varphi_{\pi}^{I*}(K,\ \varphi_{\pi}^{I})\, d\tau_{e} & \int \varphi_{\pi}^{I*}(K,\ \varphi_{\pi}^{II})\, d\tau_{e} \\ \int \varphi_{\pi}^{II*}(K,\ \varphi_{\pi}^{I})\, d\tau_{e} & \int \varphi_{\pi}^{II*}(K,\ \varphi_{\pi}^{I})\, d\tau_{e} \end{pmatrix} \tag{2.17}$$

this equation becomes

$$(K + \boldsymbol{\mathcal{E}}(X))\, \vec{\psi} = E\, \vec{\psi} \tag{2.18}$$

This equation is similar in form to the equation (2.4). Here, however, our quasi-potential $\boldsymbol{\mathcal{E}}(X)$ is a matrix operator instead of just a simple function of the nuclear coordinates as was the $\mathcal{E}(X)$ of (2.4). Our wavefunction $\vec{\psi}$, furthermore, has a "spinor" character as it now has two components instead of just one.

E. Definition of the Born-Renner Approximation.

The ansatz $\psi(X)\,\varphi(x,X)$ which leads to the Born-Oppenheimer approximation and the ansatz

$$\psi^{I}(X)\ \varphi^{I}(x,\ X) + \psi^{II}(X)\ \varphi^{II}(x,\ X)$$

which was utilized by Renner are both approximations to the general expansion

$$\psi = \sum_{n} \psi_n(X) \; \varphi_n(x, X) \qquad (2.19)$$

of the wavefunction in terms of the eigenfunctions $\varphi_n(x, X)$ of H_X. The first approximation is equivalent to assuming that one need only take one term in this expansion, the second assumes that only two terms are needed. Barring any unforseen mathematical fine-points any eigenfunction of $K + H_X$ can be expanded in the form (2.19) as the $\varphi_n(x, X)$ form a complete set in the electronic coordinates for any configuration X.

The equations for the $\psi(X)$ may be determined in a manner similar to that in which we obtained (2.13) and (2.14) by first substituting expression (2.19) into the total Schroedinger equation

$$(H_X + K) \Psi = E \Psi$$

and subsequently multiplying by one of the $\varphi_n(x, X)$ and integrating over the electronic coordinates. This yields a set of coupled equations of the form

$$\begin{aligned}(K + \mathcal{E}_n(X)) \psi_n(X) &= \sum_{n'} C_{nn'}(X) \psi_{n'}(X) \\ &= E \psi_n(X) \qquad (2.20)\end{aligned}$$

where

$$C_{nn'} = \int \varphi_n^*(K, \varphi_{n'}) \, d\tau_e \qquad (2.21)$$

and $\mathcal{E}_n(X)$ is the eigenvalue of H_X associated with the eigenfunction $\varphi_n(x, X)$.

Equations (2.20) were first obtained by Pelzer and Wigner [Ref. 20] as a means for initiating a method of justifying the Born-Oppenheimer approximation. They were, however, later obtained independently by Born [Ref. 31] who recognized that these equations might have some application to problems lying outside the limitations of the ordinary Born-Oppenheimer approximation[1]. In this thesis we shall refer to these equations as Born's equations and to the expansion (2.19) as *Born's expansion*, as Born's derivation is the more commonly cited in the literature.

Within this context the Born-Oppenheimer approximation and Renner's ansatz appear as no more than truncated Born expansions. In general, we shall refer to any ansatz of the form

$$\Psi = {\sum_n}' \psi_n(X) \; \varphi_n(x, X) \qquad (2.22)$$

where the sum only contains a finite (but greater than one) number of terms as a *generalized Born-Oppenheimer* or *Born-Renner approximation*. Thus Renner's theory forms a particular instance of the Born-Renner approximation.

---

[1] "It appears to me that this method may be of importance for the development of Frohlich's theory of superconductivity."

The matrix operators of the form $\mathcal{E}(X)$ such as appeared in (2.18) shall be called Born-Renner quasi-potentials or vibronic[1] quasi-potentials as they implicitly contain the vibrational and electronic interactions. In general we shall write

$$\mathcal{E}(X)_{nn'} = \mathcal{E}_n(X)\,\delta_{nn'} + C_{nn'}$$

and

$$(\vec{\Psi})_n = \Psi_n(X)$$

Thus an equation of the form

$$(K + \mathcal{E}(X))\,\vec{\Psi} = E\,\vec{\Psi} \qquad (2.23)$$

results from any Born-Renner approximation. We shall call the operators

$$K + \mathcal{E}(X)$$

Born-Renner quasi-Hamiltonians and any equation of the form (2.23) a vibronic equation.

F. Discussion.

The Born-Oppenheimer approximation presupposes that the system is in a single electronic state, i. e., an eigenstate of H_X. The Born-Renner approximation, however, presupposes that the system can be in a limited

---

[1] The word vibronic (vibration-electronic), to the best of our knowledge, was first coined by Moffitt and Thorson [Ref. 32].

number of electonic states. Thus it appears as a logical generalization of the Born-Oppenheimer approximation.

The relaxation of the constraint of confining the system to a single electronic state, however, causes us to generalize our notion of quasi-potential. The resulting Born-Renner quasi-potentials are no longer simply functions of the nuclear coordinates but are instead matrices whose components will in general be operators.

The quasi-Hamiltonian $K + \mathcal{E}(X)$ arising from the Born-Oppenheimer approximation has a classical analogue, i.e., the classical Hamiltonian for particles moving in the potential $\mathcal{E}(X)$. The existence of this classical analogue is the basis for all theories in which classical mechanics is applied to atomic particles (for example, the kinetic theory of gases). In those instances when the Born-Oppenheimer approximation breaks down and we apply instead the Born-Renner approximation we find, however, that the resulting quasi-Hamiltonian has no classical analogue. Such instances of quantum mechanical operators with no classical analogues are not new in physics, the most famous instance being that of the electron spin. Just as the introduction of the spin caused us to revise our concept of angular momentum, so must we here revise our concept of potential energy.

Chapter 3

THEORY OF THE INTERACTION OF A DOUBLY DEGENERATE VIBRATION WITH A DOUBLY DEGENERATE ELECTRONIC LEVEL

A. Introduction.

The simplest case requiring the use of the Born-Renner approximation rather than the Born-Oppenheimer approximation occur when one can consider the system as being in only two electronic states (i. e., two eigenstates of H_X) rather than being in three or more electronic states. We shall accordingly begin by restricting ourselves to the consideration of such cases.

As we mentioned in the last chapter, the greatest necessity for using the Born-Renner approximation arises when two or more electronic levels become degenerate at a configuration X which is "near" or "at" the "equilibrium" configuration. Let us therefore suppose that $\mathcal{E}^{I}(X)$ and $\mathcal{E}^{II}(X)$ are two such surfaces and that at a configuration X_o these two surfaces touch:

$$\mathcal{E}(X_o) = \mathcal{E}^{I}(X_o) = \mathcal{E}^{II}(X_o) \qquad (3.1)$$

There will generally be a range of configurations for which this is true. In this event we shall choose X_o such that (3.1) is minimal. The configuration, X_o, however, shall not necessarily be an equilibrium configuration (i. e., be a configuration for which both $\mathcal{E}^{I}(X)$ and $\mathcal{E}^{II}(X)$ are minimal). There will be two degenerate eigenfunctions[1] of H_{X_0} associated with the eigenvalue $\mathcal{E}(X_o)$. We shall assume that this degeneracy is not accidental and that these two wavefunctions transform as partners in a basis for a two-dimensional

---

[1] There are essentially two different possibilities which we may consider. The first possibility occurs when the spin-orbit terms are so small that they may be neglected. Then S and M_S are good quantum numbers and the two degenerate wavefunctions we speak of above are both associated with the same quantum numbers S and M_S. The eigenvalue $\mathcal{E}(X_o)$ will then actually be 2(2S+1)-fold degenerate but it is effectively a two-fold degeneracy as the vibrational distortions will split $\mathcal{E}(X_o)$ into at most only two levels each of which will be 2S+1 fold degenerate (M_S=S, S-1, . . ., -S).

In the second possibility, when the spin-orbit terms are not neglected, our considerations apply only to molecules containing an even number of electrons. A two-fold degenerate level for a system containing an odd number of electrons would be the Kramer's degeneracy and would not be split by any vibrational distortions. In order to have a situation requiring the Born-Renner approximation we would have to consider a four-fold degenerate level. Such an example has been treated briefly by Moffitt and Thorson [Ref. 32].

irreducible[1] representation under the operations of the symmetry group associated with the configuration X_o. These two eigenfunctions can, moreover, be chosen such that one is the complex conjugate of the other - as H_{X_o} is real[2].

The coordinates describing a nuclear configuration X can be taken as the <u>symmetry coordinates</u>[3] associated with the configuration X_o. The configuration X_o is then the configuration at which all the vibrational coordinates are zero. We may therefore associate the vibrational coordinates with irreducible representations

---

[1] In this thesis a two-dimensional representation is said to be irreducible if it cannot be decomposed by a similarity transformation into two <u>real</u> one-dimensional representations. (See Appendix B.)

[2] If the spin-orbit terms in H_{X_o} are not neglected we must modify our statement to stating that one eigenfunction may be chosen to be the result of the time reversal operator operating on the second, since H_{X_o} commutes with the time reversal operator.

[3] It is assumed that the rotational and translational motions have already been factored out and that we may ignore any interaction of the internal motions with the rotational motions. Our interest here lies solely in the electron-vibration interaction.

of the molecular symmetry group (i. e., the symmetry group of X_o) and may speak of non-degenerate vibrations, two-fold degenerate vibrations, etc.

For the present, let us assume that the molecule is constrained such that it only has two vibrational degrees of freedom, these being members of a two-fold degenerate vibration. Although this is an idealization it has certain justifications. When the configuration is distorted by a non-degenerate vibrational coordinate the electron degeneracy is in general not split. Generally, only a degenerate vibrational coordinate will split the degeneracy[1]. It turns out that the contribution to the eigenfunctions of the Born-Renner quasi-Hamiltonian from the vibrational coordinates which do not split the degeneracy can be approximately factored out leaving one with a reduced Hamiltonian which depends only on those vibrational coordinates which split the degeneracy. It is therefore fitting to begin by considering those cases in which only a two-fold degenerate vibration splits the degeneracy. We shall later give

---

[1] There are,of course, exceptions. The example of a four-atomic molecule discussed by Sponer and Teller [Ref. 15] is one such.

examples for which this is the case.

With this assumption, we may take H_X in a reduced form, assuming that it depends on only two vibrational coordinates, Q_1 and Q_2, besides the electronic coordinates. The coordinates, Q_1 and Q_2, are chosen to be linear combinations of the nuclei with real coefficients and to transform under the operations of the molecular symmetry group as a basis for a two-dimensional real orthogonal representation. For convenience we shall generally use the coordinates r and φ instead of Q_1 and Q_2 where

$$r = (Q_1^2 + Q_2^2)^{1/2}$$

and

$$\varphi = \tan^{-1}(Q_1/Q_2)$$

The coordinate r shall be called the <u>radial coordinate</u> of the two-fold degenerate vibration while φ shall be called the <u>angular coordinate</u>. The configuration X_o is then the configuration for which r = 0. Any distorted configuration X is specified by the coordinates r and φ.

The configuration of the molecule given by r = 0 must have a certain degree of symmetry - for otherwise the associated group could not have degenerate representations. All configurations associated with a group having degenerate irreducible representations must have at least a three-fold rotation axis or at least a four-fold rotation-reflection axis. Let us consider one such axis. For the present we assume that this is a P-fold axis where P is finite. The case of an infinite-fold axis shall be considered later.

If one were to rotate the configuration denoted by r, φ through an angle $2\pi/P$ about this axis (or rotate through an angle $2\pi/P$ ahd then carry out a reflection through a plane perpendicular to the axis if the axis is a rotation-reflection axis) he would obtain a configuration of twoopposible forms:

$$r, \quad \varphi + (2\pi/P)\ell \tag{3.2}$$

$$r, \quad -\varphi + (2\pi/P)\ell \tag{3.3}$$

where ℓ is an integer. It is proven in Appendix B that we can always choose an axis such that (3.2) is true and $\ell \neq 0(\text{Mod } P)$ **and** $2\ell \neq 0(\text{Mod } P)$. We assume that we have chosen such an axis. For convenience we shall refer to the particular axis chosen as the <u>symmetry axis</u> of the molecule[1]. We can furthermore always choose Q_1 and Q_2 such that ℓ is a positive integer less than P/2. The integer ℓ may be taken as a characterization of our

---

[1] In octahedral configurations, for example, we must choose the symmetry axis to be one of the three-fold rotation (or six-fold rotation-reflection) axes rather than one of the four-fold axes.

degenerate vibrations. If $\ell = 1$ we speak of π vibrations; if $\ell = 2$, we speak of Δ vibrations, etc.

We denote the operation of rotating the configuration through an angle $2\pi/P$ about the symmetry axis (or rotating through an angle $2\pi/P$ and then reflecting through a plane perpendicular to the symmetry axis) by the symbol R. Thus

$$R\varphi = \varphi + (2\pi/P)\ell \tag{3.4a}$$

$$R^2\varphi = \varphi + (2\pi/P)2\ell \tag{3.4b}$$

$$R^P\varphi = \varphi + (2\pi/P)P\ell \equiv \varphi \tag{3.4c}$$

In general, for any function $F(x, r, \varphi)$ where x denotes the electronic coordinates we define

$$RF(x, r, \varphi) = F(Rx, Rr, R\varphi) \tag{3.5}$$

where Rx symbolically denotes a rotation (or rotation-reflection) of all the electrons through an angle $2\pi/P$ about the symmetry axis. In particular, the electronic Hamiltonian H_X as well as K is invariant under R, i. e., commutes with R.

$$\begin{aligned} RH_X(x, r, \varphi)R^{-1} &= H_X(Rx, r, R\varphi) \\ &= H_X(x, r, \varphi) \end{aligned} \tag{3.6}$$

Since R commutes with H_{X_0} we may choose the eigenfunctions of H_{X_0} to be eigenstates of R:

$$R \quad \bar{\Phi}^{(0)}_{\Lambda}(x) \quad = \quad e^{i\Lambda 2\pi/P} \quad \bar{\Phi}^{(0)}_{\Lambda}(x) \qquad (3.7)$$

where Λ is an integer. Furthermore, since H_{X_0} is real,

$$\bar{\Phi}^{(0)}_{-\Lambda}(x) \quad = \quad \bar{\Phi}^{(0)}_{\Lambda}(x)^* \qquad (3.8)$$

will be degenerate with $\bar{\Phi}^{(0)}_{\Lambda}(x)$. It is shown in Appendix B that the only two-fold degenerate eigenstates of H_{X_0} are those which are associated with values of Λ for which

$$\Lambda \neq 0(\text{Mod } P) \qquad (3.9a)$$

$$2\Lambda \neq 0(\text{Mod } P) \qquad (3.9b)$$

In particular, the two degenerate states associated with the eigenvalue $\mathcal{E}(X_0)$ of (3.1) will be associated with such a value of Λ. We denote these states by

$$\bar{\Phi}^{(0)}_{\Lambda}(x) \qquad (3.10a)$$

and

$$\bar{\Phi}^{(0)}_{-\Lambda}(x) \quad = \quad \bar{\Phi}^{(0)}_{\Lambda}(x)^* \qquad (3.10b)$$

where we have chosen Λ to be positive and less than P/2. (Λ is arbitrary mod P.) The letter Λ, like ℓ for vibrational modes, will serve as a characterization of our electronic states. Thus we may speak of π, Δ, etc., electronic states.

If our axis is an infinite-fold axis, i. e., the configuration X_o has cylindrical symmetry, the integers ℓ and Λ arise analogously. If R_α denotes any rotation about the molecular axis through an angle α then the configuration (r, φ) will be transformed into a configuration

$$(r, \varphi + \ell\alpha)$$

by the operation R_α. The electronic states can similarly be chosen such that

$$R_\alpha \; \bar{\Phi}_\Lambda^{(0)}(x) = e^{i\Lambda\alpha} \; \bar{\Phi}_\Lambda^{(0)}(x)$$

and the two-fold degenerate states canbbe chosen in the form (3.10). In order that we shall not have to distinguish between the cases of infinite and finite P any more than necessary we shall from now on use R_α to denote a rotation (or rotation-reflection) through an angle $2\pi/P$ is P is finite and through any arbitrary angle α if P is infinite. Similarly α will be understood to denote the angle $2\pi/P$ if P is finite and any arbitrary angle α (but the same as in R_α) if P is infinite.

Often[1] there will be a two-fold rotation axis

---

[1] The exceptions are configurations X_o described by the groups C_p, C_{ph}, and T and all other groups not having two-dimensional irreducible representations.

perpendicular to the symmetry axis or a reflection plane which contains the symmetry axis. In this event we can speak of an operation C which denotes either a rotation through π or a reflection. For two-dimensional representations $C\varphi$ will always be of the form (3.3) and we can always choose Q_1 and Q_2 such that

$$C\varphi = -\varphi \tag{3.11}$$

and our two wavefunctions such that

$$C\,\bar{\Phi}^{(0)}_{\Lambda}(x) = \bar{\Phi}^{(0)}_{\Lambda}(Cx) = \bar{\Phi}^{(0)}_{-\Lambda}(x) \tag{3.12}$$

$$C\,\bar{\Phi}^{(0)}_{-\Lambda}(x) = \bar{\Phi}^{(0)}_{\Lambda}(x) \tag{3.13}$$

in addition to satisfying (3.10b).

B. The Moffitt-Liehr Quasi-Hamiltonian.

The general form of the Born-Renner quasi-Hamiltonian for the interaction of a two-fold degenerate vibration with a two-fold degenerate electronic level may be determined from a knowledge of the integers ℓ, Λ, and P which we have defined previously, i. e., merely from symmetry considerations. The simplest way of doing this is to first determine the form of an alternate Hamiltonian - the Moffitt-Liehr quasi-Hamiltonian -

which is related to the Born-Renner quasi-Hamiltonian and from which the Born-Renner quasi-Hamiltonian may be determined. The existence of the Moffitt-Liehr quasi-Hamiltonian follows from the following theorem:

<u>Theorem I</u>: It is possible to find functions $\bar{\Phi}_{\Lambda}(x, r, \varphi)$, $\bar{\Phi}_{-\Lambda}(x, r, \varphi)$, $U_{\Lambda\Lambda}(r, \varphi)$, $U_{\Lambda,-\Lambda}(r, \varphi)$, $U_{-\Lambda,\Lambda}(r, \varphi)$, and $U_{-\Lambda,-\Lambda}(r, \varphi)$ such that

1)
$$\bar{\Phi}_{\Lambda}(x, 0, \varphi) = \bar{\Phi}_{\Lambda}^{(0)}(x) \qquad (3.14a)$$

$$\bar{\Phi}_{-\Lambda}(x, 0, \varphi) = \bar{\Phi}_{-\Lambda}^{(0)}(x) \qquad (3.14b)$$

$$U_{\Lambda\Lambda}(0, \varphi) = U_{-\Lambda,-\Lambda}(0, \varphi) = \mathcal{E}(X_o) \qquad (3.15a)$$

$$U_{\Lambda,-\Lambda}(0, \varphi) = U_{-\Lambda,\Lambda}(0, \varphi) = 0 \qquad (3.15b)$$

2)
$$H_X \bar{\Phi}_{\Lambda} = U_{\Lambda\Lambda}(r, \varphi)\bar{\Phi}_{\Lambda} + U_{-\Lambda,\Lambda}\bar{\Phi}_{-\Lambda} \qquad (3.16a)$$

$$H_X \bar{\Phi}_{-\Lambda} = U_{-\Lambda,-\Lambda}\bar{\Phi}_{-\Lambda} + U_{\Lambda,-\Lambda}\bar{\Phi}_{\Lambda} \qquad (3.16b)$$

3) The functions $\bar{\Phi}_{\Lambda}$, $\bar{\Phi}_{-\Lambda}$ as well as $U_{\Lambda\Lambda}$, $U_{\Lambda,-\Lambda}$, $U_{-\Lambda,\Lambda}$, and $U_{-\Lambda,-\Lambda}$ are analytic[1] in $r\cos\varphi$ and $r\sin\varphi$.

4) The functions $\bar{\Phi}_{\Lambda}$ and $\bar{\Phi}_{-\Lambda}$ are orthonormal with respect to integration over the electronic coordinates, i. e.,

---

[1] By analytic we mean expandable in a Taylor series. Thus if F(x) is analytic in x then

$$F(x) = F_o + F_1 x + F_2 x^2 + \ldots$$

$$\int \bar{\Phi}_{\Lambda}^{*} \bar{\Phi}_{\Lambda} d\tau_e = 1 \qquad (3.17a)$$

$$\int \bar{\Phi}_{\Lambda}^{*} \bar{\Phi}_{-\Lambda} d\tau_e = 0 \qquad (3.17b)$$

This theorem is proven in Appendix A and its proof rests on the assumption that the other energy levels of H_{X_o} are sufficiently removed from the energy level $\mathcal{E}(X_o)$. We shall refer to this theorem as the <u>Moffitt-Liehr theorem</u> as it has been stated in the leterature previously by Moffitt and Liehr [Ref. 5].

A second theorem shows that we may choose the functions $\bar{\Phi}_{\Lambda}$, $\bar{\Phi}_{-\Lambda}$, $U_{\Lambda\Lambda}$, etc. to have definite symmetry properties:

<u>Theorem II</u>: The functions $\bar{\Phi}_{\Lambda}$, $\bar{\Phi}_{-\Lambda}$, $U_{\Lambda\Lambda}$, $U_{\Lambda,-\Lambda}$, $U_{-\Lambda,\Lambda}$, and $U_{-\Lambda,-\Lambda}$ can be chosen to satisfy the following properties in addition to those listed in Theorem I:

1) $$\bar{\Phi}_{\Lambda}(x, r, \varphi)^{*} = \bar{\Phi}_{-\Lambda}(x, r, \varphi) \qquad (3.18)$$

$$U_{\Lambda\Lambda}^{*} = U_{\Lambda\Lambda} = U_{-\Lambda,-\Lambda} \qquad (3.19a)$$

$$U_{\Lambda,-\Lambda} = U_{-\Lambda,\Lambda}^{*} \qquad (3.19b)$$

2) $$R_\alpha \bar{\Phi}_{\Lambda}(x, r, \varphi) = \bar{\Phi}_{\Lambda}(R_\alpha x, r, R_\alpha \varphi)$$

$$= e^{i\Lambda\alpha} \bar{\Phi}_{\Lambda}(x, r, \varphi) \qquad (3.20)$$

$$R_\alpha U_{\Lambda\Lambda}(r, \varphi) = U_{\Lambda\Lambda}(r, \varphi + \ell\alpha)$$
$$= U_{\Lambda\Lambda}(r, \varphi) \qquad (3.21a)$$

$$R_\alpha U_{\Lambda,-\Lambda}(r, \varphi) = e^{-i2\Lambda\alpha} U_{\Lambda,-\Lambda}(r, \varphi) \qquad (3.21b)$$

3) In addition, if there exists a reflection plane passing through the symmetry axis or a two-fold axis perpendicular to the symmetry axis

$$\bar{\Phi}_{\Lambda}(Cx, r, -\varphi) = \bar{\Phi}_{-\Lambda}(x, r, \varphi) \qquad (3.22)$$

$$U_{\Lambda\Lambda}(r, \varphi) = U_{\Lambda\Lambda}(r, -\varphi) \qquad (3.23a)$$

$$U_{\Lambda,-\Lambda}(r, \varphi) = U_{-\Lambda,\Lambda}(r, -\varphi) \qquad (3.23b)$$

Here C denotes the same operation as in equations (3.11), (3.12), and (3.13).

The proof of this theorem is also given in Appendix A and is essentially a corollary of the Moffitt-Liehr theorem.

It is clear that the eigenfunctions $\Phi^{I}(x, r, \varphi)$ and $\Phi^{II}(x, r, \varphi)$ which satisfy the equations

$$H_X \Phi^{I} = \mathcal{E}^{I} \Phi^{I}$$

and

$$H_X \Phi^{II} = \mathcal{E}^{II} \Phi^{II}$$

respectively, are linear combinations of $\bar{\Phi}_{\Lambda}$ and $\bar{\Phi}_{-\Lambda}$. If we write

$$\Phi^{I} = a^{I}_{\Lambda}\,\bar{\Phi}_{\Lambda} + a^{I}_{-\Lambda}\,\bar{\Phi}_{-\Lambda} \qquad (3.24a)$$

$$\Phi^{II} = a^{II}_{\Lambda}\,\bar{\Phi}_{\Lambda} + a^{II}_{-\Lambda}\,\bar{\Phi}_{-\Lambda} \qquad (3.24b)$$

then the coefficients are determined from the equation

$$\begin{pmatrix} U_{\Lambda\Lambda} & U_{\Lambda,-\Lambda} \\ U_{-\Lambda,\Lambda} & U_{-\Lambda,-\Lambda} \end{pmatrix} \begin{pmatrix} a^{I}_{\Lambda} & a^{II}_{\Lambda} \\ a^{I}_{-\Lambda} & a^{II}_{-\Lambda} \end{pmatrix} = \begin{pmatrix} a^{I}_{\Lambda} & a^{II}_{\Lambda} \\ a^{I}_{-\Lambda} & a^{II}_{-\Lambda} \end{pmatrix} \begin{pmatrix} \mathcal{E}^{I} & 0 \\ 0 & \mathcal{E}^{II} \end{pmatrix}$$

or, alternately, from

$$\mathcal{A}^{+}\mathcal{U}\mathcal{A} = \begin{pmatrix} \mathcal{E}^{I} & 0 \\ 0 & \mathcal{E}^{II} \end{pmatrix} \qquad (3.25)$$

where

$$\mathcal{A} = \begin{pmatrix} a^{I}_{\Lambda} & a^{II}_{\Lambda} \\ a^{I}_{-\Lambda} & a^{II}_{-\Lambda} \end{pmatrix} \qquad (3.26)$$

and

$$\mathcal{U} = \begin{pmatrix} U_{\Lambda\Lambda} & U_{\Lambda,-\Lambda} \\ U_{-\Lambda,\Lambda} & U_{-\Lambda,-\Lambda} \end{pmatrix} \qquad (3.27)$$

(The matrix $\mathcal{A}$ is required to be unitary in order that the functions Φ^{I} and Φ^{II} be orthonormal with respect to integration over the electron coordinates.) The energy surfaces Φ^{I} and Φ^{II} are therefore the eigenvalues of the matrix $\mathcal{U}$. In particular, if we let

$$U_{\Lambda\Lambda}(r,\varphi) = f(r,\varphi) \tag{3.28}$$

and

$$U_{\Lambda,-\Lambda}(r,\varphi) = g(r,\varphi) \tag{3.29}$$

then

$$\mathcal{U} = \begin{pmatrix} f & g \\ g^{*} & f \end{pmatrix} \tag{3.30}$$

from equations (3.19) and

$$\mathcal{E}^{I} = f(r,\varphi) + |g(r,\varphi)| \tag{3.31a}$$

$$\mathcal{E}^{II} = f(r,\varphi) - |g(r,\varphi)| \tag{3.31b}$$

The Born-Renner vibronic equations(i. e., that corresponding to (2.23)) in terms of $\bar{\Phi}_{\Lambda}$ and $\bar{\Phi}_{-\Lambda}$ is readily seen to be[1]

$$\begin{pmatrix} a_{\Lambda}^{I*} & a_{-\Lambda}^{I*} \\ a_{\Lambda}^{II*} & a_{-\Lambda}^{II*} \end{pmatrix} \left\{ \begin{pmatrix} \int \bar{\Phi}_{\Lambda}^{*} K \, \bar{\Phi}_{\Lambda} d\tau_e & \int \bar{\Phi}_{\Lambda}^{*} K \, \bar{\Phi}_{-\Lambda} d\tau_e \\ \int \bar{\Phi}_{-\Lambda}^{*} K \, \bar{\Phi}_{\Lambda} d\tau_e & \int \bar{\Phi}_{-\Lambda}^{*} K \, \bar{\Phi}_{-\Lambda} d\tau_e \end{pmatrix} \right.$$

$$\left. + \begin{pmatrix} U_{\Lambda,\Lambda} & U_{\Lambda,-\Lambda} \\ U_{-\Lambda,\Lambda} & U_{-\Lambda,-\Lambda} \end{pmatrix} \right\} \begin{pmatrix} a_{\Lambda}^{I} & a_{\Lambda}^{II} \\ a_{-\Lambda}^{I} & a_{-\Lambda}^{II} \end{pmatrix} \begin{pmatrix} \psi^{I} \\ \psi^{II} \end{pmatrix}$$

$$= E \begin{pmatrix} \psi^{I} \\ \psi^{II} \end{pmatrix}$$

---

[1] In this equationsas in all equations in this thesis unless specified otherwise the nuclear kinetic energy operator K is understood to operate on all functions appearing to its right.

by use of equations (3.24). In terms of matrix notation this equation becomes

$$\mathcal{A}^{\dagger}(K + \mathcal{T} + \mathcal{U})\mathcal{A}\begin{pmatrix}\psi^{I}\\ \psi^{II}\end{pmatrix} = E\begin{pmatrix}\psi^{I}\\ \psi^{II}\end{pmatrix} \qquad (3.32)$$

where $\mathcal{A}$ and $\mathcal{U}$ are defined by (3.26) and (3.27) and

$$\mathcal{T} = \begin{pmatrix}\int \bar{\Phi}^{*}_{\Lambda}(K, \bar{\Phi}_{\Lambda})\, d\tau_e & \int \bar{\Phi}^{*}_{\Lambda}(K, \bar{\Phi}_{-\Lambda})\, d\tau_e \\ \int \bar{\Phi}^{*}_{-\Lambda}(K, \bar{\Phi}_{\Lambda})\, d\tau_e & \int \bar{\Phi}^{*}_{-\Lambda}(K, \bar{\Phi}_{-\Lambda})\, d\tau_e\end{pmatrix} \qquad (3.33)$$

Our analysis therefore shows that the Born-Renner quasi-Hamiltonian is identical with the operator

$$\mathcal{A}^{\dagger}(K + n\mathcal{T} + \mathcal{U})\mathcal{A} \qquad (3.34)$$

i. e.,

$$K + \mathcal{E} = \mathcal{A}^{\dagger}(K + \mathcal{T} + \mathcal{U})\mathcal{A} \qquad (3.35)$$

Since $\mathcal{A}$ may be determined once $\mathcal{U}$ is known the properties of the Born-Renner quasi-Hamiltonian may be determined solely from those of $\mathcal{U}$ and $\mathcal{T}$. The properties of these matrices may, moreover, be studied by use of theorems I and II.

We shall call the operator

$$K + \mathcal{T} + \mathcal{U}$$

the <u>Moffitt-Liehr quasi-Hamiltonian</u>. The eigenvalues of this quasi-Hamiltonian are identical to those of the Born-Renner quasi-Hamiltonian. Furthermore, the eigenvectors of the Moffitt-Liehr quasi-Hamiltonian are related to those of the Born-Renner quasi-Hamiltonian by the matrix $\mathcal{A}$:

$$\mathcal{A}\begin{pmatrix}\psi^{I}\\ \psi^{II}\end{pmatrix} = \begin{pmatrix}\bar{\Psi}_{\Lambda}\\ \bar{\Psi}_{-\Lambda}\end{pmatrix} \tag{3.36}$$

The problem of solving the <u>Moffitt-Liehr vibronic equation</u>

$$(K + \mathcal{T} + \mathcal{U})\begin{pmatrix}\bar{\Psi}_{\Lambda}\\ \bar{\Psi}_{-\Lambda}\end{pmatrix} = E\begin{pmatrix}\bar{\Psi}_{\Lambda}\\ \bar{\Psi}_{-\Lambda}\end{pmatrix} \tag{3.37}$$

is therefore essentially equivalent to that of solving the Born-Renner vibronic equation.

In all problems considered to date the matrix $\mathcal{T}$ has either been explicitly or implicitly neglected. This neglect is equivalent to the neglect of the term

$$\int \Phi^{*}(K, \Phi)\, d\tau_{e}$$

(see equation (2.7)) in the ordinary Born-Oppenheimer approximation. In Appendix C we extend Born and Oppenheimer's method to the degenerate case and show that the contribution to the energy from this term appears only in fourth order in an ordering scheme analogous to that originally used by Born and Oppenheimer. At this point, however, we shall not neglect the matrix $\mathcal{T}$ for it is of some interest to examine the effect of its presence (or neglect) on the Born-Renner quasi-Hamiltonian.

C. Invariants of the Moffitt-Liehr Quasi-Hamiltonian.

We now consider the problem of finding the invariants of the Moffitt-Liehr quasi-Hamiltonian, i. e., of finding those operators which commute with the quasi-Hamiltonian.

The symmetry properties of the wavefunctions $\bar{\Phi}_{\Lambda}$ and $\bar{\Phi}_{-\Lambda}$ and of the matrix $\mathcal{U}$ given by Theorem II indicate that

$$R_\alpha(K + \mathcal{T} + \mathcal{U})\,R_\alpha^{-1}$$

$$= \begin{pmatrix} e^{-i\Lambda\alpha} & 0 \\ 0 & e^{i\Lambda\alpha} \end{pmatrix}(K + \mathcal{T} + \mathcal{U})\begin{pmatrix} e^{i\Lambda\alpha} & 0 \\ 0 & e^{-i\Lambda\alpha} \end{pmatrix} \tag{3.38}$$

Thus the operator Moffitt-Liehr quasi-Hamiltonian. We can

$$R_\alpha \begin{pmatrix} e^{i\Lambda\alpha} & 0 \\ 0 & e^{-i\Lambda\alpha} \end{pmatrix} \tag{3.39}$$

commutes with the Moffitt-Liehr quasi-Hamiltonian. We can formally write

$$R_\alpha = e^{\alpha \ell (d/d\varphi)} \tag{3.40}$$

and

$$\begin{pmatrix} e^{i\Lambda\alpha} & 0 \\ 0 & e^{-i\Lambda\alpha} \end{pmatrix} = e^{i\Lambda\alpha\,\sigma_z} \tag{3.41}$$

where σ_z is the Pauli matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. If we now define

$$\mathcal{J} = -i\ell\frac{d}{d\varphi} + \Lambda\sigma_z \tag{3.42}$$

we can write the expression (3.39) in the form

$$e^{i\alpha\mathcal{J}} \tag{3.43}$$

and this will commute with the quasi-Hamiltonian and thus be an invariant.

$$e^{i\alpha\mathcal{J}}(K + \mathcal{T} + \mathcal{U})\,e^{-i\alpha\mathcal{J}}$$

$$= K + \mathcal{T} + \mathcal{U} \tag{3.44}$$

The operator $\mathcal{J}$ in a certain sense corresponds to the "total angular momentum" operator if we identify $-i\ell\frac{d}{d\varphi}$ as the angular momentum operator for the vibrational motion and $\Lambda\sigma_z$ as the electronic angular momentum. The basis for this identification rests on the fact that in the limit of P being infinite, as for linear molecules, the number Λ is the electronic angular momentum about the symmetry axis and the operator $-i\ell\frac{d}{d\varphi}$ does actually denote the vibrational angular momentum operator. In this limit, moreover, $\mathcal{J}$ commutes with the Moffitt-Liehr quasi-Hamiltonian and this is interpreted as the conservation of angular momentum. The eigenvalues of $\mathcal{J}$ then become good quantum numbers with which to label the eigenfunctions. Even if $\mathcal{J}$ does not commute with the quasi-Hamiltonian it may "very nearly commute" with it. In this event the eigenvalues of $\mathcal{J}$ still remain a convenient designation for the eigenstates of the Moffitt-Liehr quasi-Hamiltonian.

Let the operation of taking a complex conjugate be denoted by $(*)$. Then the operator

$$(*)\ \sigma_x$$

where σ_x is the Pauli matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ commutes with the Moffitt-Liehr quasi-Hamiltonian. The proof of this follows from theorems I and II. We illustrate its validity by showing that it commutes with the matrix $\mathcal{U}$:

$$\begin{aligned}
(*)\ \sigma_x\ \mathcal{U} &= \sigma_x \begin{pmatrix} f & g^* \\ g & f \end{pmatrix} (*) \\
&= \begin{pmatrix} g & f \\ f & g^* \end{pmatrix} (*) \\
&= \begin{pmatrix} f & g \\ g^* & f \end{pmatrix} \sigma_x\ (*) \\
&= \mathcal{U}\,\sigma_x\ (*) = \mathcal{U}\,(*)\ \sigma_x
\end{aligned}$$

This operator $(*)\sigma_x$ is the analog of the time reversal operator for the Hamiltonian H_X. Its commuting with $K + \mathcal{T} + \mathcal{U}$ follows indirectly from the fact that $K + H_X$ is real (or that $K + H_X$ is invariant under time reversal).

If in addition an operator C commutes with H_X and we have made the appropriate choices such that equations (3.11), (3.22), and (3.23) are valid then the operator $C\sigma_x$ commutes with the Moffitt-Liehr quasi-Hamiltonian. Again we prove this for the case of $\mathcal{U}$:

$$\begin{aligned} C\,\sigma_x\,\mathcal{U} &= \sigma_x \begin{pmatrix} f(r,-\varphi) & g(r,-\varphi) \\ g^*(r,-\varphi) & f(r,-\varphi) \end{pmatrix} C \\ &= \sigma_x\,\mathcal{U}^*\,C \\ &= \mathcal{U}\,\sigma_x\,C = \mathcal{U}\,C\,\sigma_x \end{aligned}$$

as $f(r,-\varphi) = f(r,\varphi)$ and $g(r,-\varphi) = g^*(r,\varphi)$ follows from (3.23).

The operators

$$\mathcal{A}^{+} e^{i\alpha\mathcal{J}} \mathcal{A} \tag{3.45a}$$

$$\mathcal{A}^{+} (*)\,\sigma_x \mathcal{A} \tag{3.45b}$$

$$\mathcal{A}^{+} C\,\sigma_x \mathcal{A} \tag{3.45c}$$

will similarly all commute with the Born-Renner quasi-

Hamiltonian. Their forms depend on the matrix $\mathcal{A}$ and will be somewhat arbitrary as $\mathcal{A}$ is not uniquely defined by the fact that it transforms $\mathcal{U}$ to diagonal form. This is the principal reason for our initially applying our symmetry considerations to the Moffitt-Liehr quasi-Hamiltonian rather than the Born-Renner quasi-Hamiltonian.

The possibility of there being degenerate eigenfunctions of the Moffitt-Liehr quasi-Hamiltonian and therefore also of the Born-Renner quasi-Hamiltonian is demonstrated by showing that the operators $e^{i\alpha\mathcal{J}}$, $(\bigstar)\sigma_x$, and $C\sigma_x$ do not all commute. In particular

$$C\sigma_x \, e^{i\alpha\mathcal{J}} = e^{-i\alpha\mathcal{J}} \, C\,\sigma_x \, . \qquad (3.46)$$

The operators $e^{i\alpha\mathcal{J}}$ and $C\sigma_x$ together with all their distinct products and powers will form a group which will in general be a subgroup of the total symmetry group of the molecule. This group is isomorphic to the group generated by the operators C and R_α. Since the latter group has irreducible representations we expect there to be some degenerate eigenfunctions.

D. General Form of the Moffitt-Liehr Quasi-Hamiltonian.

Let us now examine the effect of the symmetries of the Moffitt-Liehr quasi-Hamiltonian on the functional dependence of the matrices $\mathcal{U}$ and $\mathcal{T}$ on the vibrational coordinates r and φ. As we have previously remarked the matrix $\mathcal{U}$ may be written in the form

$$\begin{pmatrix} f & g \\ g^* & f \end{pmatrix}$$

Since the elements of $\mathcal{U}$ are analytic in $r\cos\varphi$ and $r\sin\varphi$ (Theorem I), $f(r,\varphi)$ and $g(r,\varphi)$ may both be expanded in a series of the form

$$f(r,\varphi) = \sum_{m=-\infty}^{\infty}\sum_{t=0}^{\infty} f_{(m)}^{(t)}\, r^{|m|+2t}\, e^{im\varphi} \tag{3.47a}$$

$$g(r,\varphi) = \sum_{m=-\infty}^{\infty}\sum_{t=0}^{\infty} g_{(m)}^{(t)}\, r^{|m|+2t}\, e^{im\varphi} \tag{3.47b}$$

From equations (3.15) we find

$$f_{(0)}^{(0)} = \mathcal{E}(x_0) \tag{3.48}$$

and

$$g_{(0)}^{(0)} = 0 \tag{3.49}$$

The symmetry properties (3.21) imply that

$$f(r,\varphi+\ell\alpha) = f(r,\varphi) \tag{3.50a}$$

$$g(r, \varphi + \ell\alpha) = e^{-2i\Lambda\alpha} g(r, \varphi) \qquad (3.50a)$$

which in turn imply that

$$f^{(t)}_{(m)} = 0 \quad \text{unless } m\ell = 0 (\text{Mod } P) \qquad (3.51)$$

$$g^{(t)}_{(m)} = 0 \quad \text{unless } m\ell = -2\Lambda (\text{Mod } P) \qquad (3.52)$$

Also, since $f(r, \varphi)$ is real

$$f^{(t)}_{(m)} = f^{(t)*}_{(-m)} \qquad (3.53)$$

If we in addition can speak of an operator C as when the molecular configuration X_0 has a reflection plane passing through the symmetry axis or a rotational axis perpendicular to the symmetry axis then (3.23) implies that

$$f(r, \varphi) = f(r, -\varphi) \qquad (3.54)$$

and

$$g(r, -\varphi) = g(r, \varphi)^* \qquad (3.55)$$

These in turn, in conjunction with equation (3.53) imply that the coefficients $f^{(t)}_{(m)}$ and $g^{(t)}_{(m)}$ are all real. Thus we see that symmetry considerations greatly restrict the form of $\mathcal{U}$.

Let us now consider the matrix $\mathcal{T}$. The nuclear kinetic energy operator K may be taken in the form

$$K = -\frac{\hbar^2}{2M}\left(\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr} + \frac{1}{r^2}\frac{d^2}{d\varphi^2}\right) \qquad (3.56)$$

where M is an effective nuclear mass for the vibration. Then one can show that $\mathcal{T}$ must be of the form

$$\mathcal{T} = (\hbar^2/2M) \left[\mathcal{T}_{(0)} + i \frac{d}{dr}(r\, \mathcal{T}_{(1)}) + 2ir\, \mathcal{T}_{(1)} \frac{d}{dr} + i \frac{d}{d\varphi} (\mathcal{T}_{(2)}) + 2i\, \mathcal{T}_{(2)} \frac{d}{d\varphi} \right] \qquad (3.57)$$

where

$$(\mathcal{T}_{(0)})_{\Lambda\Lambda'} = \int \frac{d}{dr} \bar{\Phi}_{\Lambda}^{*} \frac{d}{dr} \bar{\Phi}_{\Lambda'} \, d\tau_e + \frac{1}{r^2} \int \frac{d}{d\varphi} \bar{\Phi}_{\Lambda}^{*} \frac{d}{d\varphi} \bar{\Phi}_{\Lambda'} \, d\tau_e \qquad (3.58)$$

$$(\mathcal{T}_{(1)})_{\Lambda\Lambda'} = i \int \bar{\Phi}_{\Lambda}^{*} \frac{1}{r} \frac{d}{dr} (\bar{\Phi}_{\Lambda'}) \, d\tau_e \qquad (3.59)$$

$$(\mathcal{T}_{(2)})_{\Lambda\Lambda'} = \frac{1}{r^2} \int \bar{\Phi}_{\Lambda}^{*} \frac{d}{d\varphi} (\bar{\Phi}_{\Lambda'}) \, d\tau_e \qquad (3.60)$$

It is readily verified that the matrices $\mathcal{T}_{(0)}$, $\mathcal{T}_{(1)}$, and $\mathcal{T}_{(2)}$ are Hermitian. Furthermore, they must be of the forms:

$$\mathcal{T}_{(0)} = \begin{pmatrix} h_0(r,\varphi) & j_0(r,\varphi) \\ j_0^{*}(r,\varphi) & h_0(r,\varphi) \end{pmatrix} \qquad (3.61)$$

$$\mathcal{T}_{(1)} = \frac{1}{r^2} \begin{pmatrix} h_1(r,\varphi) & 0 \\ 0 & -h_1(r,\varphi) \end{pmatrix} = \frac{1}{r^2} h_1 \sigma_z \qquad (3.62)$$

$$\mathcal{T}_{(2)} = \frac{1}{r^2} \begin{pmatrix} h_2(r,\varphi) & 0 \\ 0 & -h_2(r,\varphi) \end{pmatrix} = \frac{1}{r^2} h_2 \sigma_z \qquad (3.63)$$

where the functions $h_{(0)}$, $j_{(0)}$, $h_{(1)}$, and $h_{(2)}$ are all analytic in $r \cos \varphi$ and $r \sin \varphi$. The fact that these matrices are Hermitian, moreover, implies that h_o, h_1, and h_2 are real. It is also easily verified that the quantities h_1/r^2 and h_2/r^2 are finite in the limit as r approcaches zero from considering the fact that the wave functions $\bar{\Phi}_\Lambda(x, r, \varphi)$ and $\bar{\Phi}_{-\Lambda}(x, r, \varphi)$ are orthonormal to all orders in r.[1]

The symmetry properties of these functions h_o, h_1, h_2, and j_o are analogous to those of the functions f and g:

$$h_o(r, \varphi + \ell\alpha) = h_o(r, \varphi) \tag{3.64a}$$

$$h_1(r, \varphi + \ell\alpha) = h_1(r, \varphi) \tag{3.64b}$$

$$h_2(r, \varphi + \ell\alpha) = h_2(r, \varphi) \tag{3.64c}$$

$$j_o(r, \varphi + \ell\alpha) = e^{-2i\Lambda\alpha} j_o(r, \varphi) \tag{3.64d}$$

and considerations such as (3.51), (3.52), and (3.53) apply. If the operator C exists we find in addition

---

[1] The argument is actually more subtle than presented above. It is shown in Appendix A that we may choose Φ_Λ and $\bar{\Phi}_{-\Lambda}$ such that

$$\int \bar{\Phi}_\Lambda^{(0)*} \left(\frac{d}{dr} \bar{\Phi}_\Lambda \right)_{r=0} d\tau_e = 0 \tag{3.65}$$

etc. The orthonormality only implies that the above quantity be purely imaginary.

that

$$h_0(r, -\varphi) = h_0(r, \varphi) \quad (3.66a)$$

$$h_1(r, -\varphi) = -h_1(r, \varphi) \quad (3.66b)$$

$$h_2(r, -\varphi) =n +h_2(r, \varphi) \quad (3.66c)$$

$$j_0(r, -\varphi) = j_0(r, \varphi)^* \quad (3.66d)$$

E. The Case of Linear Molecules.

Let us now consider the case where P is infinite and ℓ is 1. We assume also that there is a reflection plane passing through the symmetry axis such that we may speak of an operator C. This example includes all linear molecules in a degenerate electronic state and, in particular, the problem considered by Renner.

From equations (3.47), (3.51), and (3.52) we find

$$f(r, \varphi) = \sum_{t=0}^{\infty} f^{(t)}_{(0)} r^2 = \bar{f}(r^2) \quad (3.67)$$

and

$$g(r, \varphi) = r^{2\Lambda} e^{-12\Lambda\varphi} \bar{g}(r^2) \quad (3.68)$$

where $\bar{f}(r^2)$ and $\bar{g}(r^2)$ are analytic functions of r^2. Equation (3.55) furthermore shows that $\bar{g}(r^2)$ is real. Similarly

$$h_0(r, \varphi) = \bar{h}_0(r^2) \quad (3.69a)$$

$$h_1(r,\varphi) = 0 \tag{3.69b}$$

$$h_2(r,\varphi) = \bar{h}_2(r^2) \tag{3.69c}$$

$$j_o(r,\varphi) = r^{2\Lambda} e^{-i2\Lambda\varphi} \bar{j}_o(r^2) \tag{3.69d}$$

where $\bar{h}_o$, $\bar{h}_2$, and $\bar{j}_o$ are all real analytic functions of r^2. Since $\bar{h}_2/r^2$ is finite as r approaches zero the function

$$\bar{\bar{h}}_2 = \frac{\bar{h}_2(r^2)}{r^2} \tag{3.70}$$

will also be an analytic function of r^2. With these facts at hand the Moffitt-Liehr quasi-Hamiltonian may now be written down:

$$\begin{aligned} K \quad + \frac{\hbar^2}{2M} \Big[\bar{h}_o(r^2) + r^{2\Lambda} \bar{j}_o(r^2) & \begin{pmatrix} 0 & e^{-i2}e^{-i2\Lambda\varphi} \\ e^{i2\Lambda\varphi} & 0 \end{pmatrix} \\ & + 2i\, \sigma_z \bar{\bar{h}}_2(r^2) \frac{d}{d\varphi} \quad \Big] \\ + \bar{f}(r^2) \quad + r^{2\Lambda}\bar{g}(r^2) & \begin{pmatrix} 0 & e^{-i2\Lambda\varphi} \\ e^{i2\Lambda\varphi} & 0 \end{pmatrix} \end{aligned} \tag{3.71}$$

For the matrix $\mathcal{A}$ which transforms this equation to the corresponding Born-Renner quasi-Hamiltonian we may take

$$\mathcal{A} = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\Lambda\varphi} & ie^{-i\Lambda\varphi} \\ e^{i\Lambda\varphi} & -ie^{i\Lambda\varphi} \end{pmatrix} \tag{3.72}$$

It is readily shown that

$$\mathcal{A}^{\dagger} \frac{d}{d\varphi} \mathcal{A} = \frac{d}{d\varphi} + i\Lambda\sigma_y \tag{3.73}$$

$$\mathcal{A}^{\dagger}\sigma_z\mathcal{A} = -\sigma_y \tag{3.74a}$$

$$\mathcal{A}^{\dagger}\begin{bmatrix} 0 & e^{-i2\Lambda\varphi} \\ e^{i2\Lambda\varphi} & 0 \end{bmatrix}\mathcal{A} = \sigma_z \tag{3.74b}$$

where σ_y is the Pauli matrix $\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$. Thus the Born-Renner quasi-Hamiltonian is

$$\begin{aligned}\mathcal{A}(K + \mathcal{T} + \mathcal{U})\mathcal{A} &= K - \frac{\hbar^2}{2M}\frac{1}{r^2}\left[2\Lambda i\sigma_y\frac{d}{d\varphi} - \Lambda^2\right] \\ &+ \frac{\hbar^2}{2M}\left[\bar{h}_0 + \sigma_z r^{2\Lambda}\bar{j}_0(r^2) - 2i\sigma_y\bar{\bar{h}}_2\frac{d}{d\varphi} - 2\Lambda\bar{\bar{h}}_2\right] \\ &+ \bar{f} + \sigma_z r^{2\Lambda}\bar{g}\end{aligned} \tag{3.75}$$

In Renner's paper [Ref.] the term

$$\frac{\hbar^2}{2M}\left[\bar{h}_0 + \sigma_z r^{2\Lambda}\bar{j}_0(r^2) - 2i\sigma_y\bar{\bar{h}}_2\frac{d}{d\varphi} - 2\Lambda\bar{\bar{h}}_2\right]$$

was ignored at the outset. Furthermore, the functions $\bar{f}$ and $r^{2\Lambda}\bar{g}$ were cut off after second order in r. Then he found

$$\bar{f} \doteq \mathcal{E}(X_0) + \alpha r^2 \tag{3.76}$$

$$r^{2\Lambda}g \doteq \begin{cases} \beta r^2 & \Lambda = 1 \\ 0 & \Lambda > 1 \end{cases} \tag{3.77}$$

where α and β were real constants. The theory we have presented here therefore shows in a very direct manner why Renner only considered the Π-electronic states. For $\Lambda > 1$, and assuming Renner's approximations, the

Born-Renner quasi-Hamiltonian may be transformed into the Moffitt-Liehr quasi-Hamiltonian

$$K + \mathcal{E}(X_o) + \alpha r^2$$

whose eigenvalues are just those obtained by the Born-Oppenheimer approximation.

The invariants of the quasi-Hamiltonian (3.75) can be readily determined by use of equations (3.45). We find that $\frac{d}{d\varphi}$, $C\sigma_z$, and (*) all commute with this operator. The fact that the complex conjugation operator is an invariant simply implies that the Born-Renner quasi-Hamiltonian is real which is readily verified from an inspection of (3.75).

F. The Case of Infinite P and $\ell = 2$.

A second example which we shall consider is that where $P = \infty$ and $\ell = 2$. Again we assume that there is a reflection plane passing through the symmetry axis such that the operator C may be defined and commutes with H_X. There are no actual molecules which correspond to this example but Longuet-Higgins et al. [Ref. 6] has constructed a theoretical model for which this is the case. The Moffitt-Liehr quasi-Hamiltonian in this case turns out to be of the form

$$K + \frac{\hbar^2}{2M}\Big[\bar{h}_o(r^2) + r^{\Lambda}\bar{j}_o(r^2) \begin{pmatrix} 0 & e^{-i\Lambda\varphi} \\ e^{i\Lambda\varphi} & 0 \end{pmatrix} + 2i\sigma_z\bar{\bar{h}}_2(r^2)\frac{d}{d\varphi}\Big] + \bar{f}(r^2) + r^{\Lambda}\bar{g}(r^2) \begin{pmatrix} 0 & e^{-i\Lambda\varphi} \\ e^{i\Lambda\varphi} & 0 \end{pmatrix} \tag{3.78}$$

and we may take the matrix $\mathcal{A}$ to be

$$\frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i(\Lambda/2)\varphi} & ie^{-i(\Lambda/2)\varphi} \\ e^{i(\Lambda/2)\varphi} & -ie^{i(\Lambda/2)\varphi} \end{pmatrix} \tag{3.79}$$

giving a Born-Renner quasi-Hamiltonian of the form

$$K - \frac{\hbar^2}{2M}\frac{1}{r^2}\Big[\sigma_y\Lambda i\frac{d}{d\varphi} - (1/4)\Lambda^2\Big] + \frac{\hbar^2}{2M}\Big[\bar{h}_o + \sigma_z r^{\Lambda}\bar{j}_o - i\sigma_y 2\bar{\bar{h}}_2\frac{d}{d\varphi} - \Lambda\bar{\bar{h}}_2\Big] + \bar{f}(r^2) + \sigma_z\bar{g}(r^2)r^{\Lambda} \tag{3.80}$$

Here as before the functions $\bar{h}_o$, $\bar{j}_o$, $\bar{\bar{h}}_2$, $\bar{f}$, and $\bar{g}$ are all real analytic functions of r^2.

We note that if Λ is even our vibronic Hamiltonians (3.78) and (3.80) are identical in form to (3.71) or (3.75), respectively, with a value of Λ twice that in (3.71) or (3.75). If Λ is odd, however, there is no such correspondince. An interesting feature which then arises out of choosing $\mathcal{A}$ in the form (3.79) is that the functions

$$\varphi^{I} = \frac{1}{2}\left[e^{-i(\Lambda/2)\varphi}\bar{\Phi}_{\Lambda} + e^{i(\Lambda/2)\varphi}\bar{\Phi}_{-\Lambda}\right] \qquad (3.81a)$$

$$\Phi^{II} = \frac{1}{2}\left[e^{-i(\Lambda/2)\varphi}\bar{\Phi}_{\Lambda} - e^{i(\Lambda/2)\varphi}\bar{\Phi}_{-\Lambda}\right] \qquad (3.81b)$$

go into their negatives on rotation of φ through 2π rather than into themselves if Λ is odd. Since the total wavefunction

$$\Psi = \psi^{I}\,\varphi^{I} + \psi^{II}\,\varphi^{II} \qquad (3.82)$$

must be invariant under rotations of φ through 2π we must impose the boundary condition on the eigenvectors of (3.80) that they transform into their negatives, i. e.,

$$\begin{pmatrix} \psi^{I}(r, \varphi + 2\pi) \\ \psi^{II}(r, \varphi + 2\pi) \end{pmatrix} = - \begin{pmatrix} \psi^{I}(r, \varphi) \\ \psi^{II}(r, \varphi) \end{pmatrix} \qquad (3.83)$$

if Λ is odd. This circumstance, however, could have been avoided if we had chosen $\mathcal{A}$ to be of the form

$$= \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\Lambda\varphi} & ie^{-i\Lambda\varphi} \\ 1 & -i \end{pmatrix} \qquad (3.84)$$

Then the Born-Renner quasi-Hamiltonian would have been of the form

$$K - \frac{\hbar^2}{2M}\,\frac{1}{r^2}\,(\sigma_y - 1)\,\Lambda\left(\frac{\Lambda}{2} + i\,\frac{d}{d\varphi}\right)$$
$$+ \frac{\hbar^2}{2M}\left[\bar{h}_o + r^{\Lambda}\bar{j}_o\sigma_z + \bar{\bar{h}}_2\Lambda(1 - \sigma_y) - 2i\sigma_y\bar{\bar{h}}_2\,\frac{d}{d\varphi}\right]$$
$$+ \bar{f} + r^{\Lambda}\bar{g}\,\sigma_z \qquad (3.85)$$

The invariants of the Born-Renner quasi-Hamiltonian (3.80) are the same as those of (3.75) - namely, $\frac{d}{d\varphi}$, $C\sigma_z$, and $(*)$. The invariants of (3.85) are different, however, and are $d/d\varphi$, $e^{i\Lambda\varphi}(*)$, and $e^{i\Lambda\varphi}\sigma_z C$.

G. The Case of $P = 6, \ell = 2, \Lambda = 1$.

Examples in this category have considered by Moffitt and Liehr [Ref. 5] and by Moffitt and Thorson [Ref. 32]. The Moffitt-Liehr quasi-Hamiltonian for this class of problems is somewhat more complicated than those of the preceding two cases and we shall in the interest of simplicity neglect the matrix $\mathbb{T}$ in our formulae. Extending our analysis to include the terms resulting from this matrix is straightforward and can readily be done using the same methods used to derive the form of $\mathcal{U}$. It does not, however, seem worthwhile in this instance as our analysis in Appendix C indicates that we may neglect $\mathbb{T}$ to a good approximation. As in the preceding two cases we shall assume that the molecular configuration has the additional symmetry element C.

The matrix $\mathcal{U}$ for this case is of the following form:

$$\mathcal{U} = \sum_{m=0}^{\infty} \bar{f}_m(r^2)\ r^{3m}\cos 3m\varphi$$

$$+ \sum_{m=-\infty}^{\infty} r^{|3m+1|}\,\bar{g}_m(r^2) \begin{pmatrix} 0 & e^{-i(3m+1)\varphi} \\ e^{i(3m+1)\varphi} & 0 \end{pmatrix} \qquad (3.86)$$

where the functions $\bar{f}_m$ and $\bar{g}_m$ are real analytic functions of r^2. Unlike the previous two cases considered the matrix $\mathcal{A}$ which diagonalizes $\mathcal{U}$ will here be in general a function of r as well of the angular coordinate φ. This gives $\mathcal{A}$ a somewhat complicated form and makes the process of obtaining the Born-Renner quasi-Hamiltonian from the Moffitt-Liehr quasi-Hamiltonian no longer trivial. To make our analysis simpler we shall take $\mathcal{U}$ in the harmonic approximation and neglect terms of order r^3 and higher. Then $\mathcal{U}$ is of the form

$$\mathcal{U} = Ar^2 + Br \begin{pmatrix} 0 & e^{-i\varphi} \\ e^{i\varphi} & 0 \end{pmatrix} + Cr^2 \begin{pmatrix} 0 & e^{i2\varphi} \\ e^{-i2\varphi} & 0 \end{pmatrix} \qquad (3.87)$$

where A, B, and C are constants. This approximation gives the eigenvalues of $\mathcal{U}$ as

$$\mathcal{E}^{I,II} = Ar^2 \pm (Br + Cr^2\cos 3\varphi) \qquad (3.88)$$

to the same order of approximation. The matrix $\mathcal{A}$

may then be taken as

$$\mathcal{A} = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i(\varphi/2)} & ie^{\pm i(\varphi/2)} \\ e^{i(\varphi/2)} & -ie^{i(\varphi/2)} \end{pmatrix} \begin{pmatrix} 1 & -\frac{C}{2B} r \sin 3\varphi \\ \frac{C}{2B} r \sin 3\varphi & 1 \end{pmatrix} \tag{3.89}$$

giving a Born-Renner quasi-Hamiltonian of the form

$$\begin{aligned} &K + (Br + Cr^2 \cos 3\varphi)\sigma_z + Ar^2 \\ &\quad + \frac{\hbar^2}{2M} \Big[\frac{2i\sigma_y C}{2B} \sin 3\varphi \frac{d}{dr} + \frac{1}{4r^2} - \frac{i\sigma_y}{r^2} \frac{d}{d\varphi} \\ &\qquad - \frac{3C \cos 3\varphi}{2Br}\Big(1 - 2i\sigma_y \frac{d}{d\varphi}\Big) - \frac{8i\sigma_y C \sin 3\varphi}{2Br}\Big] \end{aligned} \tag{3.90}$$

Both of these expressions are approximate - although consistent with approximating $\mathcal{U}$ by (3.87). If we took the constant C to be zero expression (3.90) would be but a special case of (3.80) with $\Lambda = 1$. The most interesting feature of C being non-zero is that the Born-Renner quasi-Hamiltonian is not invariant under arbitrary rotations of φ but instead only under rotations which change φ by multiples of $2\pi/3$. Other invariants of (3.90) are $(*)$ - as the quasi-Hamiltonian is real - and $\sigma_z C$. (C here denotes the operation of changing φ into its negative and is not to be confused with the constant C appearing in (3.90)). The eigenfunctions of (3.90) must also satisfy the boundary condition (3.83).

H. Possible Forms of the Moffitt-Liehr Quasi-Hamiltonian in the Harmonic Approximation.

An analysis along the lines of the preceding sections shows that when one considers all possible combinations of P, ℓ, and Λ he finds that there are essentially only four non-trivially different forms of the matrix $\mathcal{U}$ in the harmonic approximation. This relatively small number is due to the fact that the equations (3.47) along with the restrictions $\ell \neq 0$, $2\ell \neq 0$, $\Lambda \neq 0$, $2\Lambda \neq 0 (\mathrm{Mod}\ P)$ greatly limit the possible forms of the elements $f(r, \varphi)$ and $g(r, \varphi)$ of $\mathcal{U}$. The diagonal elements $f(r, \varphi)$ can only be of the form

$$\mathcal{E}(X_o) \quad + Ar^2$$

in the harmonic approximation while the possible forms of $g(r, \varphi)$ are as follows

A) $g(r, \varphi) = C\ r\ e^{-i\varphi} \quad + D\ r^2 e^{2i\varphi}$ (3.91a)

B) $g(r, \varphi) = C\ r\ e^{-i\varphi}$ (3.91b)

C) $g(r, \varphi) = C\ r^2 e^{-2i\varphi} + D\ r^2 e^{2i\varphi}$ (3.91c)

D) $g(r, \varphi) = C\ r^2 e^{-2i\varphi} \quad =$ (3.91d)

The complex conjugates of these expressions are also possibilities but it is clear that the resulting forms for $\mathcal{U}$ will only differ by a similarity transformation from the forms of ω obtained by considering the above.

Case A occurs if $\ell = 2\Lambda$(Mod P) and $2\ell = -2\Lambda$ (ModPP). It occurs only if $3\ell = 0$(Mod P) and therefore if the symmetry axis is three-fold, six-fold, etc.. Case B occurs only if $\ell \equiv 2\Lambda$(Mod P) and $2\ell \neq -2\Lambda$ (Mod P). It cannot occur if $3\ell \equiv 0$ and therefore cannot occur in molecules having trigonal symmetry axes. For $\ell = 1$ Case B can never occur and Case A can only occur if the symmetry axis is trigonal.

Case C occurs if $2\Lambda = 2\ell$ (Mod P), $2\Lambda = -2\ell$ (Mod P). It can only occur if $4\ell = 0$ and $4\Lambda = 0$ (Mod P), i. e., if the symmetry axis is four-fold, eight-fold, etc.. Case D occurs if $2\ell = 2\Lambda$ and $2\ell \neq -2\Lambda$(Mod P). It cannot occur if $4\ell \equiv 0$ or $4\Lambda \equiv 0$ and, in particular, if the symmetry axis is four-fold.

We must remind ourselves that any of the constants C and D appearing in any of the expressions (3.91) may vanish for reasons other than considered here. Our symmetry considerations simply indicate that we have not found any reason for their being zero from considering the properties mentioned in Theorems I and II. We have previously mentioned that the complex conjugates of expressions (3.91) are also possibilities. The remarks in the preceding paragraphs also apply to these possibilities - provided Λ is everywhere replaced by $-\Lambda$.

Let us now write down the four Moffitt-Liehr quasi-Hamiltonians corresponding to (3.91):

A)

$$K + \mathcal{E}(X_o) + Ar^2 + r\begin{pmatrix} 0 & Ce^{-i\varphi} \\ C^{*}e^{i\varphi} & 0 \end{pmatrix} + r^2\begin{pmatrix} 0 & De^{2i\varphi} \\ D^{*}e^{-2i\varphi} & 0 \end{pmatrix} \qquad (3.92a)$$

B)

$$K + \mathcal{E}(X_o) + Ar^2 + r\begin{pmatrix} 0 & Ce^{-i\varphi} \\ C^{*}e^{i\varphi} & 0 \end{pmatrix} \qquad (3.92b)$$

C)

$$K + \mathcal{E}(X_o) + Ar^2 + r^2\begin{pmatrix} 0 & Ce^{-2i\varphi} + De^{2i\varphi} \\ C^{*}e^{2i\varphi} + D^{*}e^{-2i\varphi} & 0 \end{pmatrix} \qquad (3.92c)$$

D)

$$K + \mathcal{E}(X_o) + Ar^2 + r^2\begin{pmatrix} 0 & Ce^{-2i\varphi} \\ C^{*}e^{2i\varphi} & 0 \end{pmatrix} \qquad (3.92d)$$

Here we have neglected the matrix $\mathcal{T}$ as our analysis in Appendix C shows that this is consistent with the harmonic approximation.

The constants C and D will be real if the configuration X_o has its symmetry axis lying in a reflection plane or a two-fold axis perpendicular to the symmetry axis.

The quasi-Hamiltonians (3.92b) and (3.92d) may be considered as particular instances of (3.78) and (3.71), respectively (even though they may have been obtained for cases in which the values of and P were different than those assumed in deriving (3.78) and (3.71)). The quasi-Hamiltonian (3.92b) is just that considered in the previous section. The only case which has not been included in our discussion up to this point is Case C. Equation (3.92c) has never, to our knowledge, appeared in the literature before. The reason for its neglect is probably due to the fact that molecules with four-fold symmetry axes are apparently very scarce in nature and also due to the consequences of the Jahn-Teller theorem which indicates that electronic degeneracies of molecules with four-fold symmetry axes will be split to first order in the vibrational coordinates by non-degenerate vibrations. Equation (3.92c) only indicates a splitting of second order. Nevertheless, one cannot exclude the possibility of physical problems for which a consideration of (3.92c) may be important and where it may describe the physical situation to a good approximation.

I. Extension of the Theory to Include More Vibrational Degrees of Freedom.

It is a relatively simple matter to extend the remarks make in the preceding sections to include more than two vibrational degrees of freedom. This follows from the fact that our two theorems are easily extended for the case of an arbitrary number of modes.

Let us denote the vibrational coordinates by $r_1, \varphi_1, \ldots, r_k, \varphi_k, \ldots$ and $\xi_1, \ldots, \xi_s, \ldots,$ where the degenerate vibrational[1] modes are distinguished by the index k and the non-degenerate modes by the index s. The coordinates associated with the degenerate modes can be chosen such that

$$R_\alpha r_k = r_k \tag{3.93a}$$

$$C r_k = r_k \tag{3.93b}$$

$$R_\alpha \varphi_k = \varphi_k + \ell_k \alpha \tag{3.93c}$$

$$C \varphi_k = - \varphi_k \tag{3.93d}$$

where ℓ_k is an integer characteristic of the particular vibrational mode. The symmetry properties of the non-

---

[1] We limit ourselves here to only two-fold degenerate modes and non-degenerate modes. If any three-fold degenerate modes should exist they may be considered as a two-fold degenerate vibration accidentally degenerate with a non-degenerate vibration. This follows as the subgroup generated by R_α and C has no three dimensional irreducible representations.

degenerate modes are even simpler. For each of the operations R_α and C there are only two possibilities - the coordinates either remain invariant or are transformed into their negatives

$$C\,\xi_s = \pm\,\xi_s \tag{3.94a}$$

$$R_\alpha\,\xi_s = \pm\,\xi_s \tag{3.94b}$$

The minus sign in (3.94b) can only occur if P is finite and even.

The stating of Theorem I is relatively unchanged by the inclusion of an arbitrary number of vibrations. The functions $\bar{\Phi}_\Lambda$, $U_{\Lambda\Lambda}$, etc., are now functions of all the vibrational coordinates. Equations (3.14) are modified to stating that $\bar{\Phi}_\Lambda$ reduces to $\bar{\Phi}_\Lambda^{(0)}$, $U_{\Lambda\Lambda}$ reduces to $\mathcal{E}(X_0)$, etc., when all the radial coordinates r_k and the ξ_s are set equal to zero. Equations (3.16) and (3.17) remain the same. Statement 3 must be modified to state that the $\bar{\Phi}_\Lambda$, $U_{\Lambda\Lambda}$, etc., are analytic in the $r_k \cos\varphi_k$, $r_k \sin\varphi_k$, and the ξ_s.

Equations (3.18) and (3.19) remain the same in the stating of Theorem II. Similarly equations (3.20), (3.21), (3.22), and (3.23) remain the same - providing we understand the operations R_α and C to be applied to each of the φ_k and ξ_s simultaneously.

The proof of this generalization of the two theorems proceeds just as outlined in Appendix A for the special case of one two-fold degenerate vibration.

The relations contained in the restating of the theorems make it possible to determine the form of the Moffitt-Liehr quasi-Hamiltonian for any given case. As an example[1] we consider a molecule consisting of four identical atoms whose "equilibrium configuration" X_o is that of a square (symmetry C_{4v}). For simplicity, we consider only the planar vibrational modes - of which there are five. There is one two-fold degenerate mode ($\ell = 1$) and three non-degenerate modes. The latter may be chosen such that (C is taken as a reflection through the line dividing the square into two identical rectangles)

$$R_\alpha \xi_1 = \xi_1 \qquad C \xi_1 = \xi_1 \tag{3.95a}$$

$$R_\alpha \xi_2 = -\xi_2 \qquad C \xi_2 = \xi_2 \tag{3.95b}$$

$$R_\alpha \xi_3 = -\xi_3 \qquad C \xi_3 = -\xi_3 \tag{3.95c}$$

where α denotes the angle $\pi/2$. We take $\Lambda = 1$ for our degenerate electronic state. The application of our two

---

[1] We have in mind the molecule C_4H_4 (cyclobutadiente). While it is not completely certain that this molecule exists in a planar form, calculations have been recently carried out by Hobey and McLachlan [Ref. 33] which indicate that if planar its equilibrium configuration is very nearly square and that its ground electronic state is degenerate at the square configuration.

theorems then indicates that elements of the matrix must be of the forms

$$U_{\Lambda\Lambda} = \mathcal{E}(X_0) + Ar^2 + \alpha_1 \xi_1^2 + \alpha_2 \xi_2^2 + \alpha_3 \xi_3^2 + Q \xi_1 \tag{3.96a}$$

$$U_{\Lambda,-\Lambda} = Br^2 e^{i2\varphi} + Cr^2 e^{-i2\varphi} + D\xi_2 + E\xi_1\xi_2 + iF\xi_3 + iG\xi_1\xi_3 \tag{3.96b}$$

where we have neglected cubic and higher order terms. The constant Q may be made to vanish by adjusting the length of the sides of the square (doing this will not lift the degeneracy). All the constants appearing in (3.96) are, furthermore, entirely real. The Moffitt-Liehr quasi-Hamiltonian is therefore of the form

$$K + \mathcal{E}(X_0) + \alpha_1 \xi_1^2 + \alpha_2 \xi_2^2 + \alpha_3 \xi_3^2 + Ar^2$$
$$+ D\xi_2\sigma_1 + E\xi_1\xi_2\sigma_1 - F\xi_3\sigma_2 - G\xi_1\xi_3\sigma_2$$
$$+ r^2 \begin{pmatrix} 0 & Be^{i2\varphi} + Ce^{-i2\varphi} \\ Be^{-i2\varphi} + Ce^{i2\varphi} & 0 \end{pmatrix} \tag{3.97}$$

where we have again neglected the matrix $\mathcal{T}$. The presence of the off diagonal terms $D\xi_2\sigma_1$ and $F\xi_3\sigma_2$ is a confirmation of the Jahn-Teller theorem. This very model, in fact, is used as an example by Sponer and Teller [Ref. 15] in their discussion of the Jahn-Teller theorem.

It is not necessary to express the quasi-Hamiltonian for a molecule in terms of vibrational coordinates. One could express it in terms of the position coordinates of the nuclei. Doing this impairs the generality of our treatment as one must consider specific molecules. However, we shall have occasion to make use of a vibronic quasi-Hamiltonian in this form when we generalize our considerations to solids.

Let us therefore obtain the quasi-Hamiltonian for the following model. We consider a hypothetical molecule consisting of g identical atoms constrained to move on the perimeter of a circle of circumference ga. The "equilibrium configuration" X_o is taken to be one in which the atoms are equally spaced a distance $\underline{a}$ apart around the perimeter. The Moffitt-Liehr quasi-Hamiltonian for such a system in a two-fold degenerate electronic state with symmetry index Λ may be taken to be of the form

$$K + \mathcal{U}$$

where

$$K = -\frac{\hbar^2}{2M} \sum \frac{d^2}{dX_n^2} \tag{3.98}$$

and $\mathcal{U}$ is given in the form (3.27). The matrix $\mathbb{T}$ has been neglected. The X_n represent the positions of the nuclei relative to their a-priori assigned equilibrium positions along the circumference of the circle.

The elements of the matrix U have the following properties which follow from our generalization of **Theorems** I and II:

$$U_{\Lambda\Lambda}(R_\alpha X_1,\ldots\ldots, R_\alpha X_g) = U_{\Lambda\Lambda}(X_1,\ldots\ldots,X_g)$$
$$= U_{-\Lambda,-\Lambda}(X_1,X_2,\ .\ .\ .\ .\ ., X_g) \qquad (3.99a)$$

$$U_{\Lambda,-\Lambda}(R_\alpha X_1,\ .\ .\ .,R_\alpha X_g) = e^{-i(4\pi\Lambda/g)}U_{\Lambda,-\Lambda}(X_1,.\ .\ .,X_g) \qquad (3.99b)$$

$$U_{\Lambda\Lambda}(CX_1,\ .\ .\ .,\ CX_g) = U^*_{\Lambda\Lambda}(X_1,\ .\ .\ .,X_g) \qquad (3.99c)$$

$$U_{\Lambda,-\Lambda}(CX_1,\ .\ .\ .,\ CX_g) = U_{-\Lambda,\Lambda}(X_1,\ .\ .\ .,X_g)$$
$$= U^*_{\Lambda,-\Lambda}(X_1,\ .\ .\ .,X_g) \qquad (3.99d)$$

In addition, these elements are analytic in each of the coordinates $X_1,.\ .\ .\ .,X_g$.

In the above R_α represents a permuation operator which replaces the displacement of an atom by that of the atom appearing immediately to its left:

$$R_\alpha X_n = X_{n-1} \qquad (3.100a)$$
$$R_\alpha X_1 = X_g \qquad (3.100b)$$

The operator C similarly denotes a permutation in which the displacement of the n-th atom is replaced by the negative of that of the (g-n+1)-th atom:

$$CX_n = -X_{g-n+1}$$

The invariants of the quasi-Hamiltonian are thus

$R_\alpha e^{i\frac{2\pi}{g}\Lambda\sigma_z}$, $(*)\sigma_x$, and $C\sigma_x$ as was true for the Moffitt-Liehr quasi-Hamiltonians studied previously.

An additional requirement is imposed by the fact that our quasi-Hamiltonian should be invariant under a simultaneous rotation of all the atoms about the perimeter of the circle, Also, we assume that only neighboring atoms interact.

With these considerations in mind it is a simple matter to write down a general form of the Moffitt-Liehr quasi-Hamiltonian in the harmonic approximation:

$$\begin{aligned} \mathcal{H} &= K + \mathcal{E}(X_o) + \sum \omega_o^2 (X_n - X_{n-1})^2 \\ &+ A \sum X_n \left[\sigma_x \sin \frac{2\pi\Lambda}{g}(2n-1) + \sigma_y \cos \frac{2\pi\Lambda}{g}(2n-1)\right] \\ &+ B \sum (X_n - X_{n-1})^2 \left[\sigma_x \cos \frac{2\pi\Lambda}{g}(2n-2) + \sigma_y \sin \frac{2\pi\Lambda}{g}(2n-2)\right] \end{aligned} \tag{3.101}$$

where A, B, and ω_o^2 are real constants.

The model for which this quasi-Hamiltonian was obtained is probably most familiar to workers in solid state theory as it is that of a one-dimensional monatomic lattice subject to periodic boundary conditions. If the constants A and B are set equal to zero it reduces to the well-known Born-Oppenheimer quasi-Hamiltonians discussed in practically every elementary textbook on solid state physics.

J. Clarification of the Terminology: Dynamic Jahn-Teller Effect and Renner Effect.

The study of the solutions of the Schroedinger equation for a molecule for cases in which the Born-Oppenheimer approximation is inapplicable is a relatively recent development - the only paper on the subject before 1957 being that of Renner. The writers of these recent papers have tended to categorize the class of problems considered in two categories - the terminology dynamic Jahn-Teller effect and Renner effect being associated with the two different categories. Thus a recent paper by Hobey and McLachlan [Ref. 33] was titled "Dynamic Jahn-Teller Effect in Hydrocarbon Molecules,"[1] while a previous paper by Pople and Longuet-Higgins was titled "Renner Effect in the NH_2 Molecule."

This terminology is somewhat confusing as there is apparently no unique effect associated with a breakdown of the Born-Oppenheimer approximation other than an anomalous vibrational spectra not explainable by the usual theories based on a harmonic oscillator potential. The universal trend, however, seems to be to lump every-

---

[1] Actually, the title is misleading, as the calculations contained therein were concerned with the static problem - although they indicated that their calculations showed that a consideration of the dynamic problem was necessary.

thing together which is somewhat affected by the consequences of the Jahn-Teller theorem as Jahn-Teller effects and to classify all effects arising from a break-down of the Born-Oppenheimer approximation in linear molecules as Renner effects.

The former terminology has been recently criticized by Liehr [Ref. 7]:

> Scientific usage has reserved the word "effect" for phenomena which are capable of unambiguous experimental verification, and in such case has, at times, attached to it a modifier - the name of the investigator(s) who were first to observe or explain the phenomena in question (witness the Raman effect). Frankly, to attempt to append a "name" to a natural phenomenon is puerile. At present we have at our disposal all that is needed to describe the phenomena which attend inherent configurational instabilities: we have (a) a theorem due to Jahn and Teller which yields reliable criteria for the possible observance of such effects and (b) extremely competent experimental and theoretical scientists who are capable of ascertaining when these criteria are met, and whether the expected consequences are actually achieved. However, it really is convenient...to be able to tag a label onto phenomena reasonably imputable to the Jahn-Teller theorem. It is therefore recommended that the appelation "Jahn-Teller effect" be restricted to the dynamical manifestations of Jahn and Teller's theorem (e.g., anomalous series of asymmetric vibrational modes and abnormal paramagnetic demeanor) and that the statical demonstrations of this lemma be charged to "intrinsic Jahn-Teller instability."

The present use of the terminology, however, will probably remain in vogue.

The distinction between dynamic Jahn-Teller effects and Renner effects as used implicitly in the literature apparently applies to the type of molecule considered - if one studies vibration-electron interaction in non-linear molecules he speaks of Jahn-Teller effects while if he studies vibration-electron interaction in linear molecules he speaks of Renner effects. Liehr [Ref. 35] has suggested relating the distinction to the form of the Moffitt-Liehr quasi-Hamiltonian. If the lowest order term of the off diagonal term g in the matrix $\mathcal{U}$ is odd he suggests the effects of this term be called "Jahn-Teller effects". If, on the other hand, the lowest order term is even in the vibrational coordinates he suggests that one speak of Renner effects. Symmetry considerations such as we have given previously (see equation (3.71)) show that the leading term in g for linear molecules is always even. Furthermore, the Jahn-Teller theorem implies that if all vibrational coordinates are considered in a nonlinear molecule the leading term in g will be linear in at least one vibrational coordinate (see equation (3.97)).

Nevertheless, in non-linear molecules it may happen that the effects of the interaction of the electronic motion with certain vibrations may be greater than others - even though these vibrations may not remove the degeneracy in first order - but in a higher order.

K. Summary.

Thus, in this chapter we have shown how, for the case of the interaction of a two-fold degenerate vibration, the form of the quasi-Hamiltonian is determined by a knowledge of the angular momentum quantum numbers Λ and ℓ. This general model contains both the cases of linear molecules and non-linear molecules; and we see that the Renner model and the dynamic Jahn-Teller model are but special cases of this model obtained by making various assignments of Λ and ℓ.

Chapter 4

PROPERTIES OF THE SOLUTIONS OF THE VIBRONIC EQUATIONS

In this chapter we shall study the quasi-Hamiltonians obtained in the previous chapter and shall examine the properties of their eigenfunctions and eigenvalues. Rather than perform a detailed calculation of the energy spectrum of any one quasi-Hamiltonian we shall make extensive use of symmetry considerations to find the general forms of the eigenfunctions and the possible degeneracies of the energy levels. In the latter part of the chapter we shall solve the vibronic equations in two limiting cases where perturbation theory is applicable.

A. Degeneracy of Levels of Born-Renner and Moffitt-Liehr Quasi-Hamiltonians.

The eigenvalues of the Moffitt-Liehr (or Born-Renner) quasi-Hamiltonians may be classified according to the irreducible representations of the symmetry group generated by the invariants of the quasi-Hamiltonian. In this section we shall show how this is done and shall, moreover, make use of symmetry considerations to derive certain properties of the solutions of the vibronic equations and to determine the form of the eigenfunctions in certain particular cases.

Let us begin by considering the Moffitt-Liehr quasi-Hamiltonian (3.78):

$$K + \frac{\hbar^2}{2M}\Big[\bar{h}_0(r^2) + r^{\Lambda}\bar{j}_0(r^2)\begin{pmatrix} 0 & e^{-i\Lambda\varphi} \\ e^{i\Lambda\varphi} & 0 \end{pmatrix} + 2i\sigma_z\bar{\bar{h}}_2(r^2)\frac{d}{d\varphi}\Big] + \bar{f}(r^2) + r^{\Lambda}\bar{g}(r^2)\begin{pmatrix} 0 & e^{-i\Lambda\varphi} \\ e^{i\Lambda\varphi} & 0 \end{pmatrix} \qquad (3.78)$$

The consideration of this quasi-Hamiltonian will enable us to simultaneously consider the quasi-Hamiltonians (3.71), (3.92b) and (3.92d) - as each of these may be regarded as special cases of (3.78) which maybe obtained from it by making various identifications of Λ, $\bar{f}(r^2)$, etc.

The invariants of this quasi-Hamiltonian are $(*)\sigma_x$, $C\sigma_x$, and

$$e^{i\alpha[-2i(d/d\varphi) + \Lambda\sigma_z]}$$

(for arbitrary α - as the "molecule" has cylindrical symmetry). In particular, the operator

$$\mathcal{J} = -2i(d/d\varphi) + \Lambda\,\sigma_z \tag{4.1}$$

commutes with (3.78). From these operators it is also possible to construct the operator

$$\tau = \mathcal{R}_{\pi/\Lambda}\; i\,\sigma_y\,(*) = \tag{4.2}$$

which also commutes with (3.78). Here $\mathcal{R}_{\pi/\Lambda}$ is an operator which sends φ into $\varphi + \pi/\Lambda$. That τ is an invariant follows from the fact that

$$i\,e^{-i(\pi/2\Lambda)[-2i(d/d\varphi) + \Lambda\sigma_z]}[(*)\sigma_x]$$

$$= \mathcal{R}_{\pi/\Lambda}\,\sigma_z(*)\sigma_x$$

$$= \mathcal{R}_{\pi/\Lambda}\; i\,\sigma_y\,(*) \tag{4.3}$$

We note that

$$\tau^{2\Lambda} = (-1)^{\Lambda} \tag{4.4}$$

for

$$\tau^{2\Lambda} = \mathcal{R}_{2\pi}(i\sigma_y)^{2\Lambda}(*)^{2\Lambda}$$

$$= (-1)^{\Lambda}\,\sigma_y^{2\Lambda}$$

$$= (-1)^{\Lambda}$$

as $\mathcal{R}_{2\pi}$, $(*)^{2\Lambda}$ and $\sigma_y^{2\Lambda}$ are each the identity operator. From this fact it follows that *the eigenvalues of (3.78) are all at least two-fold degenerate if Λ is odd*.

The proof of this statement is similar to Klein's proof [Ref. 36] of Kramer's theorem concerning time reversal degeneracy. We assume $\overset{\rightharpoonup}{\overline{\psi}}$ is an eigenvector of (3.78). If the eigenvalue associated with $\overset{\rightharpoonup}{\overline{\psi}}$ is non-degenerate then $\overset{\rightharpoonup}{\overline{\psi}}$ and $\tau\overset{\rightharpoonup}{\overline{\psi}}$ must be linearly dependent, i. e.,

$$\tau\overset{\rightharpoonup}{\overline{\psi}} = a\overset{\rightharpoonup}{\overline{\psi}}$$

since τ commutes with (3.78). Then

$$\begin{aligned}\tau^2\,\overset{\rightharpoonup}{\overline{\psi}} &= a^{*}\tau\,\overset{\rightharpoonup}{\overline{\psi}}\\ &= |a|^2\,\overset{\rightharpoonup}{\overline{\psi}}\end{aligned}$$

as

$$\tau\, a = a^{*}\tau.$$

Therefore

$$\begin{aligned}\tau^{2\Lambda}\overset{\rightharpoonup}{\overline{\psi}} &= |a|^{2\Lambda}\,\overset{\rightharpoonup}{\overline{\psi}}\\ &= (-1)^{\Lambda}\,\overset{\rightharpoonup}{\overline{\psi}}\end{aligned}$$

from (4.4). This is clearly a contradiction if Λ is odd - indicating that $\overset{\rightharpoonup}{\overline{\psi}}$ and $\tau\overset{\rightharpoonup}{\overline{\psi}}$ are linearly independent. (If Λ is even, however, we cannot make this conclusion.) This fact implies, in particular, that all the eigenvectors of (3.92b) are two-fold degenerate. This has also been brought out in the papers of Moffitt and Liehr [Ref. 5], Moffitt and Thomson [Ref. 32], and Longuet-Higgins et al. [Ref. 6].

Let us now determine the eigenvalues and eigenvectors of $\mathcal{J}$. The most general eigenvector we need consider is of the form

$$\psi = \sum_{m} e^{im\varphi} \begin{pmatrix} R_m^{(1)}(r) \\ R_m^{(2)}(r) \end{pmatrix} \tag{4.5}$$

as the eigenvectors of (3.78) must be periodic in φ . If we require

$$\mathcal{J} \vec{\Psi} = j \vec{\Psi} \tag{4.6}$$

we find we must require

$$2m \begin{pmatrix} R_m^{(1)} \\ R_m^{(2)} \end{pmatrix} + \Lambda \begin{pmatrix} R_m^{(1)} \\ -R_m^{(2)} \end{pmatrix} = j \begin{pmatrix} R_m^{(1)} \\ R_m^{(2)} \end{pmatrix} \tag{4.7}$$

for each value of m. This can be satisfied only if

(A) $2m + \Lambda = j$, $R_m^{(2)} = 0$

or

(B) $2m - \Lambda = j$, $R_m^{(1)} = 0$

or

(C) Both $R_m^{(1)} = 0$ and $R_m^{(2)} = 0$.

Thus the eigenvalues of $\mathcal{J}$ are even integers if Λ is even and odd integers if Λ is odd. The most general eigenvector of $\mathcal{J}$ associated with the eigenvalue j is furthermore of the form

$$\overset{\rightleftharpoons}{\Psi}_j = \begin{pmatrix} e^{i[(j-\Lambda)/2]\varphi} & 0 \\ 0 & e^{i[(j+\Lambda)/2]\varphi} \end{pmatrix} \begin{pmatrix} R_j^{(1)}(r) \\ R_j^{(2)}(r) \end{pmatrix} \qquad (4.8)$$

If we consider this as an eigenstate of (3.78) it must be degenerate with

$$C(*)\ \overset{\rightleftharpoons}{\Psi}_j = \begin{pmatrix} e^{i[(j-\Lambda)/2]\varphi} & 0 \\ 0 & e^{i[(j+\Lambda)/2]\varphi} \end{pmatrix} \begin{pmatrix} R_j^{(1)*}(r) \\ R_j^{(2)*}(r) \end{pmatrix} \qquad (4.9)$$

since $C(*)$ commutes with the quasi-Hamiltonian. This implies that we may choose $R_j^{(1)}$ and $R_j^{(2)}$ to both be real (which we shall do).

If we now apply the operator $(*)\sigma_x$ to the $\overset{\rightleftharpoons}{\Psi}_j$ given by (4.8) we find that

$$(*)\sigma_x\ \overset{\rightleftharpoons}{\Psi}_j = \begin{pmatrix} e^{-i[(j+\Lambda)/2]\varphi} & 0 \\ 0 & e^{-i[(j-\Lambda)/2]\varphi} \end{pmatrix} \begin{pmatrix} R_j^{(2)}(r) \\ R_j^{(1)}(r) \end{pmatrix} \qquad (4.10)$$

is an eigenfunction of (3.78) which is degenerate with the eigenfunction (4.8). This function, however, is an eigenstate of $\mathcal{J}$ with eigenvalue $-j$. We can therefore make the identifications

$$\overset{\rightleftharpoons}{\Psi}_{-j} = (*)\sigma_x\ \overset{\rightleftharpoons}{\Psi}_j \qquad (4.11)$$

$$R_{-j}^{(1)}(r) = R_{j}^{(2)}(r) \tag{4.12a}$$

$$R_{-j}^{(2)}(r) = R_{j}^{(1)}(r) \tag{4.12b}$$

Since the eigenfunctions given by (4.8) and (4.10) are linearly independent if $j \neq 0$ we see that the eigenvalues of (3.78) associated with non-zero values of j must be at least two-fold degenerate. This confirms our previous conclusions that the eigenvalues of (3.78) for odd Λ are at least two-fold degenerate.

If $j = 0$, however, our symmetry considerations allow the possibility of non-degenerate levels. This can only happen if Λ is even - as j even implies Λ even. Since $(\star)\sigma_x \overset{\rightrightarrows}{\Psi}_o$ is not necessarily linearly independent of $\overset{\rightrightarrows}{\Psi}_o$ we may choose $\overset{\rightrightarrows}{\Psi}_o$ to be an eigenstate of $(\star)\sigma_x$. Furthermore, since

$$[(\star)\sigma_x]^2 = 1$$

implies that the eigenvalues of $(\star)\sigma_x$ are ± 1, we have

$$(\star)\,\sigma_x\,\psi_o = \pm\,\psi_o$$

If the plus sign holds we speak of <u>gerade</u> eigenvectors and if the minus sign holds we speak of <u>ungerade</u> eigenvectors. The corresponding subscripts g and u will be used to denote gerade and ungerade eigenvectors. The most general gerade eigenvector ($j = 0$) of (3.78) is readily seen from (4.10) to be of the form

$$\vec{\vec{\Psi}}_{o,g} = \begin{pmatrix} e^{-i(\Lambda/2)\varphi} & 0 \\ 0 & e^{i(\Lambda/2)\varphi} \end{pmatrix} \begin{pmatrix} R_{o,g}(r) \\ R_{o,g}(r) \end{pmatrix} \qquad (4.13)$$

while the most general ungerade eigenvector is of the form

$$\vec{\vec{\Psi}}_{o,u} = \begin{pmatrix} e^{-i(\Lambda/2)\varphi} & 0 \\ 0 & e^{i(\Lambda/2)\varphi} \end{pmatrix} \begin{pmatrix} R_{o,u}(r) \\ -R_{o,u}(r) \end{pmatrix} \qquad (4.14)$$

The existence of these gerade and ungerade eigenvectors was discovered by Renner in his treatment of the π-states of CO_2 [Ref. 3] and also has been emphasized by Pople and Longuet-Higgins [Ref. 34]. These states correspond to the two one-dimensional representations of the group $C_{\infty v}$, while the states for $j \neq 0$ correspond to the two-dimensional representations.

The substitution of expressions (4.8), (4.13), and (4.14) into the Moffitt-Liehr vibronic equation whose quasi-Hamiltonian is (3.78) enables us to obtain the following differential equations for $R_j^{(1)}$, $R_j^{(2)}$, $R_{o,g}$, and $R_{o,u}$:

$$\Big[-\frac{\hbar^2}{2M} \Big(\frac{\partial^2}{dr^2} + \frac{1}{r} \frac{d}{dr} - \frac{j^2}{4r^2} = \frac{\Lambda^2}{4r^2} + \frac{\Lambda j}{2r^2} \sigma_z \Big) + \frac{\hbar^2}{2M} \big(\bar{h}_o(r^2) + r^{\Lambda} \bar{j}_o(r^2) \sigma_x \big) + (\Lambda - j\sigma_z) \frac{\hbar^2}{2M} \bar{\bar{h}}_2(r^2) + \bar{f}(r^2) + r^{\Lambda} \bar{g}(r^2) \sigma_x \Big] \begin{pmatrix} R_j^{(1)} \\ R_j^{(2)} \end{pmatrix} = E \begin{pmatrix} R_j^{(1)} \\ R_j^{(2)} \end{pmatrix} \qquad (4.15)$$

$$\Big[- \frac{\hbar^2}{2M} \Big(\frac{d^2}{dr^2} + \frac{1}{r} \frac{d}{dr} - \frac{\Lambda^2}{4r^2} \Big) + \frac{\hbar^2}{2M} \Big(\bar{h}_0(r^2) + r\Lambda \bar{j}_0(r^2) + \Lambda \bar{\bar{h}}_2(r^2) \Big) + \bar{f}(r^2) + r\Lambda \bar{g}(r^2) \Big] R_{0,g} = E\, R_{0,g} \tag{4.16}$$

$$\Big[- \frac{\hbar^2}{2M} \Big(\frac{d^2}{dr^2} + \frac{1}{r} \frac{d}{dr} - \frac{\Lambda^2}{4r^2} \Big) + \frac{\hbar^2}{2M} \Big(\bar{h}_0(r^2) - r\Lambda \bar{j}_0(r^2) + \Lambda \bar{\bar{h}}_2(r^2) \Big) + \bar{f}(r^2) - r\Lambda \bar{g}(r^2) \Big] R_{0,u} = E\, R_{0,u} \tag{4.17}$$

Although we have successfully eliminated the angular coordinate φ we are still left with two coupled equations for $R_j^{(1)}$ and $R_j^{(2)}$ when $j \neq 0$. The two equations may be uncoupled, however, when $j = 0$, giving the equations (4.16) and (4.17). If we were to neglect the effects of the matrix $\mathbb{T}$, i. e., the terms

$$\frac{\hbar^2}{2M} \Big(\bar{h}_0(r^2) \pm r\Lambda \bar{j}_0(r^2) + \Lambda \bar{\bar{h}}_2(r^2) \Big)$$

we would find that these represent the radial Sbhroedinger equations for the motion of the system on the potential surfaces

$$\bar{f}(r^2) + r\Lambda \bar{g}(r^2)$$

and

$$\bar{f}(r^2) - r\Lambda \bar{g}(r^2)$$

respectively, with an angular momentum of $\Lambda/2$. While we shall not be able to solve the equations (4.15-17) it is readily seen that the functions $R_j^{(1)}$ and $R_j^{(2)}$ must go to zero at least as rapidly as

$$r^{|(j - \Lambda)/2|}$$

and

$$r^{|j + \Lambda|/2}$$

respectively, while the functions $R_{o,g}$ and $R_{o,u}$ must go to zero at least as rapidly as

$$r^{\Lambda/2}$$

as r approaches zero. Thus the only eigenvectors which give a non-zero probability of the system being at r=0 are those for which $|j| = \Lambda$.

The remarks we have made above concerning the eigenvalues of (3.78) also apply to those of the Born-Renner quasi-Hamiltonian - as the eigenvalues of the two operators are identical. The analogue of the operator $\mathcal{J}$ for the Born-Renner quasi-Hamiltonian is

$$\mathcal{A}^{+}\mathcal{J}\mathcal{A} = -2i \frac{d}{d\varphi} \qquad (4.18)$$

where $\mathcal{A}$ is given by (3.79). The eigenvalues of this operator are the same as those of $\mathcal{J}$ and may be used to label the eigenvalues of (3.80). The two degenerate eigenstates associated with a level characterized by a particular value of j are

$$\mathcal{A}^{\dagger}\,\overset{\rightleftharpoons}{\Psi}_{j} = \frac{e^{ij(\varphi/2)}}{\sqrt{2}}\begin{pmatrix} R_j^{(1)}(r) + R_j^{(2)}(r) \\ -i(R_j^{(1)}(r) - R_j^{(2)}(r)) \end{pmatrix} \tag{4.19}$$

and

$$\mathcal{A}^{\dagger}\,\overset{\rightleftharpoons}{\Psi}_{-j} = \frac{e^{-ij(\varphi/2)}}{\sqrt{2}}\begin{pmatrix} R_j^{(1)} + R_j^{(2)} \\ i(R_j^{(1)} - R_j^{(2)}) \end{pmatrix} \tag{4.20}$$

The two possible forms of eigenstates corresponding to j=0 are

$$\mathcal{A}^{\dagger}\,\overset{\rightleftharpoons}{\Psi}_{o,g} = \sqrt{2}\begin{pmatrix} R_{o,g} \\ 0 \end{pmatrix} \tag{4.21}$$

and

$$\mathcal{A}^{\dagger}\,\overset{\rightleftharpoons}{\Psi}_{o,u} = i\sqrt{2}\begin{pmatrix} 0 \\ R_{o,u} \end{pmatrix} \tag{4.22}$$

The functions $R_j^{(1)}$, $R_j^{(2)}$, $R_{o,g}$, and $R_{o,u}$, as well as the corresponding eigenvalues of (3.80) are still determined by solving (4.15-17).

These remarks can immediately be carried over to a study of equations (3.71) and (3.75) by merely substituting 2Λ for Λ in our above discussion. In particular, the eigenvalues of these quasi-Hamiltonians are always associated with even values of j.

As the quasi-Hamiltonian (3.92b) is but a special case of (3.78) with $\Lambda = 1$ its eigenvalues are associated with odd values of j and, as we have remarked previously, are all two-fold degenerate. The quasi-Hamiltonian (3.92d) is also a special case of (3.78) with $\Lambda = 2$. Its eigenvalues are associated with even values of j and may be both non-degenerate and two-fold degenerate.

The second example we shall consider is the Moffitt-Liehr quasi-Hamiltonian (3.92a)

$$K + \mathcal{E}(x_o) + Ar^2 + r\begin{pmatrix} 0 & Ce^{-i\varphi} \\ Ce^{i\varphi} & 0 \end{pmatrix} + r^2\begin{pmatrix} 0 & De^{2i\varphi} \\ De^{-2i\varphi} & 0 \end{pmatrix} \qquad (3.92a')$$

where we shall take C and D real (i. e., assume the existence of a reflection plane containing the symmetry axis). The invariants of this quasi-Hamiltonian are $(*)\sigma_x$, $C\sigma_x$, and $\mathcal{R}_{2\pi/3}\, e^{i(\pi/3)\sigma_z}$ as well as all the products of these operators. Here $\mathcal{R}_{2\pi/3}$ is the operator which sends φ into $\varphi + 2\pi/3$.

We note that

$$[\, \mathcal{R}_{2\pi/3}\, e^{i(\pi/3)\sigma_z}\,]^3 = -1 \qquad (4.23)$$

It then follows that the eigenvalues of

$$\mathcal{R}_{2\pi/3}\, e^{i\pi/3\, \sigma_z}$$

are of the form

$$e^{ij(\pi/3)} \tag{4.24}$$

where j is an odd integer. Since j is arbitrary mod 6 there are only three values of j we need consider - $j = 1, 3, 5$. We use the symbol j here to stress the analogy with the example previously considered. It is readily verified that any eigenvector of the operator given by (4.1) associated with the eigenvalue j is also an eigenvector of $\mathcal{R}_{2\pi/3}e^{i(\pi/3)\sigma_z}$ associated with the eigenvalue $e^{ij(\pi/3)}$. This analogy indicates that we can discuss the eigenvalues and eigenvectors of (3.92a') in a manner very similar to that in which we discussed those of (3.78).

Since $\mathcal{R}_{2\pi/3}\, e^{i(\pi/3)\sigma_z}$ is an invariant we may choose the eigenvectors of (3.92a') to also be eigenvectors of this operator. Let $\overset{\rightleftharpoons}{\Psi}_j$ be an eigenvector of (3.92a') which is also an eigenvector of $\mathcal{R}_{2\pi/3}\, e^{i(\pi/3)\sigma_z}$

$$\mathcal{R}_{2\pi/3}\, e^{i(\pi/3)\sigma_z}\, \overset{\rightleftharpoons}{\Psi}_j = e^{ij(\pi/3)}\, \overset{\rightleftharpoons}{\Psi}_j \tag{4.25}$$

Then $\overset{\rightleftharpoons}{\Psi}_j$ must be degenerate with $(\ast)\sigma_x\, \overset{\rightleftharpoons}{\Psi}_j$ as $(\ast)\sigma_x$ is also an invariant. We note, however, that

$$\begin{aligned} &\mathcal{R}_{2\pi/3}\, e^{i(2\pi/3)\sigma_z}\, [(\ast)\, \sigma_x\, \overset{\rightleftharpoons}{\Psi}_j\,] \\ &\qquad = e^{-ij(\pi/3)}[(\ast)\sigma_x\, \overset{\rightleftharpoons}{\Psi}_j\,] \end{aligned} \tag{4.26}$$

Thus we may take

$$(\bigstar)\ \sigma_x \overset{\Rightarrow}{\Psi}_j = \overset{\Rightarrow}{\Psi}_{-j} \qquad (4.27)$$

An eigenstate for which $j = 1$ (Mod 6) is therefore always degenerate with an eigenstate for which $j = 5$ (Mod 6). It follows also that $\overset{\Rightarrow}{\Psi}_j$ and $\overset{\Rightarrow}{\Psi}_{-j}$ are linearly independent unless $j=3$ (Mod 6).

It is also possible to choose the $\overset{\Rightarrow}{\Psi}_j$ such that

$$C(\bigstar)\ \overset{\Rightarrow}{\Psi}_j = \overset{\Rightarrow}{\Psi}_j \qquad (4.28)$$

This follows since

$$\mathcal{R}_{2\pi/3}\ e^{i(\pi/3)\sigma_z}\ [C(\bigstar)\ \overset{\Rightarrow}{\Psi}_j] = e^{ij(\pi/3)}[C(\bigstar)\overset{\Rightarrow}{\Psi}_j]$$

and since any eigenfunction of $C(\bigstar)$ may be chosen to satisfy (4.28) by simply multiplying by a phase factor. We shall assume that the choice (4.28) has been made.

In the event that $j = 3$(Mod 6) we may choose the Ψ_j to be eigenstates of $(\bigstar)\sigma_x$ and we may attach the subscripts g and u on these eigenfunctions - where

$$(\bigstar)\sigma_x \overset{\Rightarrow}{\Psi}_{j,g} = \overset{\Rightarrow}{\Psi}_{j,g} \qquad (4.29a)$$

$$(\bigstar)\sigma_x \overset{\Rightarrow}{\Psi}_{j,u} = -\overset{\Rightarrow}{\Psi}_{j,u} \qquad (4.29b)$$

Our symmetry considerations show no reason for these states being degenerate. Thus we expect the eigenvalues of (3.92a') to be either non-degenerate or two-fold degenerate.

The above considerations allow us to write down the general form of the eigenfunctions of (3.92a')

$$\Psi_j = \begin{pmatrix} e^{i(j-1)\varphi/2} & 0 \\ 0 & e^{i(j+1)\varphi/2} \end{pmatrix} \sum_m e^{i3m\varphi} \begin{pmatrix} R^{(1)}_{j,m} \\ R^{(2)}_{j,m} \end{pmatrix} \tag{4.30}$$

By our convention (4.28) the $R^{(1)}_{j,m}$ and $R^{(2)}_{j,m}$ are real. The state represented above will be degenerate with the state

$$\overset{\rightleftarrows}{\Psi}_{-j} = \begin{pmatrix} e^{i(-j-1)\varphi/2} & 0 \\ 0 & e^{i(-j+1)\varphi/2} \end{pmatrix} \sum_m e^{i3m\varphi} \begin{pmatrix} R^{(2)}_{j,-m} \\ R^{(1)}_{j,-m} \end{pmatrix}$$

Thus we identify

$$R^{(1)}_{-j,m} = R^{(2)}_{j,-m} \tag{4.31a}$$

$$R^{(2)}_{-j,m} = R^{(1)}_{j,-m} \tag{4.31b}$$

The form of the states for $\overset{\rightleftarrows}{\Psi}_{j,g}$ and $\overset{\rightleftarrows}{\Psi}_{j,u}$ (j=3,Mod6) is found by setting j=3 in (4.30) and requiring

$$R^{(1)}_{3,m} = + R^{(2)}_{3,-m-1} \tag{4.32}$$

for the $\overset{\rightleftarrows}{\Psi}_{j,g}$ states and

$$R^{(1)}_{3,m} = - R^{(2)}_{3,-m-1} \tag{4.33}$$

for the $\overset{\rightleftarrows}{\Psi}_{j,u}$ states.

Equations for these radial functions can be determined by substituting expression (4.30) into the Moffitt-Liehr vibronic equation corresponding to (3.92a') and subsequently multiplying by each of the $e^{im\varphi}$ and integrating. This gives us an infinite set of coupled equations, however, instead of just two as we obtained in (4.15).

Often it may be legitimate to treat the term

$$r^2 \begin{pmatrix} 0 & De^{2i\varphi} \\ De^{-2i\varphi} & 0 \end{pmatrix} \qquad (4.34)$$

as a perturbation, the unperturbed Hamiltonian being (3.92b). An interesting question which arises then is that of which levels of (3.92b) are split by this perturbation[1]. Our considerations indicate that only those levels for which $j=3$ (Mod 6) will be split. This follows since only the eigenstates associated with these levels can be simultaneously chosen to be eigenstates of the invariants of (3.92a'). We can, furthermore, use our symmetry considerations to determine the correct zeroth order eigenfunctions. If $\overset{\rightleftharpoons}{\Psi}_j$ and $\overset{\rightleftharpoons}{\Psi}_{-j}$ are two degenerate

---

[1] Moffitt and Liehr [Ref. 5] have also briefly discussed this problem.

eigenstates associated with a $j = 3$ (Mod 6) level and chosen to satisfy

$$(\star)\,\sigma_x\,\overset{\rightleftharpoons}{\Psi}_j \;=\; \overset{\rightleftharpoons}{\Psi}_{-j}$$

as in (4.11) it is clear that

$$\overset{\rightleftharpoons}{\Psi}_{j,g} \;=\; \overset{\rightleftharpoons}{\Psi}_j + \overset{\rightleftharpoons}{\Psi}_{-j} \tag{4.35a}$$

and

$$\overset{\rightleftharpoons}{\Psi}_{j,u} \;=\; \overset{\rightleftharpoons}{\Psi}_j - \overset{\rightleftharpoons}{\Psi}_{-j} \tag{4.35b}$$

are the correct zeroth order wavefunctions as they are also eigenstates of $(\star)\,\sigma_x$ in addition to being eigenstates of $\mathcal{R}_{2\pi/3}\, e^{i\pi/3\,\sigma_z}$. From (4.8) and (4.10) these wavefunctions are simply

$$\Psi_{j,g} \;=\; \begin{pmatrix} e^{-i(\varphi/2)} & 0 \\ 0 & e^{i(\varphi/2)} \end{pmatrix}\begin{pmatrix} e^{ij(\varphi/2)} & e^{-ij(\varphi/2)} \\ e^{-ij(\varphi/2)} & e^{ij(\varphi/2)} \end{pmatrix}\begin{pmatrix} R_j^{(1)} \\ R_j^{(2)} \end{pmatrix} \tag{4.36}$$

and

$$\overset{\rightleftharpoons}{\Psi}_{j,u} \;=\; \begin{pmatrix} e^{-i(\varphi/2)} & 0 \\ 0 & e^{i(\varphi/2)} \end{pmatrix}\begin{pmatrix} e^{ij(\varphi/2)} & -e^{-ij(\varphi/2)} \\ -e^{-ij(\varphi/2)} & e^{ij(\varphi/2)} \end{pmatrix}\begin{pmatrix} R_j^{(1)} \\ R_j^{(2)} \end{pmatrix} \tag{4.37}$$

For the most general (P finite) Moffitt-Liehr quasi-Hamiltonian for the interaction of a two-fold degenerate electronic level with a two-fold degenerate vibration the invariants are $\mathcal{R}_{(2\pi\ell/P)} e^{i(2\pi\Lambda/P)\sigma_z}$, $(\star)C$, and $(\star)\sigma_x$. Symmetry considerations such as those given previously show that the eigenfunctions of such a quasi-Hamiltonian can be chosen such that

$$\mathcal{R}_{(2\pi\ell/P)} e^{i(2\pi\hat{\Lambda}/P)\sigma_z} \overset{\rightleftharpoons}{\Psi}_j = e^{ij(2\pi/P)} \overset{\rightleftharpoons}{\Psi}_j \tag{4.38}$$

where j can take on the values $m\ell+\Lambda$ for all integer m. Each state $\overset{\rightleftharpoons}{\Psi}_j$ will be degenerate with a state $\overset{\rightleftharpoons}{\Psi}_{-j}$ and these states will be linearly independent unless

$$2j = 0 \ (\text{Mod } P) \tag{4.39}$$

For all finite P there are always allowed values of j which satisfy this relation[1]. Thus we may always expect the possibility of non-degenerate levels if P is finite. The non-degenerate states can always be chosen to be eigenstates of $(\star)\sigma_x$

$$(\star)\sigma_x\Psi_j = \star \overset{\rightleftharpoons}{\Psi}_j \tag{4.40}$$

and we may speak of gerade and ungerade states. All states may be chosen to be eigenstates of $C(\star)$ with eigenvalue 1.

$$C(\star) \overset{\rightleftharpoons}{\Psi}_j = \overset{\rightharpoonup}{\Psi}_j \tag{4.41}$$

---

[1] Consider, for example, the very hypothetical case $P=127$, $\ell=2$, $\Lambda=2$. Then (4.39) will be satisfied if $j = 0$ (Mod 254).

This discussion could have also been phrased in group theoretical terminology. While all the energy levels correspond to representations of the group C_{pv} not all representations of this group are in general allowed because of the nature of our generating operators. For example, if $P = 6$, $\Lambda = 1$, ℓ =21we must allow only those representations of C_{6v} for which the rotation of the molecule about 180° is represented by the matrix

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

for two-dimensional representations or

$$-1$$

for one-dimensional representations. There are only three such representations. In particular, the identity representation is not allowed. In the case of $\ell = 2$, $P = \infty$, and Λ odd the operation R_{π} must be represented by (-1) also. As there are no non-degenerate representations of $C_{\infty v}$ which have this property we accordingly found all the eigenvalues to be at least two-fold degenerate.

B. The Potential Surfaces.

The choice of a method for attacking the problem of solving the Born-Renner or Moffitt-Liehr vibronic equations depends for the most part on the nature of the two surfaces $\mathcal{E}^{I}(X)$ and $\mathcal{E}^{II}(X)$ which arise from the degenerate level. It is therefore necessary at this point to examine these surfaces and to make a tentative classification of their possible forms.

For the case of a two-fold degenerate level the energy surfaces are just the eigenvalues of the matrix $\mathcal{W}$ and simply

$$f \pm |g| \tag{4.42}$$

Thus the form of the energy surfaces is determined up to a certain point simply from symmetry considerations. We list below the forms of the energy surfaces for some of the cases explicitly considered in the preceding section. First, the case for which $P = \infty$:

1) $P = \infty$, $\ell = 1$ (one two-fold degenerate vibration)

$$\mathcal{E}(X_o) + \bar{f}(r^2) \pm r^{2\Lambda}\bar{g}(r^2) \tag{4.43a}$$

2) $P = \infty$, $\ell = 2$

$$\mathcal{E}(X_o) + \bar{f}(r^2) \pm r^{\Lambda}\bar{g}(r^2) \tag{4.43b}$$

The energy surfaces for the four cases corresponding to (3.92) are

A) $$\mathcal{E}(X_o) + Ar^2 \pm (C^2r^2 + D^2r^4 + 2CDr^3\cos 3\varphi)^{1/2} \qquad (4.44a)$$

B) $$\mathcal{E}(X_o) + Ar^2 \pm Cr \qquad (4.44b)$$

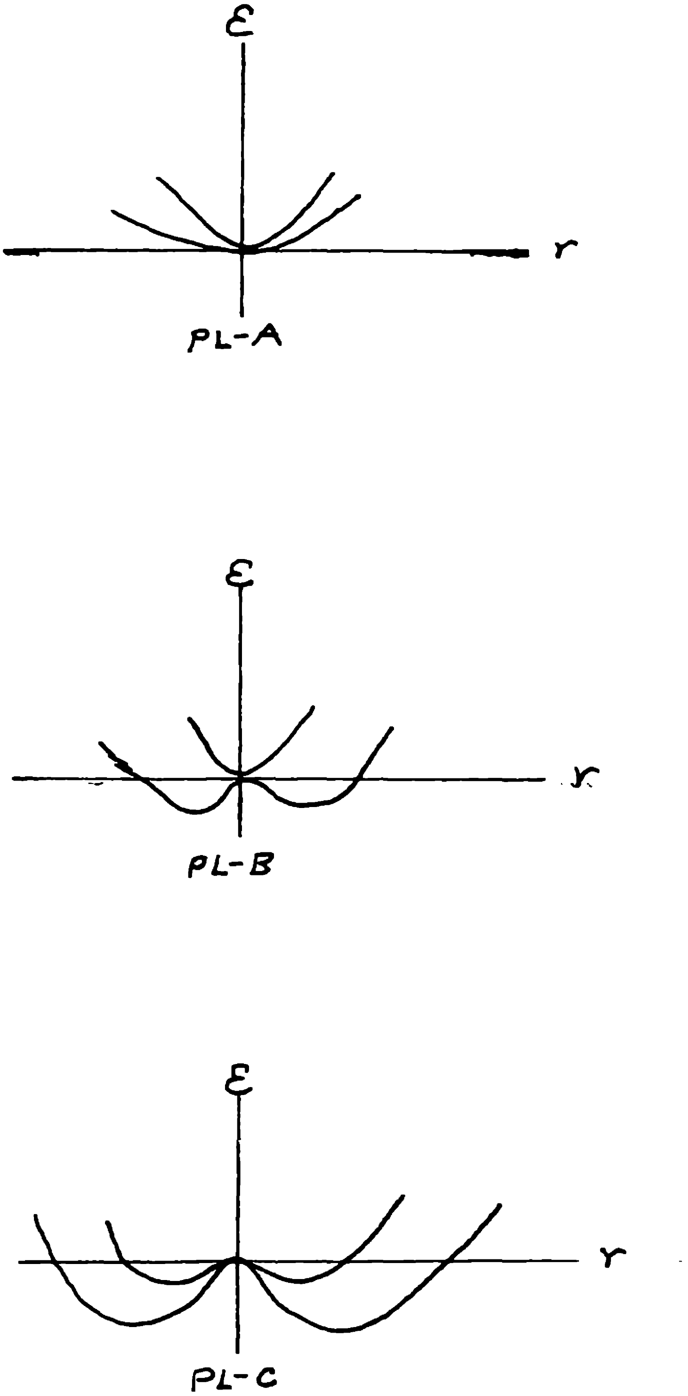
C) $$\mathcal{E}(X_o) + Ar^2 \pm r^2(C^2+D^2+2CD\cos 4\varphi)^{1/2} \qquad (4.44c)$$

D) $$\mathcal{E}(X_o) + Ar^2 \pm Cr^2 \qquad (4.44d)$$

For the four-atom molecule considered in section I of Chapter 3 the energy surfaces are

$$\begin{aligned}
&\mathcal{E}(X_o) + Ar^2 + \alpha_1 \xi_1^2 + \alpha_2 \xi_2^2 + \alpha_3 \xi_3^2 \\
&\pm [r^4(B^2 + C^2 + 2BC\cos 4\varphi) + D^2 \xi_2^2 + F^2 \xi_3^2 + G^2 \xi_1^2 \xi_2^2 \\
&+ 2DE\, \xi_1 \xi_2^2 + 2FG\, \xi_1 \xi_3^2 + 2r^2 \xi_2 D(B+C)\cos 2\varphi \\
&+ 2r^2 \xi_1 \xi_2 E(B+C)\cos 2 + 2r^2 \xi_3 (B-C)F\sin 2\varphi \\
&+ 2r^2 \xi_1 \xi_3 (B-C)F \sin 2\varphi\,]^{1/2} \qquad (4.45)
\end{aligned}$$

In a study of such surfaces the first question that arises is at what configurations are the energy surfaces a maximum or a minimum. Let us consider first the surfaces described by (4.43a). We note that the configuration r=0 is an extremal for both surfaces are a function of r^2 only. There are essentially only three distinct cases of interest and these have been classified



PL-A

PL-B

PL-C

Figure 3

by Pople and Longuet-Higgins [Ref. 34] as follows:

Case PL-A. The configuration $r=0$ is a minimum for both surfaces.

Case PL-B. One surface (the upper surface) has a minimum at $r=0$ while the other has a maximum at this configuration.

Case PL-C. Both surfaces have a maximum at $r=0$.

These three possibilities are illustrated in Fig. 3 .

The case PL-A was that considered by Renner [Ref. 3] in his theory of the degenerate electronic states of CO_2. Pople and Longuet-Higgins have studied the vibronic equations resulting from PL-B and have used the resulting theory to explain certain anomalies in the vibrational spectra of NH_2. The case PL-C has never been explicitly considered in the literature.

In constructing the diagrams in Fig. 3 it has been assumed that the surfaces must have a minimum at some value of r - otherwise, the molecule would be unstable. It has also been assumed that the surfaces touch only at $r=0$ - implying that $\bar{g}(r^2)$ is either always positive or always negative (although possibly zero at $r\neq 0$). This means we exclude accidental degeneracies.

If one wishes to include only harmonic terms, i.e., neglect terms of order r^4 and higher, the only case of interest is PL-A. The cases PL-B and PL-C would apparently lead to instability in this approximation. Furthermore, the surfaces for PL-A would appear to be always degenerate in the harmonic approximation unless $\Lambda = 1$. This explains why Renner only considered the π-states in his theory of CO_2. It is, of course, entirely conceivable that consideration of the higher order terms may lead to a significant modification of the spectra - particularly in cases PL-B and PL-C.

Let us now consider the surfaces given by (4.43b). A distinction will be made between the cases of even and odd Λ, respectively. For Λ even, the surfaces are identical in form to a surface of the form (4.43a) and the preceding remarks may be applied. If Λ is odd, however, we must modify our considerations somewhat. If Λ is one the surfaces have a cuspidal shape at r=0. The most likely possibility for the form of the surfaces is that depicted in Fig. 4a. An alternate possibility may conceivably occur when the upper surface has a true minimum at a value of r which is nonzero. In general, we shall not need make a distinction between these two possibilities unless this minimum

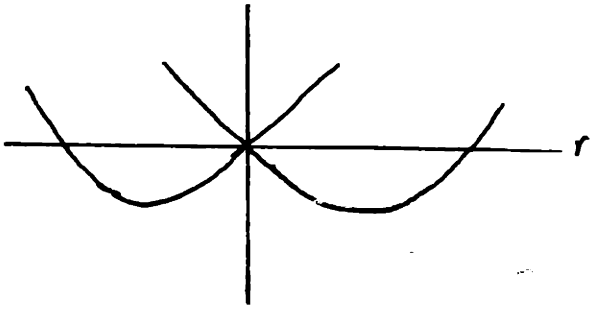


Fig. 4a

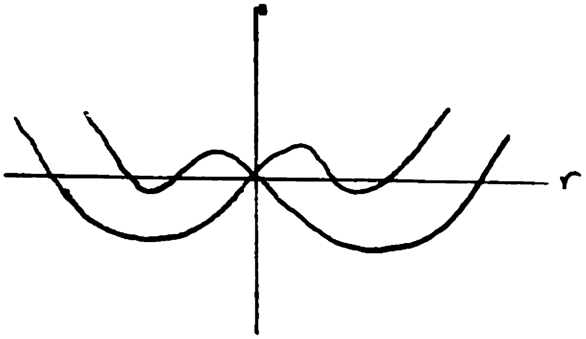


Fig. 4b

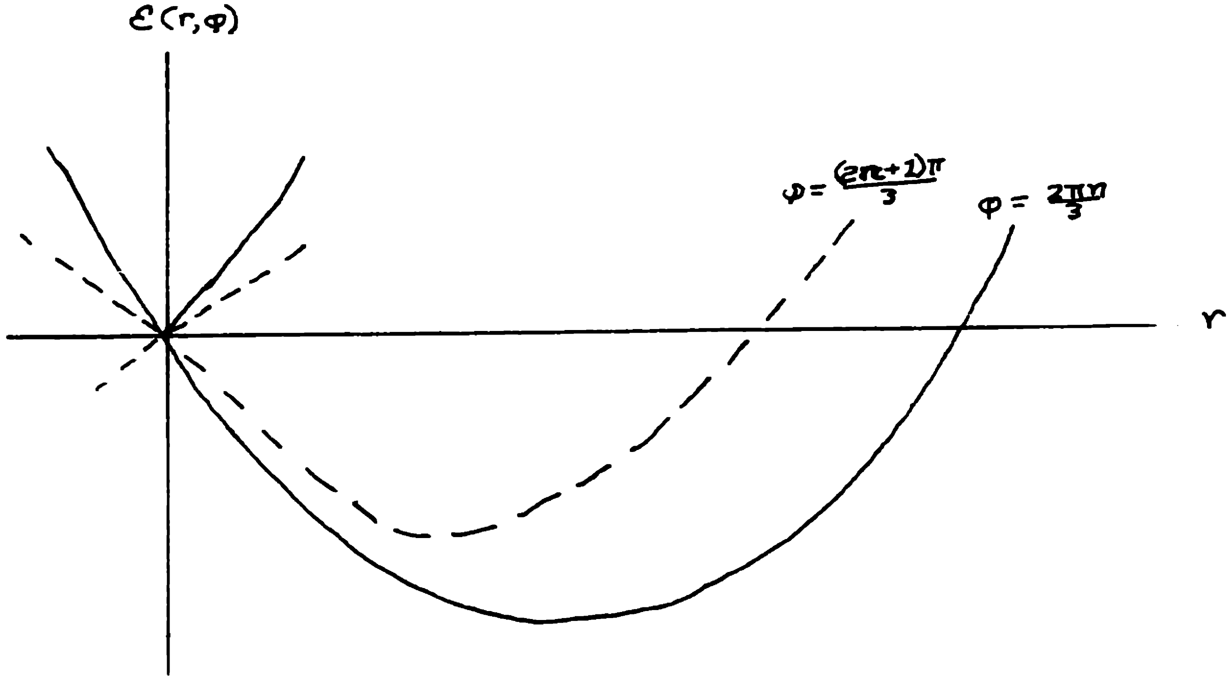


Fig. 5

lies below $\mathcal{E}(X_0)$. This situation is depicted in Fig. 4b. The lower surface must always have a minimum at a non-zero value of r in order that the molecule be stable. We note that there is a certain similarity in the cases depicted by Figs. 4a and 4b to those of PL-B and PL-C, respectively. If $\Lambda = 3, 5$, etc., there is no such cuspidal appearance at r = 0. The configuration r=0 is either a minimum point for both surfaces or a maximum for both surfaces - depending on the behavior of the function $\bar{f}(r^2)$. The form of the potential surfaces in these cases is analogous to those depicted for PL-A and PL-C. If we restrict ourselves to the harmonic approximation, the only case of interest is that where $\Lambda = 1$. If Λ is 3 or higher the levels appear always degeneracies and we expect no anomalies to appear in the spectra - providing the harmonic approximation is valid.

If we consider the surfaces described by (4.44b) and (4.44d) we see that they are but special cases of (4.43a) and (4.43b). In order that the approximations made in obtaining these surfaces (i.e., the harmonic approximation) be valid we must require A to be positive in both equations and $|C| < A$ in equation (4.44d). The resulting surfaces are then those depicted by Fig. 4a and by the diagram for Case PL-A.

The surfaces described by (4.44a) and (4.44c), on the other hand, depend on the angle φ. The energy surfaces corresponding to (4.44a) have already been discussed to some extent by Moffitt and Liehr [Ref. 5]. Let us assume for definiteness that A, B, and C are all positive. Analogous results are obtained if B or C or both B and C are negative. If is assumed that

$$A > |B| + |C|$$

Otherwise the lower sheet would have no minima. With this assumption the lower sheet has three minima at

$$r = \frac{C}{2(A-D)}, \qquad \varphi = \frac{2\pi n}{3} \tag{4.46a}$$

and three saddle points at

$$r = \frac{C}{2(A+D)}, \qquad \varphi = \frac{(2n+1)\pi}{3} \tag{4.46b}$$

where n is an integer. The values of $\mathcal{E} - \mathcal{E}(X_o)$ at these points are

$$-\frac{C^2}{4(A-D)} \qquad \text{and} \qquad -\frac{C^2}{4(A+D)} \tag{4.47}$$

respectively. The curvatures of the energy surface in r direction at these two classes of points are

$$2(A+D) \qquad \text{and} \qquad 2(A-D)$$

respectively. These results are diagrammed in Fig. 5 . These surfaces will degenerate to those for (4.44a) if on allows D to go to zero. The surfaces are then independent of φ and the lower surface has a minimum at

$$r = c/2A \qquad (4.48)$$

with a corresponding value of $\mathcal{E} - \mathcal{E}(X_0)$ of

$$- \frac{c^2}{2A} \qquad (4.49)$$

If D is small one may take the surfaces in this form as an initial step - taking the terms in the quasi-Hamiltonian depending on D as a perturbation. This was essentially the procedure followed by Moffitt and Liehr [Ref. 5] in their solution for the eigenstates of (3.92a).

The surfaces described by (4.44c) must both have their only minimum at r=0. Otherwise an instability would be indicated. We must assume

$$A > |C| + |D|$$

as before in order that this be satisfied. The upper curve has maximum curvature in the r direction for

$$\varphi = \frac{2\pi n}{4}$$

and minimum curvature for

$$\varphi = \frac{(2n+1)\pi}{4}$$

while the converse is true for the lower surface (We assume C and D are positive). In Fig. 6 we plot the cross section of these two surfaces. Whenever C or D should vanish the dotted curves coincide with the solid

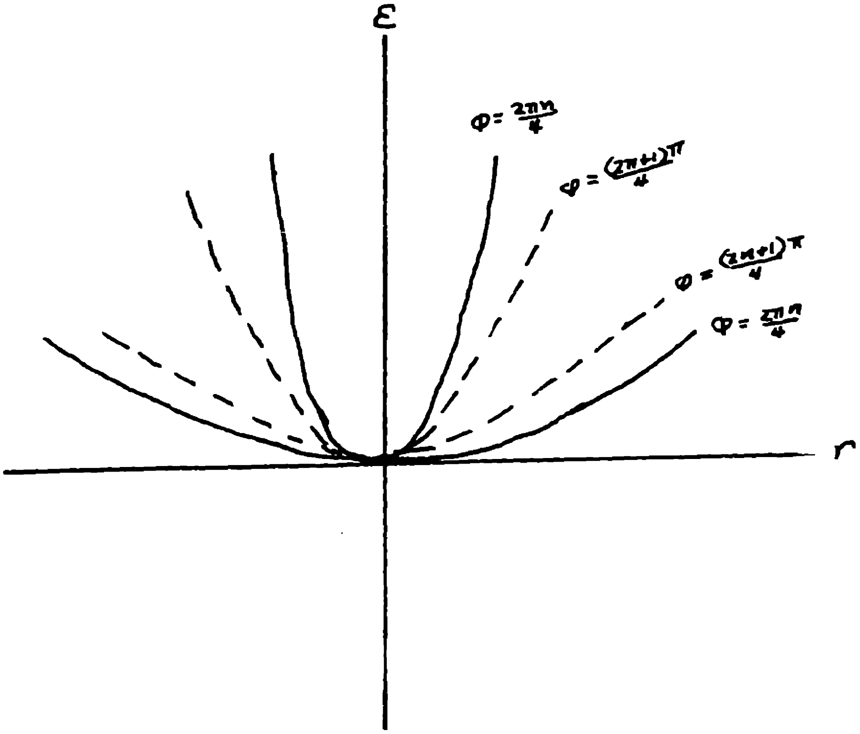

ε
$\varphi = \frac{2\pi n}{4}$
$\varphi = \frac{(2n+1)\pi}{4}$
$\varphi = \frac{(2n+1)\pi}{4}$
$\varphi = \frac{2\pi n}{4}$
r

Fig. 6

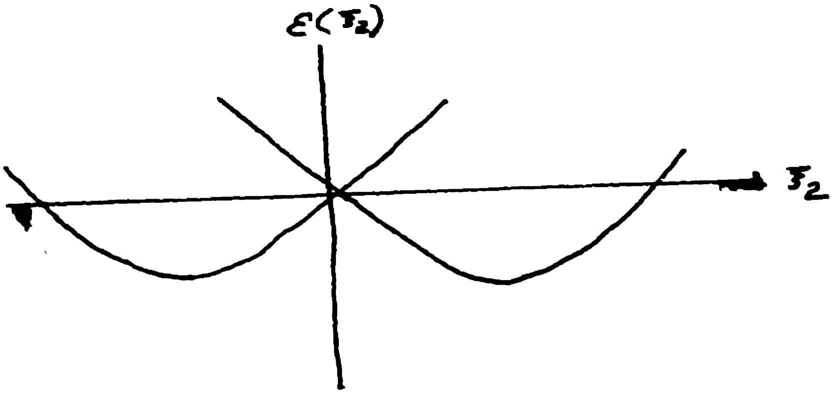

$\varepsilon(\xi_2)$
ξ_2

Fig. 7

curves and one has the situation depicted in Fig. in Chapter 2.

The surfaces represented by (4.45) are very complex because of the large number of vibrational coordinates considered. The directions of greatest splitting of the degeneracy at the configuration X_o are clearly the ξ_2 and ξ_3 directions as in only these coordinates is the splitting of first order. Let us, for simplicity, examine these surfaces as a function of ξ_2 only. Setting r, ξ_1, and ξ_3 equal to zero the surfaces become

$$\mathcal{E}(X_o) + \alpha_2 \xi_2^2 \pm D \xi_2$$

Thus along the intersection of the energy surfaces with the "plane" r=0, ξ_1=0, ξ_3=0, the lower surface has a minimum at

$$\xi_2 = D/2\alpha_2$$

and a corresponding value of $\mathcal{E} - \mathcal{E}(X_o)$ of $-D^2/4\alpha_2$. These surfaces are plotted in Fig. 7 and are readily seen to be identical in form to those in Fig. 4a when D =0. Here, however, we are dealing with a non-degenerate vibration rather than a two-fold degenerate vibration.

The fact that the lower surfaces in Figs. 4a and 7 have their minima at configurations other than

X_o may be interpreted as a consequence of the Jahn-Teller theorem. The tendency of nonlinear molecules in degenerate electronic states to distort in order to lift their degeneracy is called the Jahn-Teller effect.[1] The problem of determining the manner in which the electronic energies vary with the vibrational coordinates and the new stable equilibrium configurations has been called the static Jahn-Teller problem by Opik and Pryce [Ref. 38]. Examples of the static Jahn-Teller problem have been considered in the literature by Van Vleck [Ref. 37], Opik and Pryce [Ref. 38], Hobey and McLachlan [Ref. 33], Teller [Ref. 39], Liehr [Ref. 40], and Snyder [Ref. 41].

Thus this section has been concerned in part with the static Jahn-Teller problem. The examples represented by equations (4.43a), (4.44c), (4.44d) as well as (4.43b) for $\Lambda \neq 1$ do not fall into the category of static Jahn-Teller problems as the splitting of the potential surfaces is not linear in the vibrational coordinates. One of our purposes, however, is to stress that there is a formal similarity in all the examples we are considering.

---

[1] This terminology is evidently due to Van Vleck [Ref. 37].

Actually the second order splitting of the energy surfaces may lead to more prominent effects than the first order splitting and any realistic calculation should consider second order splitting terms as well as first order terms.

C. Symmetry of the Ground State.

A question of considerable interest in molecular physics is that of the symmetry properties of the ground state and, in particular, whether or not this state is degenerate. In general, the answer to this question requires a detailed solution of the Schroedinger equation for the molecule. However, in certain simple cases this question may be partly or entirely answered by relatively simple considerations. This is true for the case when we limit ourselves to the case of the interaction of a two-fold degenerate level with a two-fold degenerate vibration and when the ground state eigenvalue is given to a good approximation by that of a Moffitt-Liehr quasi-Hamiltonian of the form (3.78) for some Λ.

In particular, we find the ground state eigenvalue of (3.78) must be associated with a value of j which satisfies the relation

$$|j| \leqslant \Lambda \qquad (4.50)$$

If $\Lambda = 1$ the ground state symmetry is unambiguous - it must be $|j| = 1$. For higher values of Λ there is a narrow choice of ground state symmetries - governed by equation (4.50). We present our proof of these statements below.

The eigenvalues of (3.78) associated with a particular value of j are given by solving the equation (4.15). Let $E_{j,o}$ be the lowest eigenvalue of this equation and

$$\vec{R}_{j,o} = \begin{pmatrix} R^{(1)}_{j,o} \\ R^{(2)}_{j,o} \end{pmatrix}$$

be the eigenvector (or one of the eigenvectors) of (4.15) associated with this lowest eigenvalue. In general we shall abbreviate

$$\mathbb{H}_j = -\frac{\hbar^2}{2M}\left[\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr} - n\frac{j^2}{4r^2} - \frac{\Lambda^2}{4r^2} + \frac{\Lambda j}{2r^2}\sigma_z\right] + \frac{\hbar^2}{2M}\left[\bar{h}_o(r^2) + r^{\Lambda}\bar{j}_o(r^2)\sigma_x\right] + (\Lambda - j\sigma_z)\frac{\hbar^2}{2M}\bar{\bar{h}}_2(r^2) + \bar{f}(r^2) + r^{\Lambda}\bar{g}(r^2)\sigma_x \qquad (4.51)$$

for the "Hamiltonian" operator of (4.15). We note that this operator depends on j only through the terms

$$\frac{\hbar^2}{2M}\left[\frac{j^2}{4r^2} - \frac{\Lambda j}{2r^2}\sigma_z\right] - j\,\sigma_z\,\frac{\hbar^2}{2M}\,\bar{\bar{h}}_2(r^2) \qquad (4.52)$$

In what follows the presence of the term $-j\sigma_z(\hbar^2/2M)\bar{\bar{h}}_2$

shall be ignored. We expect this term to be small (Appendix C) and dodnot expect it to affect the relative ordering of the energy levels.

The variational principle for a Hermitian operator indicates that

$$\frac{\int \vec{R}^{\dagger}_{j,o} \; \mathcal{H}_{j'} \; \vec{R}_{j,o} \; r dr \; d\varphi}{\int \vec{R}^{\dagger}_{j,o} \; \vec{R}_{j,o} \; r dr \; d\varphi} \geqslant E_{j',o}$$

Therefore

$$E_{jo} + \frac{\int \vec{R}^{\dagger}_{j,o} \, [\, \mathcal{H}_{j'} - \mathcal{H}_{j} \,] \, \vec{R}_{j,o} \; r dr \; d\varphi}{\int \vec{R}^{\dagger}_{j,o} \vec{R}_{j,o} \; r dr \; d\varphi} \geqslant E_{j',o}$$

From this it follows that

$$E_{j,o} \geqslant E_{j',o} \tag{4.53}$$

if

$$\frac{\int \vec{R}^{\dagger}_{j,o} \, [\, \mathcal{H}_{j'} - \mathcal{H}_{j} \,] \, \vec{R}_{j,o} \; r dr \; d\varphi}{\int \vec{R}^{\dagger}_{j,o} \; \vec{R}_{j,o} \; r dr \; d\varphi} \leqslant 0 \tag{4.54}$$

A glance at (4.52) shows that this equation will always be satisfied if both of the conditions

$$j^2 - 2\Lambda j - j'^2 + 2\Lambda j' \geqslant 0$$

$$j^2 + 2\Lambda j - j'^2 - 2\Lambda j' \geqslant 0$$

are simultaneously satisfied. These conditions reduce, however, to

$$j \geqslant 2\Lambda - j' \qquad (4.55a)$$

$$j \geqslant -2\Lambda - j' \qquad (4.55b)$$

if $j > j'$, or

$$j \leqslant 2\Lambda - j' \qquad (4.56a)$$

$$j \leqslant -2\Lambda - j' \qquad (4.56b)$$

if $j < j'$. (The case $j=j'$ obviously need not be considered.) Letting $j' = \Lambda$ in the above relations gives

$$E_{j,o} \geqslant E_{\Lambda,o} \qquad \text{if} \quad j > \Lambda \qquad (4.57)$$

Since the eigenvalues of $\mathbb{H}_j$ are identical to those of $\mathbb{H}_{-j}$ this statement can be generalized to

$$E_{j,o} \geqslant E_{\Lambda,o} \qquad \text{if} \quad |j| > 0 \qquad (4.58)$$

From this the assertions associated with (4.50) follow directly.

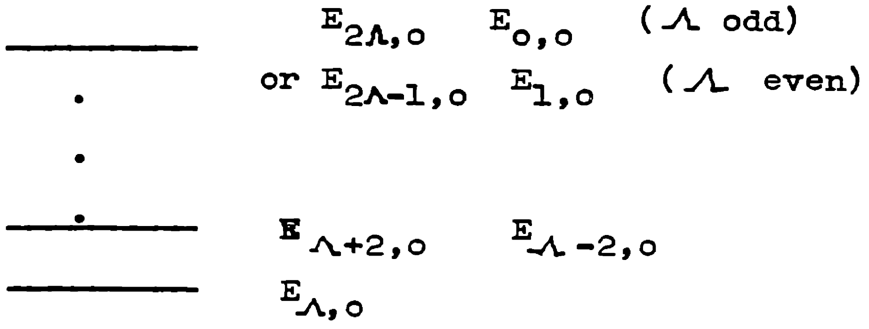
This proves, in particular, that the lowest eigenvalue of (3.92b) is always a $|j| = 1$ eigenvalue. This fact is borne out by the detailed numerical calculations of Longuet-Higgins et al. [Ref. 6] and of Moffitt and Thorson [Ref. 42].

Also, since the $j = 1$ level is not split by a perturbation of the form (4.34) we have good reason to believe that the ground state of (3.92a') is also degenerate. It is very interesting in this context to note that a recent elaborate calculation outside the framework of the Born-Renner approximation by McConnel and McLachlan [Ref. 43] found the ground vibronic state of the benzene negative ion to be doubly degenerate. It was for the study of this molecule that Moffitt and Liehr [Ref. 5] first derived the quasi-Hamiltonian (3.92a).

For higher values of Λ the symmetry of the ground state depends on the nature of the functions $\bar{f}$ and $\bar{g}$ (and also to a minor extent on $\bar{h}_o$, $\bar{j}_o$, and $\bar{\bar{h}}_2$). In the limit, for example, when g is identically zero the ground state must be of symmetry $|j| = \Lambda$. The ground states of the various $\mathbb{H}_j$ then satisfy the following relations

$$E_{\Lambda,o} \leqslant E_{\Lambda+2,o} = E_{\Lambda-2,o} \leqslant E_{\Lambda+4,o} = E_{\Lambda-4,o} \leqslant \cdot \quad \cdot \quad \cdot \tag{4.59}$$

This situation is illustrated on the "energy level" diagram below:



This diagram shows the accidental degeneracy of the $\mathcal{H}_{\Lambda+j}$ eigenvalues with the $\mathcal{H}_{\Lambda-j}$ eigenvalues which occurs in this limit. When g is non-zero we expect this degeneracy to be removed. Our preceding remarks suggest the possibility, furthermore, of the levels $E_{\Lambda-2,0}$, $E_{\Lambda-4,0}$, . . . , $E_{0,0}$ (or $E_{1,0}$ if Λ is odd) being lower than $E_{\Lambda,0}$ - although they preclude the possibility of the $E_{\Lambda+2,0}$, etc., levels falling below $E_{\Lambda,0}$.

A second possibility which demonstrates that the ground state need not necessarily be a $|j| = \Lambda$ state occurs when the lower energy surface does not have its minimum at r=0 (as in the case PL-B discussed in the previous section). If the energy at athe bottom of the trough formed by the lower surface is sufficiently low compared to that of $\mathcal{E}(X_0)$ we expect that the system will be almost entirely confined to the lower electronic state when it is in its ground state. In this event the ground state will be a $|j|=1$ state if Λ is odd and a $|j| = 0$ state if Λ is even. This is readily seen to be so by subjecting $\mathcal{H}_j$ to a unitary transformation such that the matrix $r^{\Lambda}g\sigma_x$ is diagonal. Doing this gives

$$\mathcal{H}_j' = -\frac{\hbar^2}{2M}\left[\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr} - \frac{j^2}{4r^2} - \frac{\Lambda^2}{4r^2} + \frac{\Lambda j}{2r^2}\sigma_y\right] + \frac{\hbar^2}{2M}\left[\bar{h}_o + r^{\Lambda}\bar{j}_o\sigma_z\right] + (\Lambda - j\sigma_y)\frac{\hbar^2}{2M}\bar{\bar{h}}_2 + \bar{f} + r^{\Lambda}\bar{g}\sigma_z \qquad (4.60)$$

Here we can treat $\frac{\Lambda j}{2r^2}\sigma_y$ as a negligible perturbation (and also $j\sigma_y\bar{\bar{h}}_2$) provided

$$\frac{\hbar^2}{2M}\frac{\Lambda j}{2r_o^2} \ll \mathcal{E}(x_o) - \mathcal{E}^{II}(x))_{r=r_o}$$

where r_o is the value of r which corresponds to the bottom of the trough on the lower surface. The unperturbed Hamiltonian then depends on j primarily through the term

$$\frac{\hbar^2}{2M}\frac{j^2}{4r^2}$$

(It is assumed that $|h_2(r_o)| \ll (1/r_o^2)$) indicating that

$$E_{o,o} \leq E_{2,o} \leq \cdot \quad \cdot \quad \cdot \quad \cdot \leq E_{\Lambda,o}$$

if is even, or

$$E_{1,o} \leq E_{3,o} \leq \cdot \quad \cdot \quad \cdot \leq E_{\Lambda,o}$$

if is odd. This verifies our assertion.

This shows, in particular, that the ground state may be either non-degenerate or two-fold degenerate if Λ is even. Our preceding remarks therefore apply to that of the ground states of (3.71) or (3.92d).

D. Solution of the Vibronic Equations in Some Limiting Cases.

The detailed solution of the Born-Renner or Moffitt-Liehr vibronic equations is in general a most difficult problem and lengthy numerical calculations are required for a complete solution for even the simplest models [Ref. 6 , 42]. However, in certain limiting cases perturbation techniques may be used and it is such limiting cases that shall be discussed here.

Consider the lowest quasi-potential surface arising from the interaction of a two-fold degenerate vibration with a two-fold degenerate electronic level. As we mentioned in Section B this function may either have its minimum at r=0 or elsewhere. If it has its minimum at another configuration there exists the possibility that we may consider the system as being confined entirely to the lower surface in the first approximation. When this situation occurs we shall speak of very strong electron-vibration coupling.

A second possibility is that the upper and lower energy surfaces may nearly coincide. Then the splitting of the two surfaces may be treated as a perturbation.

This shall be called the case of very weak electron-vibration coupling.

Let us first consider the case of very strong electron-vibration coupling. In the neglect of the matrix[1] $\mathbb{T}$ the Moffitt-Liehr quasi-Hamiltonian is

$$\mathbb{H} = K + \begin{pmatrix} f(r,\varphi) & g(r,\varphi) \\ g^*(r,\varphi) & f(r,\varphi) \end{pmatrix} \tag{4.61}$$

To discuss the strong coupling case it is advantageous to transform this to a form where the matrix $\mathbb{U}$ is diagonal. A matrix $\mathbb{A}$ which accomplishes this (see equation (3.34)) is readily seen to be

$$\mathbb{A} = \frac{1}{\sqrt{2}}\frac{1}{|g|}\begin{pmatrix} |g| & -g \\ g^* & |g| \end{pmatrix} \tag{4.62}$$

Then

$$\begin{aligned} \mathbb{H}' &= \mathbb{A}^+ \mathbb{H} \mathbb{A} \\ &= K + \frac{1}{2}\begin{pmatrix} \frac{g}{|g|}(K, g^*/|g|) & -(K, g/|g|) \\ (K, g^*/|g|) & \frac{g^*}{|g|}(K, g/|g|) \end{pmatrix} \\ &\qquad + f(r,\varphi) \quad + \sigma_z\, |g(r,\varphi)| \end{aligned} \tag{4.63}$$

which is one form of the Born-Renner quasi-Hamiltonian. This operator has the following invariants which follow from the symmetry properties of the functions $f(r,\varphi)$ and $g(r,\varphi)$ (see equations (3.45)):

$$\mathcal{R}_{l\alpha} \begin{pmatrix} e^{i\lambda\alpha} & 0 \\ 0 & e^{-i\lambda\alpha} \end{pmatrix} \tag{4.64a}$$

$$\begin{pmatrix} g/|g| & 0 \\ 0 & -g^*/|g| \end{pmatrix} (*) \tag{4.64b}$$

$$\begin{pmatrix} g/|g| & 0 \\ 0 & -g^*/|g| \end{pmatrix} \mathcal{C} \tag{4.64c}$$

If the minimum of $\mathcal{E}^{II}$ is sufficiently lower than the minimum of $\mathcal{E}^{I}$ we expect the energy levels and eigenfunctions of (4.63) and also the energy levels of (4.61) to be very nearly those of

$$\mathbb{H}'^{(0)} = K + \frac{1}{2} \begin{pmatrix} (g/|g|)(K, g^*/|g|) & 0 \\ 0 & (g^*/|g|)(K, g/|g|) \end{pmatrix}$$

$$+ f(r, \varphi) + \sigma_z |g(r, \varphi)| \tag{4.65}$$

for sufficiently low energies. This suggests that we may break up $\mathbb{H}'$ as follows

$$\mathbb{H} = \mathbb{H}'^{(0)} + \mathbb{H}'^{(1)} \tag{4.66}$$

where $\mathbb{H}'^{(0)}$ is given by (4.65) and $\mathbb{H}'^{(1)}$ is given by

$$\frac{1}{2} \begin{pmatrix} 0 & -(K, g/|g|) \\ (K, g^*/|g|) & 0 \end{pmatrix} \tag{4.67}$$

We take the eigenfunctions and eigenvalues of $\mathbb{H}'^{(0)}$ to be of the forms

$$\vec{\psi}^{(0)\ \mathrm{I}}_{kn} = \begin{pmatrix} \psi^{(0)\ \mathrm{I}}_{k,n} \\ 0 \end{pmatrix}, \quad E^{(0)\ \mathrm{I}}_{k,n} \tag{4.68a}$$

$$\vec{\psi}^{(0)\ \mathrm{II}}_{k'n'} = \begin{pmatrix} 0 \\ \psi^{(0)\ \mathrm{II}}_{k'n'} \end{pmatrix}, \quad E^{(0)\ \mathrm{II}}_{k'n'} \tag{4.68b}$$

Here the letter k denotes how the states transform under $\mathcal{R}_{\ell\alpha}$:

$$\mathcal{R}_{\ell\alpha}\psi_k = \psi_k(r, \varphi + \ell\alpha) = e^{ik\alpha}\,\psi_k(r, \varphi)$$

Then the correction to an eigenvalue $E^{(0)\ \mathrm{II}}_{kn}$ is found to be

$$-\ \frac{1}{2}\sum \frac{|\langle k'n'\mathrm{I}|(K,\ g/g\ )|k\ n\ \mathrm{II}\rangle|^2}{E^{(0)\ \mathrm{I}}_{k'n'} - E^{(0)\ \mathrm{II}}_{kn}} \tag{4.69}$$

while the correction to a state $\vec{\psi}^{(0)\mathrm{II}}_{k'n'}$ is found to be

$$-\ \sum \frac{\langle k'n'\mathrm{I}|(K,\ g/g\ )\ |\ kn\mathrm{II}\rangle}{E^{(0)\mathrm{II}}_{kn} - E^{(0)\mathrm{I}}_{k'n'}}\ \vec{\psi}^{(0)\mathrm{I}}_{k'n'} \tag{4.70}$$

To gain some insight into the nature of the meaning of (4.69) and (4.70) we replace the denominators in these expressions by Δ where Δ is the difference in energy between the lowest eigenvalues of the two components of $\mathcal{H}'^{(0)}$. Let us note also that

$$\Delta \doteq (\text{Min. } \mathcal{E}^{\mathrm{I}} - \text{Min. } \mathcal{E}^{\mathrm{II}})$$

The denominators in (4.69) and (4.70) will always be larger than this number if $E^{(0)II}_{kn}$ refers to the unperturbed ground state energy. While this replacement is a crude approximation it will clarify our future discussions somewhat. With this approximation (4.69) becomes

$$\frac{1}{2\Delta} \langle kn\ II \,|\, (K, g^*/|g|)(K, g/|g|) \,|\, kn\ II \rangle \tag{4.71}$$

while (4.70) becomes

$$\frac{1}{\Delta} (K, g/|g|) \begin{pmatrix} \psi^{(0)II}_{k,n} \\ 0 \end{pmatrix} \tag{4.72}$$

The approximate eigenstate of the Born-Renner quasi-Hamiltonian is then

$$\begin{pmatrix} 0 \\ \psi^{(0)II}_{k,n} \end{pmatrix} + \frac{1}{\Delta} (K, g/|g|) \begin{pmatrix} \psi^{(0)II}_{kn} \\ 0 \end{pmatrix} \tag{4.73}$$

While the states given by the above are somewhat crude approximations to the actual eigenstates they do give us a convenient means of estimating the probability of the system being in the electronic state associated with the upper quasi-potential surface. This probability is the expectation value of

$$\frac{1}{\Delta^2} (K, g^*/|g|)(g/|g|, K) \tag{4.74}$$

between eigenstates of

$$K + \tfrac{1}{2}(g^*/|g|)(K, g/|g|) + f - |g| \qquad (4.75)$$

The criteria for the validity of the perturbation procedure outlined here is that this number be sufficiently less than unity. Later we shall return and apply this approximation to a specific example.

The perturbation $\mathbb{H}'^{(1)}$ cannot split degenerate eigenvalues associated with two degenerate eigenstates of $\mathbb{H}'^{(0)}$ of the forms

$$\begin{pmatrix} 0 \\ \psi_{k,n}^{(0)II} \end{pmatrix} \quad , \quad \begin{pmatrix} 0 \\ \frac{g^*}{g}\psi_{k,n}^{(0)II^*} \end{pmatrix} \qquad (4.76)$$

These two functions will always be linearly independent providing

$$2j \equiv 2(k - \Lambda) \neq 0 \quad (\mathrm{Mod}\ P) \qquad (4.77)$$

If this condition is not satisfied then we may always choose our wavefunctions such that

$$\frac{g^*}{g}\psi_{k,n}^{(0)II^*} = \pm \quad \psi_{k,n}^{(0)II} \qquad (4.78)$$

in addition to

$$C\ \psi_{k,n}^{*} = \psi_{k,n} \qquad (4.79)$$

where the upper or lower sign holds for any given state. Such states will in general be non-degenerate.

Let us now consider the case of very weak coupling. Here we use the Moffitt-Liehr quasi-Hamiltonian (4.61) taking

$$\mathcal{H}^{(1)} = \begin{pmatrix} 0 & g(r,\varphi) \\ g^{*}(r,\varphi) & 0 \end{pmatrix} \tag{4.80}$$

as a weak perturbation, the unperturbed Hamiltonian being

$$\mathcal{H}^{(0)} = K + f(r,\varphi) \tag{4.81}$$

Let $\Psi^{(0)}_{n,k}$ be an eigenstate of this unperturbed Hamiltonian associated with the energy value $E^{(0)}_{n,k}$. Such functions may be chosen to have the properties

$$\bar{\Psi}^{(0)}_{n,k}(r,\varphi+\ell\alpha) = e^{ik\alpha}\,\bar{\Psi}^{(0)}_{n,k}(r,\varphi) \tag{4.82a}$$

$$\bar{\Psi}^{(0)}_{n,k}(r,\varphi)^{*} = \bar{\Psi}^{(0)}_{n,-k}(r,\varphi) \tag{4.82b}$$

$$\bar{\Psi}^{(0)}_{n,k}(r,-\varphi) = \bar{\Psi}^{(0)}_{n,k}(r,\varphi) \tag{4.82c}$$

These follow from the symmetry properties of the function $f(r,\varphi)$ and arguments similar to those given in Appendix B.

For the present we shall assume none of the eigenstates of $K + f(r,\varphi)$ are accidentally degenerate - although we shall later discuss this possibility. If k is zero (Mod P) the associated eigenvalue of $K + f$ is non-degenerate and we have two different types of wavefunctions, $\bar{\Psi}^{(0)}_{n,g}$ and $\bar{\Psi}^{(0)}_{n,u}$, where both satisfy (4.82c) and $\bar{\Psi}^{(0)}_{n,g}$ is real and $\bar{\Psi}^{(0)}_{n,u}$ is imaginary.

If k is not zero (Mod P) then there are four degenerate eigenstates of $\mathcal{H}^{(0)}$

$$\Psi^{(0)}_{n,\,k+\Lambda} = \begin{pmatrix} \bar{\Psi}^{(0)}_{n,k} \\ 0 \end{pmatrix} \tag{4.83a}$$

$$\Psi^{(0)}_{n,\Lambda-k} = \begin{pmatrix} \bar{\Psi}^{(0)}_{n,-k} \\ 0 \end{pmatrix} \tag{4.83b}$$

$$\Psi^{(0)}_{k-\Lambda,n} = \begin{pmatrix} 0 \\ \bar{\Psi}^{(0)}_{n,k} \end{pmatrix} \tag{4.83c}$$

$$\overset{\rightleftarrows}{\Psi}{}^{(0)}_{n,-k-\Lambda} = \begin{pmatrix} 0 \\ \overset{\rightleftarrows}{\Psi}{}^{(0)}_{n,-k} \end{pmatrix} \tag{4.83d}$$

The subscripts k+Λ, k−Λ, etc., are values of j where

$$\begin{pmatrix} e^{i\Lambda\alpha} & 0 \\ 0 & e^{-i\Lambda\alpha} \end{pmatrix} \overset{\rightleftarrows}{\Psi}_{n,j} = e^{i\alpha j}\ \overset{\rightleftarrows}{\Psi}_{n,j} \tag{4.84}$$

If k is zero we need only consider the states (4.83a) and (4.83c) and the unperturbed level is only two-fold degenerate. We note that only states with the same value of j will mix under the perturbation. There is therefore no first order correction to the energy unless

$$2(\Lambda - k) = 0 \quad (\text{Mod } P) \tag{4.85}$$

$$2(\Lambda + k) = 0 \quad (\text{Mod } P) \tag{4.86}$$

If (4.85) is satisfied the four-fold degenerate level splits into two non-degenerate levels and a two-fold degenerate level. The corrected zeroth order wave-functions and the associated first order energies are

$$\frac{1}{2} \begin{pmatrix} \overline{\Psi}^{(0)}_{n,-k} \\ \overline{\Psi}^{(0)}_{n,k} \end{pmatrix} \qquad E^{(0)}_{n,k} + \mathcal{R}e\langle n,-k \mid g \mid n,k\rangle \tag{4.87a}$$

$$\frac{1}{\sqrt{2}} \begin{pmatrix} \overline{\Psi}^{(0)}_{n,-k} \\ -\overline{\Psi}^{(0)}_{n,k} \end{pmatrix} \qquad E^{(0)}_{n,k} - \mathcal{R}e\langle n,-k \mid g \mid n,k\rangle \tag{4.87b}$$

$$\vec{\overline{\Psi}}^{(0)}_{n,k+\Lambda} \quad , \quad \vec{\overline{\Psi}}^{(0)}_{n,-k-\Lambda} \quad , \quad E^{(0)}_{n,k} \tag{4.87c}$$

Analogous results hold if $2(\Lambda + k) = 0(\text{Mod } P)$. It is easily seen that higher order perturbations will not split the degeneracy any further.

If neither (4.85) or (4.86) are satisfied the four-fold degenerate level splits into two two-fold degenerate levels - although the splitting is zero in first order. The energy eigenvalues in second order are

$$E^{(0)}_{k,n} + \frac{|\langle k,n \mid g \mid k'n'\rangle|^2}{E^{(0)}_{n,k} - E^{(0)}_{k'n'}} \tag{4.88}$$

and

$$E^{(0)}_{n,k} \quad + \quad \frac{|\langle k,n \,|\, g^{*} \,|\, k'n'\rangle|^{2}}{E^{(0)}_{n,k} \; - \; E^{(0)}_{n'k'}} \tag{4.89}$$

The two degenerate eigenfunctions corresponding to (4.88) to first order are

$$\begin{pmatrix} \bar{\Psi}^{(0)}_{n,-k} \\ 0 \end{pmatrix} + \sum \frac{\langle n'k' \,|\, g^{*} \,|\, n,-k\rangle}{E^{(0)}_{nk} \; - \; E^{(0)}_{n'k'}} \begin{pmatrix} 0 \\ \bar{\Psi}^{(0)}_{n'k'} \end{pmatrix} \tag{4.90a}$$

$$\begin{pmatrix} 0 \\ \bar{\Psi}^{(0)}_{n,k} \end{pmatrix} + \sum \frac{\langle n'k' \,|\, g \,|\, nk\rangle}{E^{(0)}_{n,k} \; - \; E^{(0)}_{n'k'}} \begin{pmatrix} \bar{\Psi}^{(0)}_{n'k'} \\ 0 \end{pmatrix} \tag{4.90b}$$

while the two degenerate eigenfunctions corresponding to (4.89) are

$$\begin{pmatrix} \bar{\Psi}^{(0)}_{n,k} \\ 0 \end{pmatrix} + \sum \frac{\langle n'k' \,|\, g^{*} \,|\, nk\rangle}{E^{(0)}_{nk} \; - \; E^{(0)}_{n'k'}} \begin{pmatrix} 0 \\ \bar{\Psi}^{(0)}_{n'k'} \end{pmatrix} \tag{4.91a}$$

$$\begin{pmatrix} 0 \\ \bar{\Psi}^{(0)}_{n,-k} \end{pmatrix} + \sum \frac{\langle n'k' \,|\, g \,|\, n,-k\rangle}{E^{(0)}_{nk} \; - \; E^{(0)}_{n'k'}} \begin{pmatrix} \bar{\Psi}^{(0)}_{n'k'} \\ 0 \end{pmatrix} \tag{4.91b}$$

If k is zero (Mod P) then expressiond (4.88) and (4.89) are identical -nthe degeneracy not being lifted by the perturbation. The two wavefunctions to first order corresponding to the degenerate level are given by (4.90b) and (4.91a).

So far we have ignored accidental degeneracies. These will occur, however, if we take

$$f(r,\varphi) = \mathcal{E}(X_o) + Ar^2 \qquad (4.92)$$

as in the harmonic approximation. Then the energy eigenvalue of $\mathcal{H}^{(0)}$ are of the form

$$\mathcal{E}(X_o) + (n+1)\hbar\omega , \quad n = o, 1, 2, . \; . \qquad (4.93)$$

where

$$\omega = \sqrt{\frac{2A}{M}} \qquad (4.94)$$

The n-th energy level is n+1 -fold degenerate - the values of k being ℓn, $\ell(n-2)$, ..., $-\ell n$. This degeneracy will be accidental if $n \geqslant 2$.

For illustrative purposes we discuss the case $n = 2$. The energy level $3\hbar\omega + \mathcal{E}(X_o)$ of $\mathcal{H}^{(0)}$ will be six-fold degenerate - four of the states being of the form given by (3.181) with $k=2\ell$ and the remaining two being

$$\begin{pmatrix} \bar{\Psi}^{(0)}_{2,0} \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ \bar{\Psi}^{(0)}_{2,0} \end{pmatrix}$$

The function $\bar{\Psi}^{(0)}_{2,0}$ is independent of φ and may be taken to be real. The values of j corresponding to these six states are $2\ell + \Lambda$, $\Lambda - 2\ell$, $2\ell - \Lambda$, $-2\ell - \Lambda$, Λ and $-\Lambda$, respectively. These states will not mix under the perturbation unless any two of these values of j are identical (Mod P). This will happen only if

$$2\ell = \pm\Lambda \quad (\text{Mod } P) \tag{4.95a}$$

or

$$2\ell = \pm 2\Lambda \quad (\text{Mod } P) \tag{4.95b}$$

If one of the relations (4.95a) is true we encounter the same situation discussed previously. The two states for which $j=0$ may be chosen in the forms (4.87a) and (4.87b) and will not mix with any of the remaining four degenerate states.

If one of the relations (4.95b) is true we must take appropriate linear combinations of our states. If $2\ell = 2\Lambda$ the states $j = 2\ell + \Lambda$ and $j = -\Lambda$ will mix as will those for which $j = -2\ell - \Lambda$ and $j = \Lambda$. The resulting states, which will consist of two degenerate pairs, will be of the following form in zeroth order

$$\begin{pmatrix} a\,\psi^{(0)}_{2,-2\ell} \\ b\,\psi^{(0)}_{2,0} \end{pmatrix} \qquad \begin{pmatrix} b^{*}\psi^{(0)}_{2,0} \\ a^{*}\psi^{(0)}_{2,2\ell} \end{pmatrix} \tag{4.96a}$$

$$\begin{pmatrix} a^{*}\psi^{(0)}_{2,0} \\ -b^{*}\psi_{2,2\ell} \end{pmatrix} \qquad \begin{pmatrix} -b^{*}\psi^{(0)}_{2,-2\ell} \\ a^{*}\psi_{2,0} \end{pmatrix} \tag{4.96b}$$

The coefficients a and b as well as the first order corrections to the energy for the states (4.96a) and (4.96b), respectively, are given by

$$\begin{bmatrix} 0 & \langle 2,-2\ell \mid g \mid 2,0\rangle \\ \langle 2,0 \mid g^* \mid 2,-2\ell\rangle & 0 \end{bmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = E^{(1)} \begin{pmatrix} a \\ b \end{pmatrix} \tag{4.97}$$

where

$$E^{(1)} = \pm \left| \langle 2,-2\ell \mid g \mid 2,0\rangle \right| \tag{4.98}$$

Thus the six-fold degenerate level splits in first order into three two-fold degenerate levels. Higher order perturbations will not remove this degeneracy. Detailed calculations of this very sort were made by Renner. In the example he considered $\ell = 1$, $\Lambda = 1$, $g = Br^2 e^{-2i}$ and our formulae above apply directly.

If the first order splitting of the degenerate level vanishes for reasons other than the symmetry considerations given here one must proceed to second order. The procedure is analogous - only one must subject $\mathbb{H}$ to a unitary transformation in order that we may directly apply first order perturbation procedures to the resulting Hamiltonian. Let

$$\widehat{\mathbb{H}} = e^{-\mathcal{S}} \mathbb{H} e^{\mathcal{S}} \tag{4.99}$$

where $\mathcal{S}$ is an operator (assumed to be of first order) such that

$$[\mathcal{S}, \mathbb{H}^{(0)}] = \mathbb{H}^{(1)} \tag{4.100}$$

and $\mathcal{S} = -\mathcal{S}^{+}$. Then the transformed Hamiltonian is

$$\hat{\mathbb{H}} = \mathbb{H}^{(0)} - \tfrac{1}{2}[\, \mathcal{S} \,,\, \mathbb{H}^{(1)}] \qquad (4.101)$$

to second order. It is quickly seen that σ_z commutes with $\hat{\mathbb{H}}$. For since σ_z anticommutes with $\mathbb{H}^{(1)}$ and commutes with $\mathbb{H}^{(0)}$ it follows that $\mathcal{S}$ may be chosen to anticommute with σ_z and then σ_z will commute with $[\, \mathcal{S} \,,\, \mathbb{H}^{(1)}]$. It thus follows that we may choose our unperturbed states to be eigenstates of σ_z and that states of the form $\begin{pmatrix} \psi \\ 0 \end{pmatrix}$ will not mix with states of the form $\begin{pmatrix} 0 \\ \psi \end{pmatrix}$ under the perturbation.

The above remarks apply in particular to the Moffitt-Liehr quasi-Hamiltonian

$$K + \mathcal{E}(X_0) + Ar^2 + Br \begin{pmatrix} 0 & e^{-i\varphi} \\ e^{i\varphi} & 0 \end{pmatrix} \qquad (4.102)$$

Here $\mathcal{S}$, as was discovered by Moffitt and Thorson [Ref. 32], may be found directly

$$\mathcal{S} = \mathbb{H}^{(1)} \frac{1}{M\omega^2} \left(\frac{1}{r} \frac{d}{dr} + \frac{1}{r^2} \sigma_z i \frac{d}{d\varphi} \right) \qquad (4.103)$$

and therefore

$$[\, \mathcal{S} \,,\, \mathbb{H}^{(1)}] = 2B^2 \frac{1}{M\omega^2} \left(1 - \sigma_z i \frac{d}{d\varphi} \right) \qquad (4.104)$$

The Hamiltonian (4.101) then becomes

$$\hat{\mathbb{H}} = K + \mathcal{E}(X_0) + Ar^2 - \frac{B^2}{M\omega^2} \left(1 - \sigma_z i \frac{d}{d\varphi} \right) \qquad (4.105)$$

Since σ_z and $i(d/d\varphi)$ commute with $\mathbb{H}^{(0)}$ the eigenstates of $\mathbb{H}^{(0)}$ may be chosen such that they are also eigenstates of $\hat{\mathbb{H}}$. The resulting energy levels are then

$$E_{n,m(\pm)} = \hbar\omega(n+1) - \frac{B^2}{M\omega^2}(1 \pm m) + \mathcal{E}(X_o) \qquad (4.106)$$

The two degenerate eigenstates of $\mathbb{H}$ to first order which correspond to the energy value given by the plus sign in (4.106) are

$$(1 + \mathcal{S})\, e^{im\varphi} \begin{pmatrix} R_{n,m}(r) \\ 0 \end{pmatrix} \qquad (4.107a)$$

$$(1 + \mathcal{S})\, e^{-im\varphi} \begin{pmatrix} 0 \\ R_{n,m}(r) \end{pmatrix} \qquad (4.107b)$$

where $R_{n,m}$ is the solution of

$$\left\{ -\frac{\hbar^2}{2M} \left[\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr} - \frac{m^2}{r^2} \right] + Ar^2 \right\} R_{n,m} = (n+1)\,\hbar\omega\, R_{n,m} \qquad (4.108)$$

The states corresponding to the negative sign can be found by merely replacing m by -m in (4.107).

Let us now find the eigenstates and eigenvalues of (4.102) in the strong coupling limit. Here

$$\frac{g}{|g|} = e^{-i\varphi} \qquad (4.109)$$

and

$$(K,\ g/|g|) = \frac{\hbar^2}{2Mr^2}\, e^{-i\varphi} \left(2i \frac{d}{d\varphi} + 1\right) \qquad (4.110)$$

Thus the expressions $\mathcal{H}'^{(0)}$ and $\mathcal{H}'^{(1)}$ defined by (4.65) and (4.67) are

$$\mathcal{H}'^{(0)} = K + \mathcal{E}(X_o) + \frac{\hbar^2}{4Mr^2}\left(1 + 2i\sigma_z \frac{d}{d\varphi}\right) + Ar^2 + Br\sigma_z \tag{4.111}$$

and

$$\mathcal{H}'^{(1)} = \frac{\hbar^2}{4Mr^2} \begin{pmatrix} 0 & -e^{-i\varphi}\left[2i\frac{d}{d\varphi} + 1\right] \\ e^{i\varphi}\left[-2i\frac{d}{d\varphi} + 1\right] & 0 \end{pmatrix} \tag{4.112}$$

If we let

$$\psi^{(0)II} = e^{im\varphi} R^{II}_m(r)$$

we find $R^{II}_m(r)$ satisfies

$$\left(-\frac{\hbar^2}{2M}\left[\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr}\right] + \frac{\hbar^2}{2Mr^2}\left(m + \frac{1}{2}\right)^2 + \mathcal{E}(X_o) + Ar^2 - Br\right) R^{II}_m = E^{(0)II} R^{II}_m \tag{4.113}$$

This equation has been solved by Longuet-Higgins et al. [Ref. 6] who find the solution in the limit

$$B \gg \sqrt[4]{\frac{A^3 \hbar^2}{M}} \tag{4.114}$$

(i. e., the strong coupling limit) to be approximately

$$R^{(0)II}_{n,m} = \frac{1}{r}\mathcal{R}^{II(0)}_n(r - r_o) \tag{4.115}$$

with the associated eigenvalue

$$E^{II(0)}_{n,m} = \mathcal{E}(X_o) - \frac{B^2}{4A} + \left(n + \frac{1}{2}\right)\hbar\omega - \frac{\hbar^2}{8Mr_o^2} + \frac{\hbar^2}{2Mr_o^2}\left(m + \frac{1}{2}\right)^2 \tag{4.116}$$

Here r_o is the value of r for which Ar^2-Br is a minimum

$$r_o = \frac{B}{2A} \tag{4.117}$$

and the function $\mathcal{R}_n^{II(0)}(x)$ is the solution of the equation

$$\left(- \frac{\hbar^2}{2M} \frac{d^2}{dx^2} + Ax^2\right) \mathcal{R}_n^{(0)II}(x) = \epsilon_n \mathcal{R}_n^{(0)II}(x) \tag{4.118}$$

which vanishes at $x = - r_o$. The eigenvalue ϵ_n is approximately $(n+1)\hbar\omega$ with the assumption (4.114) and $\mathcal{R}_n^{II(0)}$ coincides with the corresponding harmonic oscillator eigenfunction in the region where x is small. Therefore the zeroth order eigenfunctions of $\mathcal{H}'$ which correspond to the low lying energy states are given approximately by

$$\psi_{n,m}^{(0)} = \frac{e^{im\varphi}}{\sqrt{r}} \mathcal{R}_n^{II(0)}(r-r_o) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{4.119}$$

and the energy eigenvalues by (4.116).

Let us now use the approximation (4.73) to obtain the eigenstates of $\mathcal{H}'$ to first order. We take

$$\Delta = \frac{B^2}{4A} = n\mathrm{Min}\,\mathcal{E}^{I} - \mathrm{Min}\,\mathcal{E}^{II} \tag{4.120}$$

Thus (4.73) becomes

$$\vec{\psi}_{n,m} = \frac{1}{\sqrt{r}} \mathcal{R}_n^{(0)II}(r - r_o)\, e^{im\varphi} \begin{pmatrix} \frac{\hbar^2 4A}{2Mr^2B^2}(1-2m)e^{-i\varphi} \\ 1 \end{pmatrix} \tag{4.121}$$

The probability of the system being in the upper electronic state is approximately

$$\left\{ \frac{\hbar^2}{2Mr_0^2} \quad \frac{4A}{B^2} (1-2m) \right\}^2 \quad = \quad \frac{8\hbar^4 A^3}{MB^4} (1-2m)^2 \qquad (4.122)$$

The approximate state which is degenerate with (4.121) is obtained by replacing m by -m-1. In the next chapter we shall make use of (4.121) to estimate the electronic current carried by a low energy state in the strong-coupling case.

E. Summary.

Thus, in this chapter, we have used symmetry considerations to find the possible degeneracies of the eigenvalues and the general forms of the quasi-Hamiltonians obtained in Chapter 3. We have classified the various forms of the two potential surfaces which become degenerate at the equilibrium configuration and have shown under what circumstances may one expect the ground state to be two-fold degenerate. From considering the possible forms of the potential surfaces we have distinguished two limiting cases - strong-coupling and weak-coupling - and have examined the forms of the eigenfunctions of the vibronic Hamiltonians in both limits.

Chapter 5

ELECTRIC CURRENT AND THE BORN-RENNER APPROXIMATION

An interesting question is whether or not there is a net current flow about the symmetry axis of a molecule when it is in a pure quantum mechanical state. If the Born-Oppenheimer approximation were rigorously applicable there would never be any current as the electronic state could always be taken to be purely real. As we shall see, however, the consideration of electron-vibration interaction allows the possibility of our choosing any two degenerate states such that they carry non-zero electronic current in opposite directions.

Our motivation in discussing this topic is that of examining whether or not the molecular models discussed in the previous chapters will have some features which, when incorporated into solid state models, will explain the nature and properties of the ground state of a superconductor. From examining the predictions following from his formulation of the macroscopic thermodynamics of equilibrium, Tisza (Ref. 8) has advanced the view that the ground state may be doubly degenerate and that the two degenerate states may be chosen such that they carry currents in opposite directions.

Before proceeding, we shall clarify the question as to whether or not this view is in contradiction to the Bloch theorem (see Shafroth, Ref. 44). This theorem states that a system in thermodynamic equilibrium cannot carry a spontaneous current and, in particular, that the ground state (the state of thermodynamic equilibrium at absolute zero) cannot carry a spontaneous current. This theorem may be interpreted in either of two ways. In the narrow sense it may be interpreted as implying that the ground state is always non-degenerate. This view seems to be tacitly assumed in much of the existing literature. However, a wider interpretation may be given in which the theorem does not exclude the possibility of degenerate ground states. Instead the theorem implies that - if the ground state is degenerate - the weights of the two states carrying currents in opposite directions must be taken to be equal - such that the net current is zero. Thus, in the wider interpretation, Tisza's view is not in contradiction with the Bloch theorem.

The electric current flowing about the symmetry axis of a molecule is formally the expectation value of the operator

$$\langle I \rangle_{op} = \frac{1}{2\pi} \frac{\hbar e}{2m} \sum \frac{1}{\rho_i^2} \frac{d}{d\theta_i} \qquad (5.1)$$

where e and m^o are the charge and mass of an electron, θ_i and ρ_i are the angular and radial coordinates of an electron in cylindrical coordinates (the symmetry axis is taken as the z-axis).

Let us suppose that the wavefunction of the molecule is given by the Born-Renner approximation rather than by the Born-Oppenheimer approximation and that the situation studied in the previous chapters prevails. Only one two-fold degenerate vibration is considered and the electronic state is degenerate at $r = 0$. The wavefunction is therefore taken to be of the form

$$\bar{\Psi}_{\Lambda}\,\bar{\Phi}_{\Lambda} + \bar{\Psi}_{-\Lambda}\,\bar{\Phi}_{-\Lambda} \qquad (5.2)$$

where $\bar{\Phi}_{\Lambda}$ and $\bar{\Phi}_{-\Lambda}$ are the functions satisfying the conditions of theorems I and II, and $\bar{\Psi}_{\Lambda}$ and $\bar{\Psi}_{-\Lambda}$ are functions of r and φ only. Matrix elements of $\langle I\rangle_{op}$ between two states of the form (5.2) are then

$$\int\Big\{\bar{\Psi}'^{*}_{\Lambda}\,\bar{\Psi}_{\Lambda}\;\Big[\int\bar{\Phi}^{*}_{\Lambda}\,I_{op}\,\bar{\Phi}_{\Lambda}\,d\tau_e\Big]$$
$$+\;\bar{\Psi}'^{*}_{-\Lambda}\,\bar{\Psi}_{\Lambda}\;\Big[\int\bar{\Phi}^{*}_{-\Lambda}\langle I\rangle_{op}\,\bar{\Phi}_{\Lambda}\,d\tau_e\Big]$$
$$+\;\bar{\Psi}'^{*}_{\Lambda}\,\bar{\Psi}_{-\Lambda}\;\Big[\int\bar{\Phi}^{*}_{\Lambda}\langle I\rangle_{op}\,\bar{\Phi}_{-\Lambda}\,d\tau_e\Big]$$
$$+\;\Psi'^{*}_{-\Lambda}\,\bar{\Psi}_{-\Lambda}\;\Big[\int\bar{\Phi}^{*}_{-\Lambda}\,I_{op}\,\bar{\Phi}_{-\Lambda}\,d\tau_e\Big]\Big\}\,r\,dr\,d\varphi \qquad (5.3)$$

It may be readily shown, however, that

$$\int\bar{\Phi}^{*}_{\Lambda}\langle I\rangle_{op}\,\bar{\Phi}_{-\Lambda}\,d\tau_e$$
$$= \int\bar{\Phi}^{*}_{-\Lambda}\langle I\rangle_{op}\,\bar{\Phi}_{\Lambda}\,d\tau_e = 0 \qquad (5.4)$$

by partial integration and the fact that

$$\bar{\Phi}_{-\Lambda}^{*} = \bar{\Phi}_{\Lambda}.$$

Furthermore

$$\int \bar{\Phi}_{\Lambda}^{*} \langle I \rangle_{op} \, \bar{\Phi}_{\Lambda} \, d\tau_{e} = - \int \bar{\Phi}_{-\Lambda}^{*} \langle I \rangle_{op} \, \bar{\Phi}_{-\Lambda} \, d\tau_{e} \qquad (5.5)$$

If we define

$$I(r, \varphi) = \int \bar{\Phi}_{\Lambda}^{*} \langle I \rangle_{op} \, \bar{\Phi}_{\Lambda} \, d\tau_{e} \qquad (5.6)$$

then (5.3) may be rewritten in the form

$$\int (\bar{\Psi}'^{*}_{\Lambda}, \bar{\Psi}'^{*}_{-\Lambda}) \quad \sigma_{z} \, I(r, \varphi) \begin{pmatrix} \bar{\Psi}_{\Lambda} \\ \bar{\Psi}_{-\Lambda} \end{pmatrix} r \, dr \, d\varphi \qquad (5.7)$$

- matrix multiplication being implied in the integrand. This shows that the expectation value of the electric current carried by an eigenstate of the Moffitt-Liehr quasi-Hamiltonian is the expectation value of the operator

$$I(r, \varphi) \, \sigma_{z} \qquad (5.8)$$

The function $I(r, \varphi)$ is readily seen to have all the symmetry properties of the equilibrium configuration. In particular

$$I(r, \varphi + \ell \alpha) = I(r, \varphi) \qquad (5.9a)$$

$$I(r, -\varphi) = I(r, \varphi) \qquad (5.9b)$$

$$I(r, \varphi) = I(r, \phi) \qquad (5.9c)$$

It is well known that a non-degenerate state cannot carry an electric current. In our formalism this follows from the fact that the expectation value of (5.8) with respect to any eigenstate of $\sigma_x(*)$ or $\sigma_x C$ canishes - for any non-degenerate eigenstate of the Moffitt-Liehr quasi-Hamiltonian must be an eigenstate of both of these operators as they are invariants. We prove this for the case of $(*)\sigma_x$, taking

$$\sigma_x(*)\begin{pmatrix}\bar{\Psi}_{\Lambda}\\ \bar{\Psi}_{-\Lambda}\end{pmatrix} = a\begin{pmatrix}\bar{\Psi}_{\Lambda}\\ \bar{\Psi}_{-\Lambda}\end{pmatrix} \qquad (5.10)$$

where

$$a\,a^{*} = 1$$

Then

$$\begin{aligned}\int \vec{\bar{\Psi}}^{\dagger} I\,\sigma_z\,\vec{\bar{\Psi}}\,r\,dr\,d\varphi &= \int (a^{*}\vec{\bar{\Psi}}^{\dagger} I\sigma_z a\,\vec{\bar{\Psi}})^{T}\,r\,dr\,d\varphi\\ &= \int\left\{[(*)\sigma_x\vec{\bar{\Psi}}]^{\dagger} I\sigma_z(*)\sigma_x\vec{\bar{\Psi}}\right\}^{T} r\,dr\,d\varphi\\ &= \int \vec{\bar{\Psi}}^{*T} I\,\sigma_x\sigma_z\sigma_x\,\vec{\bar{\Psi}}\,r\,dr\,d\varphi\\ &= -\int \vec{\bar{\Psi}}^{\dagger} I\sigma_z\,\vec{\bar{\Psi}}\,r\,dr\,d\varphi\\ &= 0 \qquad (5.11)\end{aligned}$$

as $\sigma_x\sigma_z\sigma_x = -\sigma_z$.

Another interesting fact is that a linear combination of any two degenerate states can always be taken in general such that the resultant state is a current carrying state - regardless of the particular form of the Moffitt-Liehr quasi-Hamiltonian. This shall be verified by considering some limiting cases.

Let us first find that linear combination of two degenerate states which carries the maximum current. We can always choose the two degenerate states $\overset{\rightleftharpoons}{\Psi}_j$ and $\overset{\rightleftharpoons}{\Psi}_{-j}$ such that

$$\begin{pmatrix} e^{i\Lambda\alpha} & 0 \\ 0 & e^{-i\Lambda\alpha} \end{pmatrix} \overset{\rightleftharpoons}{\Psi}_j = e^{ij\alpha} \overset{\rightleftharpoons}{\Psi}_j \tag{5.12}$$

$$\sigma_x \, (*) \, \overset{\rightleftharpoons}{\Psi}_j = \overset{\rightleftharpoons}{\Psi}_{-j} \tag{5.13}$$

$$C \, (*) \, \overset{\rightleftharpoons}{\Psi}_j = \overset{\rightleftharpoons}{\Psi}_j \tag{5.14}$$

where j is an integer not equal to 0 (Mod $\mathcal{B}$). The matrix of $I(r, \varphi)\sigma_z$ between these two states is

$$\begin{pmatrix} \langle j \mid I(r,\varphi)\sigma_z \mid j\rangle & \langle j \mid I(r,\varphi)\sigma_z \mid -j\rangle \\ \langle -j \mid I(r,\varphi)\sigma_z \mid j\rangle & \langle -j \mid I(r,\varphi)\sigma_z \mid -j\rangle \end{pmatrix} \tag{5.15}$$

This may be shown to be diagonal as

$$\langle j \mid I(r,\varphi)\sigma_z \mid -j \rangle = 0 \tag{5.16}$$

follows from (5.12) and the fact that

$$I(r, \varphi + \mathcal{L}\alpha) = I(r, \varphi)$$

Furthermore

$$\langle j \mid I(r,\varphi)\sigma_z \mid j\rangle = - \langle -j \mid I(r,\varphi)\sigma_z \mid -j \rangle \tag{5.17}$$

follows from considerations similar to those of (5.11) and from the fact that $I(r, \varphi)$ is real. We shall define

$$I_j = \langle j \mid I(r,\varphi)\sigma_z \mid j \rangle \tag{5.18}$$

and therefore write (5.15) in the form

$$\begin{pmatrix} I_j & 0 \\ 0 & -I_j \end{pmatrix} \qquad (5.19)$$

If $\overset{\rightleftarrows}{\Psi}$ is any normalized linear combination of $\overset{\rightleftarrows}{\Psi}_j$ and $\overset{\rightleftarrows}{\Psi}_{-j}$

$$\overset{\rightleftarrows}{\Psi} = a\overset{\rightleftarrows}{\Psi}_j + b\,\overset{\rightleftarrows}{\Psi}_{-j}$$

where

$$|a|^2 + |b|^2 = 1$$

then the expectation value of $I(r,\varphi)$ is readily seen to be

$$(|a|^2 - |b|^2)I_j$$

However

$$\big||a|^2 - |b|^2\big| \leqslant |a|^2 + |b|^2 = 1.$$

Thus the maximum current that can be carried by a linear combination of the two degenerate states is I_j and the states $\overset{\rightleftarrows}{\Psi}_j$ and $\overset{\rightleftarrows}{\Psi}_{-j}$ satisfying (5.12-15) are the states carrying the maximum current.

Let us note that our symmetry considerations give no reason to expect I_j to be zero. This suggests that one can always take a linear combination of two degenerate states and obtain a current carrying state. We confirm this by considering the two limiting cases discussed in the previous chapter.

Consider first the weak coupling case. If there are no accidental degeneracies of K + f the degenerate eigenstates of the Moffitt-Liehr quasi-Hamiltonian which satisfy (5.12-15) may be taken to be approximately of the forms (4.83a) and (4.83d), $j = k + \Lambda$. Then

$$I_j = \int |\bar{\psi}^{(0)}_{n,k}|^2 \, I(r, \varphi) \, r dr \, d\varphi \qquad (5.20)$$

This can be expected to vanish only for very special forms of $I(r, \varphi)$ and we can therefore assume I_j being zero to be extremely unlikely. If we approximate $I(r, \varphi)$ by a constant

$$I(r, \varphi) \doteq I_o \qquad (5.21)$$

then $I_j = I_o$ as we have taken the wavefunction $\psi^{(0)}_{n,k}$ to be normalized. This approximation will be legitimate if the functions $\bar{\Phi}_\Lambda$ and $\bar{\Phi}_{-\Lambda}$ are very nearly independent of r - as might be expected when the spacing between the degenerate level $\mathcal{E}(X_o)$ of H_{X_o} and the other levels of H_{X_o} is very large (See Appendix A.).

If the levels of K+f are accidentally degenerate and states with equivalent values of j correspond to the same level analogous results are obtained. If the degenerate states are not mixed by the first order perturbation the correct zeroth order states are still of the forms (4.83a) and (4.83d) and expression (5.20) still approximately applies.

If, on the other hand, the states mix in first order we cannot take our zeroth order states in this form. As an example, we consider the case discussed in the previous chapter (equations (4.96), (4.97), (4.98) where n=2. Taking the solution of (4.97) to be $a = 1/\sqrt{2}$, $b = 1/\sqrt{2}$ the correct zeroth order wavefunction becomes

$$\frac{1}{\sqrt{2}} \begin{pmatrix} \bar{\Psi}^{(0)}_{2,-2\ell} \\ \bar{\Psi}^{(0)}_{2,0} \end{pmatrix}$$

with a current equal to

$$I_{j=-\Lambda} \overset{\circ}{=} \frac{1}{2} \int I(r,\varphi) \left[\left| \bar{\Psi}^{(0)}_{2,-2\ell} \right|^2 - \left| \bar{\Psi}^{(0)}_{2,0} \right|^2 \right] r\,dr\,d\varphi \qquad (5.22)$$

This expression is zero if I is identically a constant. However, I is never identically a constant and we expect that this expression is merely small. Furthermore, if higher order corrections to the state were included the resulting expression for the current would be non-zero even if I were identically a constant.

The strong coupling case is the most interesting for here one would ordinarily approximate the low energy states by the Born-Oppenheimer approximation. Although the molecule is almost entirely confined to the electronic state associated with the lower quasi-potential surface it will have a component in the upper state which is not identically zero. This fact enables a degenerate state to carry a current - which, although small, is in general non-zero.

For an estimate of the current carried by a degenerate state in the strong coupling limit we take our two degenerate eigenstates of $\mathbb{H}'$ in the approximate form (4.73)

$$\begin{pmatrix} 0 \\ \psi^{(0)II}_{k,n} \end{pmatrix} + \frac{1}{\Delta}(K,\ g/|g|)\begin{pmatrix} \psi^{(0)II}_{k,n} \\ 0 \end{pmatrix} \tag{5.23a}$$

$$\begin{pmatrix} 0 \\ -\frac{g^*}{|g|}\psi^{(0)II}_{k,n} \end{pmatrix} - \frac{1}{\Delta}(K,\ g/|g|)\begin{pmatrix} \frac{g^*}{|g|}\psi^{(0)II*}_{k,n} \\ 0 \end{pmatrix} \tag{5.23b}$$

The corresponding eigenstates of $\mathbb{H}$ may be obtained by applying $\mathbb{A}$ given by (4.62) to these states. We find that (5.23a) is transformed into

$$\psi_j = \frac{1}{\sqrt{2}}\begin{pmatrix} -\frac{g}{|g|}\psi^{(0)II}_{k,n} \\ \psi^{(0)II}_{k,n} \end{pmatrix} + \frac{1}{\Delta\sqrt{2}}\begin{pmatrix} (K,\ \frac{g}{|g|})\psi^{(0)II}_{k,n} \\ \frac{g^*}{|g|}(K,\ \frac{g}{|g|})\psi^{(0)II}_{k,n} \end{pmatrix} \tag{5.24}$$

where $j = k-\Lambda$. The expectation value of $I(r,\varphi)$ with respect to this state is then

$$I_j = -\frac{2}{\Delta}\ \mathcal{R}e\int I(r,\varphi)\psi^{(0)II}_{k,n}\ \frac{g^*}{|g|}(K,\ \frac{g}{|g|})\psi^{(0)II}_{k,n}\,r\,dr\,d\varphi \tag{5.25}$$

Let us apply this relation to the special case of the Moffitt-Liehr quasi-Hamiltonian (4.102) which was discussed in the preceding chapter. Here $\ell=2$, and

$$\frac{g}{|g|}^{\blacktriangle}\left(K,\ \frac{g}{|g|}\right) = \frac{\hbar^2}{2Mr^2}\left(2i\,\frac{d}{d\varphi} + 1\right) \qquad (5.26)$$

Taking $\psi_{k,n}^{(0)II}$ in the form

$$\frac{1}{\sqrt{2\pi}}\, e^{im\varphi}\, R(r)$$

where m is k/2 expression (5.25) becomes

$$I_J = -\frac{\hbar^2}{M\Delta}\,\frac{(1-2m)}{2\pi}\int I(r,\varphi)\; R(r)^2\; \frac{1}{r^2}\, r dr\, d\varphi \qquad (5.27)$$

Since R(r) is appreciably different from zero only in the region where $r = r_o$ we may approximate this expression by taking

$$\frac{1}{2\pi}\int I(r,\varphi)\; R(r)^2\; \frac{1}{r^2}\, r dr\, d\varphi$$

$$\doteq \frac{1}{2\pi r_o^2}\int I(r_o,\varphi)\, d\varphi \int R(r)^2 r dr$$

$$\doteq \frac{1}{2\pi r_o^2}\int I(r_o,\varphi)\, d\varphi \qquad (5.28)$$

as

$$\int R(r)^2\, r dr = 1$$

is our normalization condition. As a further approximation we may take

$$\frac{1}{2\pi}\int I(r_o,\varphi)\, d\varphi = I_o$$

giving

$$I_j = \frac{\hbar^2 j}{M \Delta r_0^2} I_o \tag{5.30}$$

Although this expression was obtained with the particular Moffitt-Liehr vibronic equation (4.102) in mind it is clear that it can apply equally as well to any Moffitt-Liehr quasi-Hamiltonian of the forms (3.71) or (3.78) - providing that Δ is sufficiently large and we are interested only in the current carried by the low lying energy states. While the approximations made were somewhat rough we expect that the relation (5.30) will generally give a reasonable order of magnitude estimate of the current.

One interesting feature of (5.30) is its dependence on j, the total internal angular momentum of the system. That I_j is proportional to j shows that the current is proportional to the angular momentum. If we consider the system from a classical viewpoint as moving along the bottom of the trough of $f - |g|$ we reach the interesting conclusion that such nuclear motion literally <u>drags</u> an electric current along with it - the magnitude of the electric current being roughly proportional to the nuclear angular momentum. Let us note that this current is <u>not</u> carried by the nuclei - but by the electrons. It is the nuclear motion, however, that induces this current.

One feature not considered in our discussion of the current carried in the strong coupling case is that the trough of the lower energy surface generally has "ripples", i. e., the function $f - |g|$ is not independent of φ. This is the case, for example, in Fig. 5. The system, however, will have at least three (since $\ell \neq 0$, $2\ell \neq 0 (\mathrm{Mod}\ P)$) equivalent configurations at which the lower surface is a minimum. If the ripples are sufficiently small we may ignore them. In any event, however, the system may be considered as resonating among the equivalent equilibrium configurations. The fact that we must have three or more such configurations allows the possibility of degenerate vibrational states and one may choose these states such that, figuratively speaking, the distorted equilibrium configuration revolves around the symmetry axis (although the individual nuclei remain in the vicinity of their original equilibrium configurations). Here, just as in the examples considered previously, we expect the "rotation" of the new configuration to induce an electric current. It is to be expected, however, that the expression (5.30) will be a poor estimate of the current if the ripples are large.

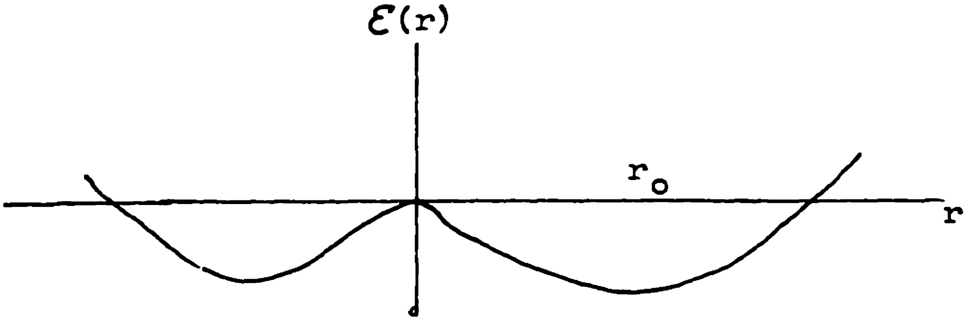
Thus we have shown that degenerate eigenstates of the Born-Renner or Moffitt-Liehr quasi-Hamiltonians generally can always be chosen such that they carry currents in opposite directions. Furthermore, in the last chapter we have shown that in many cases the ground state is degenerate and thus in these cases the ground state wavefunctions can generally be taken as states carrying currents in opposite directions.

A well-known model of a superconductor originally discussed by Fröhlich (Ref. 9) may be analyzed in a manner similar to that given above. This is a model of a one-dimensional lattice. Frohlich's original discussion of this model showed that it has low energy current carrying states arising because of electron-lattice interaction. That this feature may be interpreted in a manner similar to that of the molecular case described above is not evident from Frohlich's discussion but becomes clear if we parallel the steps for the molecular case. This we have done in Appendix D.

There we find that there is an analogue of the Jahn-Teller effect in this model and that the lowest quasi-potential surface attains its minimum when the atoms are not equally spaced. The form of this surface as a function of the radial coordinate r of a two-fold degenerate vibration whose wavenumber is twice that of an electron at the Fermi level is given by

$$\mathcal{E} = \mathcal{E}(X_o) + Ar^2 + Cr^2 \ln r \qquad (5.31)$$

where A and C are constants. This curve is sketched below



It thus has a form similar to that of Fig. 4a in Chapter 4. Since this potential does not depend on the vibrational angular coordinate the molecular configuration will revolve around the trough of the surface when the system is in its low energy states. Since, however, the system cannot be identically in a single electronic state this motion will induce an electric current - just as in the molecular case.

Using approximations similar to those used in deriving (5.30) we may estimate this current and may write it in a form identical to (5.30):

$$I_j = \frac{\hbar^2 j}{M \Delta r_o^2} I_o$$

where r_o is the value of r at which the lowest quasi-potential surface attains its minimum and Δ is the spacing between the minimum of the lowest surface and that of the first excited electronic state. For Frohlich's model, however, I_o may be expressed in terms of constants appropriate to the energy band model of a solid:

$$I_o = - \frac{\pi}{2} \frac{v_F \, q_e \, N_e}{L} e^{-1/2} e^{- \frac{9}{8} \frac{N}{N_e} \frac{M}{m} (\hbar\omega/C)^2} \qquad (5.32)$$

where v_F is the velocity of an electron at the Fermi level, q_e is the charge on an electron, N_e/L is the number of electrons per unit length, N_e/N is the number of electrons per nucleus, M and m are the masses of the nuclei and electrons, respectively, ω is the vibrational frequency, and C is the familiar Sommerfeld-Bethe parameter giving the strength of the electron-lattice coupling. Δ and r_o may similarly be expressed in terms of the same constants - this is also done in Appendix D.

The Jahn-Teller-type effect exhibited by this model does not appear in two and three dimensions and it was essentially for this reason that Frohlich never generalized his theory to higher dimensional models. Frohlich's one-dimensional theory is now apparently significant only from the standpoint of the history of the development of the more recent theories but it is nevertheless interesting to note that the mechanism which Frohlich ascribed as being responsible for superconductivity is exactly that which we have been discussing in this chapter - in particular, that encountered in the strong-coupling case.

Chapter 6

A NEW HAMILTONIAN FOR ELECTRON-LATTICE INTERACTION IN SOLIDS

A. Introduction.

The problem to be considered here is that of constructing the analogue of the Born-Renner quasi-Hamiltonian for a solid. One possible method of doing this is to consider the solid as a single large molecule and to approximate the wavefunction for the solid by a Born expansion of the form

$$\sum_n \Phi_n(x, X)\ \psi_n(X) \tag{6.1}$$

where the Φ_n are electronic states for the solid as a whole and the sum runs over a sufficiently large number of electronic states.

This approach to the problem presents two severe difficulties which we list below:

1) In the case of metals, the number of electronic states needed in the Born expansion is extremely large and increases exponentially with the size of the crystal. The resulting quasi-Hamiltonian will therefore appear in the form of a matrix operator of very large dimension and would seem to be unmanageable.

2) The form of the resulting quasi-Hamiltonian appears to be strongly dependent on the size of the crystal. If one were to obtain the quasi-Hamiltonian for a crystal of given dimensions he would apparently find it difficult to generalize this quasi-Hamiltonian to one for a crystal of, say twice the dimensions.

A similar example in which both these difficulties would in principle appear is the well-known Heisenberg model for ferromagnetism where the quasi-Hamiltonian for a crystal of N atoms is a 2^N by 2^N matrix. If such a matrix were written down explicitly for a given N it would appear unmanageable for all but the smallest N and the relationship between matrices corresponding to different N would appear, at the very best, to be extremely subtle. On the other hand, if this matrix were written in the form of an expansion in terms of spin matrices both of these difficulties would be overcome. The relation between matrices corresponding to different N would appear almost trivial and, as is well-known, the quasi-Hamiltonian is amenable to various approximation methods and is in no way unmanageable.

This example of the Heisenberg model suggests that similar considerations might be successfully applied in the writing of the quasi-Hamiltonianffor a solid along the lines of the Born-Renner approximation. Let us remember that, in addition to being considered as a single large molecule, a solid may also be considered as an aggregate of small molecules. Just how these molecules are chosen is to a certain extent arbitrary - although, obviously, it is most advantageous to take these molecules as each comprising an integral number of the unit cells of the crystal - such that the lattice formed by these molecules exhibits the periodicsstructure of the crystal. The simplest choice of these molecules is clearly that in which each molecule comprises a unit cell of the crystal . This sort of viewpoint may be considered as the starting point for the formulation of such well-known physical theories as the theory of lattice vibrations in insulators, the Heisenberg theory of ferromagnetism, the theory of excitons in molecular crystals, etc.. Nevertheless, it is always possible to choose these molecules as comprising more than one unit cell.

Considering the crystal as an aggregate of molecules provides the tie-in with our previous discussion of the Born-Renner approximation applied to molecules. The quasi-Hamiltonian for a molecule restricted to two electronic states has two predominant characteristics: (1) the coordinates of the electrons do not appear explicitly and (2) the quasi-Hamiltonian is a two-by-two matrix. Thus it may always be written in the form

$$H(P, X, \vec{\sigma}) \tag{6.2}$$

i. e., as a function of the nuclear momenta (P), the nuclear coordinates (X), and the Pauli spin matrices ($\vec{\sigma} = \sigma_x, \sigma_y, \sigma_z$). This interpretation of the quasi-Hamiltonian as being a function of the spin matrices follows from the fact that any two-by-two matrix may be written as a linear combination of the unit matrix and the three Pauli matrices.

If one could, in some sense, completely uncouple the molecules in different cells such that the state of each cell would be independent of those of its neighbors and if one could consider each molecule as being restricted to two electronic states it is clear that the quasi-Hamiltonian for the solid as a whole would be simply the sum of the quasi-Hamiltonians for the individual molecules

$$H = \sum_{n} H(P_n, X_n, \vec{\sigma}_n) \tag{6.3}$$

where the sum extends over all cells comprising the crystal. Here P_n represents the momenta of the nuclei in the n-th cell while X_n represents their displacements relative to their equilibrium positions. The $\vec{\sigma}_n$ are again matrices with the commutation and anti-commutation properties of the Pauli spin matrices - here, however, they are 2^N-by-2^N matrices, where N is the total number of cells. The spins corresponding to different cells commute. Thus, in this limiting case, the Born-Renner quasi-Hamiltonian may be written in such a manner that both the difficulties mentioned previously are circumvented.

Only in molecular crystals may one expect the quasi-Hamiltonian given by (6.3) to even be a fair approximation to the actual Hamiltonian obtained from an ansatz of the form (6.1). In most cases the cells are strongly coupled and it cannot be assumed that the states of molecules in different cells are independent of one another. Nevertheless, these considerations do suggest that there may exist a method of writing the Born-Renner quasi-Hamiltonian in more general cases in such a way that we overcome our previously mentioned difficulties.

This method rests on what might be called the concept of vibronic spin. It is clear that a generalization of the quasi-Hamiltonian (6.3) which formally accounts for the coupling between cells is one of the form

$$H(P_1,X_1,\vec{\sigma}_1,\ P_2,X_2,\vec{\sigma}_2,\ .\ .\ .\ .\ .,\ P_N,X_N,\vec{\sigma}_N\ ) \qquad (6.4)$$

Here we have taken the Hamiltonian to be a function of the momenta and coordinates of all the nuclei in the crystal and of N (the total number of cells) spin vectors. From analogy with (6.3) it is to be assumed that spin vectors with different indices commute. Thus (6.4), just as was the case for (6.3), represents a 2^N-by-2^N matrix. Any 2^N-by-2^N matrix may be written as a function of N commuting spin vectors and our Hamiltonian (6.4) in the absence of any additional restrictions would be of an extremely general form.

However, one very stringent restriction is implied by the manner in which we wrote the Hamiltonian - associating each spin vector with a particular cell. Our Hamiltonian should exhibit the symmetry of the crystal. As we found in the molecular case there exists for every symmetry property of the molecule an operator which commutes with the quasi-Hamiltonian. Let us recall also that in the case of the quasi-Hamiltonian resulting from the Born-Oppenheimer approximation the symmetry operation corresponding to translational invariance was a permutation of particles. For a linear lattice, for example, the Born-Oppenheimer quasi-Hamiltonian would be invariant if one replaced the displacements (relative to their equilibrium positions) and momenta of the particles in each cell by those of the cell to its right (periodic boundary conditions being assumed, as usual). The symmetry of the Hamiltonian (6.3) is an obvious generalization of the symmetry of the Born-Oppenheimer quasi-Hamiltonian. Here the Hamiltonian is invariant under permutation of cells - providing one permutes the spins of the cells as well as the nuclear displacements and momenta. It is clear that such symmetry is a very desirable requirement for our Hamiltonian (6.4). Thus we should require that this Hamiltonian be invariant under all permutations of cell coordinates (displacements, momenta, and spin) which correspond to the translational symmetries of the lattice formed by the cells. This requirement appears as a necessity if we are to adopt the viewpoint of the vibronic spin being a <u>true</u> internal coordinate of the cell.

Similar restrictions should be placed on the Hamiltonian for every additional symmetry of the lattice and the possible forms of the Hamiltonian will be narrowed with each restriction. In each case we should be guided by the principle that the vibronic spin behave under symmetry operations as a true internal coordinate of the cell.

A further symmetry desired in our quasi-Hamiltonian is that it be invariant under translating each nucleus by the same amount in the same direction. This is also satisfied by the Born-Oppenheimer quasi-Hamiltonians and is necessary if the energy of solid is not to depend on the absolute position of the solid in space. This requirement is in accord with the view that the vibronic spin is associated with the atoms making up the molecule and not with the equilibrium position of the molecule.

An additional requirement comes from the following considerations. If the Hamiltonian (6.3) were considered as a special case of the Hamiltonian (6.4) (which indeed it is) it is readily seen that it has the property

$$\frac{1}{N} H(P,X,\vec{\sigma}, P,X,\vec{\sigma}, . \quad . \quad . \quad ., P,X,\vec{\sigma})$$

$$= H(P,X,\vec{\sigma}) \qquad (6.5)$$

i. e., if the momenta and displacements of corresponding particles in each cell as well as the vibronic spin vectors are formally taken as being equal and the resulting Hamiltonian is divided by the total number of cells one will obtain the Hamiltonian for a single molecule (that for the case N=1). This requirement is also satisfied by the Born-Oppenheimer quasi-Hamiltonian for a crystal - providing one interprets the Hamiltonian for a single cell as being that obtained by applying the method of periodic boundary conditions to a minature crystal of the size of a single cell. It is also satisfied by the quasi-Hamiltonian for the Heisenberg model of ferro-magnetism, and, indeed, by any quasi-Hamiltonian for a model of a crystal based on the concept of the crystal consisting of an aggregate of coupled molecules. It is therefore practically mandatory that this requirement be satisfied by the general quasi-Hamiltonian (6.4)[1]. Thus, if the Hamiltonian for a single cell is

---

[1] It appears that there is an ambiguity in the requirement (6.5) as the expression would seem to depend on the order in which the factors in terms such as

$$\sigma_{1x}\ \sigma_{2y}$$

are taken - for when one sets

$$\vec{\sigma}_1 = \vec{\sigma}_2 = \vec{\sigma}$$

this becomes

$$\sigma_{1x}\sigma_{2y} \rightarrow i\sigma_z$$

known in advance, a great deal of information about the Hamiltonian for an arbitrary number of cells is already known and a stringent requirement is placed on this Hamiltonian.

All of these requirements are reasonable requirements that would be expected of any Hamiltonian for a solid which depicts the crystal as a lattice of molecules - each molecule having in addition to the usual classical coordinates of the nuclei a non-classical

---

While

$$\sigma_{2y}\sigma_{1x} \rightarrow -i\sigma_z$$

In this particular instance this ambiguity is spurions, however, when remembers that the quasi-Hamiltonians must always be Hermitian. We must write

$$\sigma_{1x}\sigma_{2y} = (1/2)(\sigma_{1x}\sigma_{2y} + \sigma_{2y}\sigma_{1x})$$

Thus

$$\sigma_{1x}\sigma_{2y} \rightarrow 0$$

One encounters difficulties, however, when he considers products of the spins of three or more cells, i. e.,

$$\sigma_{1x}\sigma_{2y}\sigma_{3y}$$

These difficulties might be avoided by assuming that the quasi-Hamiltonian never contains terms of this type - as would be the case if we limit it to interactions between neighboring cells only. If we are unable to make this assumption, however, there is no clear-cut answer to the question. Analogous problems are encountered when one attempts to quantize classical expressions. For example, how would one quantize the expression x^2p_x? In terms of quantum-mechanical operators both xp_xx and $(1/2)(x^2p_x + p_xx^2)$ would appear to be equally valid. Similarly $(1/2)(\sigma_{1x}\sigma_{2y}\sigma_{3y} + \sigma_{2y}\sigma_{3y}\sigma_{1x})$ and $\sigma_{2y}\sigma_{1x}\sigma_{3y}$ both go over into Hermitian operators when all the spins are formally set equal - but one expression is the negative of the other. This ambiguity must be kept in mind when applying (6.5) coordinate which we have called vibronic spin. In addition, it would also be desirable to make several simplifying assumptions about the form of the quasi-Hamiltonian which are intuitively appealing and which correspond to the usual assumptions made concerning the Born-Oppenheimer quasi-Hamiltonians. One might assume, for example, that only neighboring cells interact. One would expect the interaction between two cells to be small if they are sufficiently removed from each other and such an assumption would furnish the simplest case in accord with such an expectation. Another assumption would be that the quasi-Hamiltonian contains terms of at most second order in the nuclear displacements, i. e., that cubic, quartic, etc., terms may be neglected. Such an assumption is permissible as long as the displacements of the nuclei from their equilibrium positions are sufficiently small.

The question we wish to consider in this chapter is whether or not it is possible to construct a model quasi-Hamiltonian meeting these requirements. If the cells of the crystal are chosen to be a single unit cell of the crystal the answer to this question turns out to be trivially yes. The case, however, when each cell comprises more than one unit cell is much more interesting. Then the quasi-Hamiltonian should have additional symmetries which do not correspond to a mere permutation of cells.

If we make the apparently reasonable assumption that the symmetry elements of the quasi-Hamiltonian are a product of two factors - one operating only on a spin coordinate and the other operating only on the nuclear coordinates - the problem reduces to finding a representation of the crystal group in terms of the spin matrices which has certain properties.

Our deliberations are carried out for the case of a monatomic linear lattice - chosen because of its relative simplicity compared to other models of a solid. Our analysis for this model indicates that it may not be possible to construct a quasi-Hamiltonian satisfying all the requirements listed above if one takes his basis cell to contain more than one unit cell. The chief difficulty occurs because of simultaneously requiring that only neighboring cells interact and requiring that the quasi-Hamiltonian exhibit the full symmetry of the lattice.

While one may always relax the requirement of nearest neighbor interactions only it becomes apparent that the resulting model will be too complicated to serve as a useful model for electron-lattice interaction in solids. A prime requirement on our model quasi-Hamiltonian shall be simplicity. The other possibility - of relaxing the requirement that the quasi-Hamiltonian display the full symmetry of the lattice - proves much more fruitful. Also, there are physical arguments and precedents which justify doing this. These shall be discussed in the course of our development of the theory.

B. Properties of the Spin Matrices.

Since many of the relationships which we shall derive in this chapter will depend on the properties of the spin matrices it is appropriate that we begin by reviewing some of their properties. Detailed proofs shall be omitted for brevity.

We may always define 3N matrices σ_{1x}, σ_{1y}, σ_{1z}, σ_{2x}; σ_{2y}, σ_{2z}, ..., σ_{Nx}, σ_{Ny}, σ_{Nz}, each having 2^N rows and 2^N columns which have the following properties:

$$\sigma_{nx}\sigma_{nx} = \sigma_{ny}\sigma_{ny} = \sigma_{nz}\sigma_{nz} = 1 \quad (n = 1, 2, ..N) \qquad (6.6)$$

$$[\sigma_{nx}, \sigma_{n'y}] = 2i\, \delta_{nn'}\, \sigma_{nz} \qquad (6.7a)$$

$$[\sigma_{ny}, \sigma_{n'z}] = 2i\, \delta_{nn'}\, \sigma_{nx} \tag{6.7b}$$

$$[\sigma_{nz}, \sigma_{n'x}] = 2i\, \delta_{nn'}\, \sigma_{ny} \tag{6.7c}$$

$$[\sigma_{nx}, \sigma_{n'x}] = [\sigma_{ny}, \sigma_{n'y}] = [\sigma_{nz}, \sigma_{n'z}] \tag{6.7d}$$

$$\begin{aligned}[\sigma_{nx}\sigma_{ny} + \sigma_{ny}\sigma_{nx}] &= [\sigma_{ny}\sigma_{nz} + \sigma_{nz}\sigma_{ny}] = \\ &= [\sigma_{nz}\sigma_{nx} + \sigma_{nx}\sigma_{nz}] = 0\end{aligned} \tag{6.7e}$$

$$\sigma_{ni} = \sigma_{ni}^{+} \tag{6.8}$$

This set of matrices is not uniquely defined for the set

$$\{\sigma'_{ni}\}$$

where

$$\sigma''_{ni} = \mathcal{U}\, \sigma_{ni}\, \mathcal{U}^{\dagger}$$

and $\mathcal{U}$ is any 2^N-by-2^N matrix will also satisfy the relations given. Conversely, if any two sets separately satisfy these relations there will always exist a unitary matrix such that corresponding members of the two sets are related by the above equation.

The trace of each of the spin matrices is zero and, furthermore, the trace of products of spin matrices is always zero

$$\mathrm{tr}\ \sigma_{ni}\sigma_{n'i'} \cdot \cdot \cdot \sigma_{n_r i_r} = 0 \tag{6.9}$$

if no two of the indices n, n', . . ., n_r are equal.

Any 2^N-by-2^N matrix may be expanded in terms of the spin matrices and their products:

$$\begin{aligned}\mathcal{A} = a^{(0)} &+ \sum a^{(1)}_{ni}\sigma_{ni} + \sum_{n>n'} a^{(2)}_{nin'i'}\sigma_{ni}\sigma_{n'i'} \\ &+ \cdot \cdot \cdot + \sum a^{(N)}_{N,i_N,\ldots,1,i_1}\sigma_{Ni_N} \cdot \cdot \cdot \sigma_{2i_2}\sigma_{1i_1}\end{aligned} \tag{6.10}$$

The coefficients in this expansion are unique, being given by

$$a^{(0)} = \frac{1}{2^N} \operatorname{tr} \mathcal{A}$$

$$a^{(1)}_{ni} = \frac{1}{2^N} \operatorname{tr}(\mathcal{A}\, \sigma_{ni})$$

$$a^{(2)}_{nin'i'} = \frac{1}{2^N} \operatorname{tr}(\mathcal{A}\, \sigma_{ni}\sigma_{n'i'})$$

etc.. If the matrix $\mathcal{A}$ is Hermitian the coefficients will be real and if $\mathcal{A}$ is an operator which is Hermitian the coefficients will be Hermitian operators. For brevity, we shall generally write such expansions. in the form

$$A(\vec{\sigma}_1, \vec{\sigma}_2, \ldots\ldots, \vec{\sigma}_N)$$

thus implying that the matrix is a function of the components of the N spin vectors.

If such a matrix were subjected to a similarity transformation it is evident from examining the expansion that the following relation holds

$$\mathcal{T} \mathcal{A} \mathcal{T}^{-1} = A(\mathcal{T} \vec{\sigma}_1 \mathcal{T}^{-1}, \ldots\ldots, \mathcal{T} \vec{\sigma}_N \mathcal{T}^{-1}) \qquad (6.11)$$

providing the elements of $\mathcal{T}$ commute with the elements of the matrix.

Certain operators constructed from the spin matrices will be of special interest. Such operators are the permutation operators

$$\mathbb{P}_{nn'} = \tfrac{1}{2}(1 + \vec{\sigma}_n \cdot \vec{\sigma}_{n'}) \qquad (6.12)$$

which have the following properties

$$\mathbb{P}_{nn'}\,\mathbb{P}_{nn'} = 1$$

$$\mathbb{P}_{nn'}\sigma_{ni} = \sigma_{n'i}\,\mathbb{P}_{nn'}$$

$$\mathbb{P}_{nn'}\sigma_{n'i} = \sigma_{ni}\,\mathbb{P}_{nn'}$$

and, in general,

$$\mathbb{P}_{nn'}A(\vec{\sigma}_1, \ldots,\vec{\sigma}_n,\ldots,\vec{\sigma}_{n'},\ldots.\vec{\sigma}_N)\,\mathbb{P}_{nn'}$$

$$= A(\vec{\sigma}_1,\ldots.,\vec{\sigma}_{n'}, \ldots.,\vec{\sigma}_n,\ldots.\vec{\sigma}_N) \qquad (6.13)$$

Thus the operator $\mathbb{P}_{nn'}$ corresponds to a permutation of the indices n and n'. From these operators one may construct the operator for cyclic permutations

$$\mathbb{B} = \mathbb{P}_{1,2}\,\mathbb{P}_{1,3}\cdot\;\cdot\;\cdot\;\cdot\mathbb{P}_{1,N} \qquad (6.14)$$

which has the properties

$$\mathbb{B}^N = 1$$

$$\mathbb{B}\,\sigma_{ni} = \sigma_{n-1,i}\,\mathbb{B} \qquad (n \neq 1)$$

$$\mathbb{B}\,\sigma_{1,i} = \sigma_{N,i}\,\mathbb{B}$$

or in general

$$\mathbb{B}\,A(\vec{\sigma}_1, \vec{\sigma}_2, \ldots.,\vec{\sigma}_N)\,\mathbb{B}^{-1}$$

$$= A(\vec{\sigma}_N, \vec{\sigma}_1, \vec{\sigma}_2; \;.\;\;.\;\;., \vec{\sigma}_{N-1}) \qquad (6.15)$$

C. Symmetries of the Quasi-Hamiltonian for a Monatomic Linear Lattice.

We shall consider the simplest type of linear lattice in which X_o (the tentative equilibrium configuration) corresponds to all the nuclei (assumed identical) being equally spaced a distance a apart. The lattice as a whole is divided into super cells of equal length - each containing an integer number g of unit cells. We assume that the lattice is composed of N super cells. The nuclei in the n-th super cell are denoted n,1, n,2, n,3, . . ., n,g while the displacement of the n,j nucleus from its equilibrium position is denoted $X_{n,j}$. The symbol X_n shall denote the set of displacements $(X_{n1}, X_{n2}, \ldots, X_{ng})$ of the nuclei in the n-th super cell. Analogous notational conventions shallbbe used for the momentum operators of the nuclei.

The quasi-Hamiltonian for this lattice, in accordance with the remarks made in the introduction, is assumed to be a 2^N-by2^N matrix operator and may therefore be written as a function of the spin matrices defined in the previous section:

$$H(P_1, X_1, \vec{\sigma}_1, P_2, X_2, \vec{\sigma}_2, \ . \ . \ ., P_N, X_N; \vec{\sigma}_N) \qquad (6.16)$$

The exact particle Hamiltonian involving the nuclear and electronic coordinates (from which, conceptually, the above Hamiltonian may be derived) has the following symmetries:

1) Translational invariance: If the displacement X_{nj} is replaced by that of the nucleus to its left and each of the electrons are translated through a distance a to the right, the Hamiltonian remains unchanged.

2) Reflectional invariance: If the displacement X_{nj} is replaced by the negative of that of the (N-n+1, g-j+1)-th nucleus and each electron coordinate x is replaced by L-x where L=Nga is the length of the lattice the Hamiltonian will remain invariant. (It is assumed that the origin for electron displacements lies a distance a/2 to the left of the (1,1)-th nucleus.)

3) Time-reversal invariance: The time reversal operator consisting of the product of the complex conjugation operator (*) with the product of the y components of the spin operators of all the electrons commutes with the Hamiltonian.

In addition, it must be borne in mind that products and powers of the operations corresponding to these symmetries will also commute with the Hamiltonian.

Operators built up from the operations of translational and reflectional symmetry will be referred to as operators corresponding to geometrical symmetry.

Neglecting for the present the time reversal operator it is immediately seen that all of the remaining symmetry operations of the particle Hamiltonian may be written as a product of two operators - one of which acts only on the nuclear coordinates and the second which acts only on the electron coordinates. The factor which acts on the nuclear coordinates turns out to be an invariant of the Born-Oppenheimer quasi-Hamiltonians and is of the major interest.

We define the following two operators T and C which act only on the nuclear coordinates and momenta

$$T\, X_{n,j} = X_{n,j-1} \quad (j \neq 1) \tag{6.16a}$$

$$T\, X_{n,1} = X_{n-1,g} \quad (n \neq 1) \tag{6.16b}$$

$$T\, X_{1,1} = X_{N,g} \tag{6.16c}$$

$$C\, X_{n,j} = -\, X_{N-n+1,g-j+1} \tag{6.17}$$

They are additionally defined such that the corresponding momentum operators transform analogously. These two operators are the appropriate factors operating on the nuclear coordinates which appear in the operators describing translational and reflectional symmetry of the particle Hamiltonian. For these operators the following relations can be readily shown to hold:

$$T^{Ng} = 1 \qquad (6.18)$$

$$C^2 = 1 \qquad (6.19)$$

$$CT = T^{-1}C \qquad (6.20)$$

In the molecular case it was found that it was generally preferable from the standpoint of symmetry considerations to take our quasi-Hamiltonian in the form of a Moffitt-Liehr quasi-Hamiltonian rather than a Born-Renner quasi-Hamiltonian. The symmetry invariants of the Moffitt-Liehr quasi-Hamiltonian which corresponded to geometrical symmetries turned out to be representable also as a product of two commuting operators - one operating on the nuclear coordinates and momenta and the other (being a two-by-two matrix) operating on the spin matrices. The factors operating on the nuclear coordinates and momenta in the invariants of the particle Hamiltonian and the Moffitt-Liehr quasi-Hamiltonian were identical.

Such simplicity is highly desirable for our quasi-Hamiltonian. Let us therefore assume that it is in a certain sense in the Moffitt-Liehr form such that its symmetry invariants can be factored into the product of a factor operating only on the nuclear displacements and momenta and a factor operating only on the spins. Assuming that there is an operator commuting with the quasi-Hamiltonian for each operator commuting with the particle-Hamiltonian we may take the operator corresponding to translation of the lattice through a to be of the form

$$T\,\Theta$$

and the operator corresponding to reflection in the form

$$C\,\Gamma$$

where Θ and Γ are 2^N-by-2^N matrices which commute with the nuclear displacements and momenta. These matrices must satisfy relations similar to those satisfied by T and C in order that they generate a group isomorphic to the symmetry group of the linear lattice

$$\Theta^{Ng} = 1 \tag{6.21}$$

$$\Gamma^2 = 1 \tag{6.22}$$

$$\Gamma\,\Theta = \Theta^{-1}\,\Gamma \tag{6.23}$$

Let us now consider the time reversal operator. In the molecular case the operator corresponding to time reversal was the product of the comples conjugation operator (*) and the matrix which appeared with C in the operator corresponding to reflectional symmetry. Thus, if we wish to preserve our analogy we must take the **time reversal operator to be**

$$(*)\Gamma$$

Our choosing of the invariants of the quasi-Hamiltonian to be of these forms may appear to be an ad hoc choice. However, it is possible to give this choice a firmer basis than that given above. If we were to take the wavefunction for the solid as a whole to be given by a sum of the form

$$\sum \psi_{nk}(X) \quad \varphi^{(0)}_{nk}(x) \tag{6.24}$$

where the $\varphi^{(0)}_{nk}(x)$ were eigenstates of H_{X_0} with the property

$$\varphi^{(0)}_{nk}(x+a) = e^{ik}\varphi^{(0)}_{nk}(x)$$

$$k = \frac{2\pi m}{N}, \text{ m integer}$$

for simultaneous translation of all the electrons through $\underline{a}$ and if the sum contained exactly 2^N terms we would find the quasi-Hamiltonian to be the sum of the total nuclear kinetic energy and a matrix whose elements were of the form

$$\langle k,n \mid H_X \mid k'n'\rangle = \int \varphi^{(0)}_{kn}(x)^* H_X(x,X)\, \varphi^{(0)}_{k'n'}(x)\, d\tau_e$$

The $\varphi^{(0)}_{kn}$ could be chosen such that

$$\varphi^{(0)}_{k,n}(L-x)^* = \varphi^{(0)}_{k,n}(x)$$

$$\varphi_{kn}^{(0)}(x)^{*} = \varphi_{-k,n}^{(0)}(x) \qquad k \neq 0,\pi$$
$$= \pm\ \varphi_{k,n}^{(0)}(x) \qquad k = 0,\pi$$

The plus sign corresponds to g (gerade) states while the minus sign corresponds to u (ungerade) states. The invariants of the quasi-Hamiltonian would then be $T\Theta$, $C\Gamma$, and $(*)\Gamma$ where Θ was the matrix with elements

$$\langle k\ n \,|\, \Theta \,|\, k'n' \rangle = \delta_{nn'}\ \delta_{k,k'}\ e^{-ik}$$

while Γ was the matrix with elements

$$\langle k\ n \,|\, \Gamma \,|\, k'n' \rangle = \delta_{n,n'}\ \delta_{k,-k'} \qquad k\neq 0,\pi$$
$$= -\ \delta_{n,n'}\delta_{k,k'} \qquad k = (0,u),\ (\pi,u)$$
$$= \delta_{n,n'}\delta_{k,k'} \qquad k = (0,g),\ (\pi,g)$$

It therefore appears that our choosing our invariants to be in the forms given is a realistic choice.

The quasi-Hamiltonian for a single cell has already been obtained in Section I of Chapter 3 and is given by (3.101). The operator R_α appearing there is the same as the operator T appearing in this chapter. Thus, for the case N=1 the matrices Γ and Θ are

$$\Theta = e^{i(2\pi\nu/g)\sigma_z}$$
$$\Gamma = \sigma_x$$

For the case of arbitrary N we may expect the matrices Θ and Γ to be in a certain sense generalizations of the above expressions. In addition, we should impose the requirement that these matrices be unitary. This is satisfied for the case N=1 and is necessary if matrices of the form $\Theta \mathcal{A} \Theta^{-1}$ are to be Hermitian if $\mathcal{A}$ is Hermitian. Additional restrictions are imposed since the spatial operations of translation or feflection must commute with the time reversal operator. Thus, in addition to their group-like properties, we must require the following for these matrices

$$\Theta^{+} = \Theta^{-1} \tag{6.24}$$

$$\Gamma^{+} = \Gamma^{-1} \tag{6.25}$$

$$\Gamma = \Gamma^{*} \tag{6.26}$$

$$\Gamma \Theta \Gamma = \Theta^{*} \tag{6.27}$$

Furthermore, these together with the group properties imply that

$$\Gamma = \Gamma^{+} \tag{6.28}$$

$$\Theta^{*} = \Theta^{+} \tag{6.29}$$

It is readily seen that all of the above relations are satisfied for the case N=1.

Let us now note that if we apply the operator $T^{g}\,\Theta^{g}$ to the quasi-Hamiltonian (6.16) we obtain

$$H(P_N, X_N, \Theta^{g}\vec{\sigma}_1\Theta^{-g}, P_1, X_1, \Theta^{g}\vec{\sigma}_2\Theta^{-g}, \;.\;.\;., P_{N-1}, X_{N-1}, \Theta^{g}\vec{\sigma}_N\Theta^{-g})$$

$$= H(P_1, X_1, \vec{\sigma}_1, P_2, X_2, \vec{\sigma}_2, \ldots, P_N, X_N, \vec{\sigma}_N)$$

Thus, if we desire the vibronic spin to transform as a true internal coordinate it is clear that we must require

$$\Theta^{g}\,\sigma_1\,\Theta^{-g} = \sigma_N$$

$$\Theta^{g}\,\sigma_n\,\Theta^{-g} = \sigma_{n-1}$$

Then the Hamiltonian will be invariant under cyclic interchange of the coordinates (nuclear displacements and momenta and vibronic spin) of the super cells. The above equations will be satisfied if we take

$$\Theta^{g} = \mathbb{B} \tag{6.30}$$

where $\mathbb{B}$ is the operator defined by (6.14). It may be shown that it is the only operator, apart from a multiplicative factor, satisfying the above relations.

Since Θ is unitary its g-th power must also be unitary. This requirement is satisfied, however, by the matrix $\mathbb{B}$. We must also require that the complex conjugate of $\mathbb{B}$ be equal to its Hermitian conjugate as this is also satisfied by Θ. This can be accomplished only if we impose additional requirements on our representation of the spin matrices. The simplest requirement accomplishing this is to take

$$\sigma^{*}_{nx} = \sigma_{N-n+1,x} \tag{6.31a}$$

$$\sigma^{*}_{ny} = - \sigma_{N-n+1,y} \qquad (6.31b)$$

$$\sigma^{*}_{nz} = \sigma_{N-n+1,z} \qquad (6.31c)$$

It is clearly satisfied for the case $N = 1$ and, furthermore, it is always possible to choose our representation such that this is so[1]. Further justification for this choice shall be given shortly.

Now let us consider the effect of applying the operator $C\boldsymbol{\Gamma}$ to the quasi-Hamiltonian. If we define the operator $\bar{C}$ by the relations

$$\bar{C}\, X_{nj}\, \bar{C} = -X_{n,g-j+1}$$

$$\bar{C}\, P_{nj}\, \bar{C} = -P_{n,g-j+1}$$

we find that

$$\begin{aligned} C\boldsymbol{\Gamma}\, H\, C\boldsymbol{\Gamma} &= H(.\ \ .\ \ .P_n, X_n, \vec{\sigma}_n,\ .\ \ .\ \ .) \\ &= H(.\ \ .\ \ .\bar{C}P_{N-n+1}\bar{C},\ \bar{C}X_{N-n+1}\bar{C},\ \boldsymbol{\Gamma}\vec{\sigma}_n\boldsymbol{\Gamma}\ ,\ .\ \ .) \end{aligned}$$

---

[1] The proof of this is straightforward but somewhat messy. One may convince himself that he need only consider the case $N=2$. The spin matrices may then be written out in a standard representation in which the x and z components of both spin vectors are real and the y components are purely imaginary. A unitary matrix $\mathcal{A}$ is then found such that $\mathcal{A}^{\mp}\mathcal{A} = \mathbb{P}_{12}$ where $\mathbb{P}_{1,2}$ is given by (6.12). The representation $\{\sigma'_{nj}\}$ where

$$\sigma'_{n,j} = \mathcal{A}\sigma_{n,j}\mathcal{A}^{\dagger}$$

will then satisfy the requirements (6.31)

The operation $\bar{C}$ corresponds to a reflection within a super cell. It does not mix displacements and momenta of nuclei in different super cells. Thus, if we are to maintain our conceptual view of the vibronic spin being an internal coordinate of the super cell it is clear that we must require the components of $\Gamma\vec{\sigma}_n\Gamma$ to be functions of the components of the N-n+1 cell only. This requirement will be satisfied if we take

$$\Gamma = A(\vec{\sigma}_1)A(\vec{\sigma}_2)\cdot\cdot\cdot A(\vec{\sigma}_N)\ \mathbb{P}_{N,1}\ \mathbb{P}_{N-1,2}\cdot\ \cdot\ \cdot$$

where the $_{n,n'}$ are defined by (6.12) and the $A(\sigma_n)$ are non singular matrices which may be expanded in the form

$$A(\vec{\sigma}_n) = A_o + \sum A_i\sigma_{n,i}$$

the coefficients being independent of n. It may be readily verified that the relation

$$\Gamma\mathbb{B} = \mathbb{B}^{-1}\Gamma$$

is satisfied with this general form for Γ. The requirement that its square be the unit matrix, furthermore, is satisfied by taking

$$A(\vec{\sigma}_n)A(\vec{\sigma}_n) = 1$$

It may be shown that any matrix Γ' which has the property

$$\Gamma'\vec{\sigma}_n\Gamma' = A(\vec{\sigma}_{N-n+1})\ \vec{\sigma}_{N-n+1}\ A(\vec{\sigma}_{N-n+1})$$

is a constant times our expression for Γ. Thus the form given for Γ is the most general form which maintains the principle of vibronic spin being a true internal coordinate of the super cell and which preserves the equivalence of the super cells.

The matrices $A(\vec{\sigma}_n)$ may be determined by considering the case N = 1. If our matrix Γ is to be a generalization of the matrix for this case it is clear that our only logical choice for $A(\vec{\sigma}_n)$ is

$$A(\vec{\sigma}_n) = \sigma_{nx}$$

Thus our matrix Γ must be of the form

$$\Gamma = \sigma_{1x} \cdot \cdots \cdot \sigma_{Nx} \, \mathcal{P}_{N,1} \, \mathcal{P}_{N-1,2} \cdot \; \cdot \; \cdot \qquad (6.32)$$

It is to be noted that the requirement $\Gamma = \Gamma^T$ is satisfied with this choice. Furthermore, the matrix will be entirely real as required if the conditions (6.31) on the choice of the representation of the spin matrices are imposed.

D. The Matrix Θ.

Let us note that as of yet we have no guarantee that a matrix Θ satisfying all the properties requested in the preceding section actually exists. If the number g of unit cells in a super cell is taken to be identically unity the matrix Θ is simply $\mathcal{B}$. A more interesting question is that of the form of Θ when g is not identically one.

In seeking such a matrix we discover at the outset that it does not exist if both g and N are even. This may be seen by considering the case for N=2. Here $\mathcal{B}$ may be shown to be equivalent totthe matrix

$$\begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{bmatrix}$$

Since Θ and Γ must both commute with $\mathcal{B}$ when N=2 they will both be reduced to the following form when $\mathcal{B}$ is in the form given above

$$\Gamma = \begin{pmatrix} & \mathcal{X} & & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \\ 0 & 0 & 0 & x \end{pmatrix} \qquad \Theta = \begin{pmatrix} & \mathcal{Y} & & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \\ 0 & 0 & 0 & y \end{pmatrix}$$

where $\mathcal{X}$ and $\mathcal{Y}$ are three-by-three matrices. Since $\Gamma^2 = 1$, x^2 must be unity, and since $\Gamma \Theta \Gamma = \Theta^{+}$, y must be real. However, if we satisfy $\Theta^g = \mathcal{B}$ we must take $y^g = -1$. No real y will accomplish this if g is even.

As our results must be valid for an arbitrary number of super cells it is apparent that if we are to require that our Hamiltonian be invariant under the operation $T\Theta$ and hence that Θ exist we must require at the outset that g be an odd integer.

Returning to the problem of determining Θ we now seek to find a matrix which is unitary, symmetric, and satisfies

$$\Gamma \Theta \Gamma = \Theta^{\dagger}$$

$$\Theta^{g} = \mathbb{B}$$

and which reduces to $e^{i(2\pi\Lambda/g)\sigma_z}$ for the case $N = 1$.

A plausible form for this matrix which satisfies these conditions is

$$\Theta = e^{i2\pi(\Lambda/g)(\sigma_{1z} + \sigma_{2z} + \cdot\cdot + \sigma_{Nz})} M(\mathbb{B}) \qquad (6.33)$$

where $M(\mathbb{B})$ is a function of $\mathbb{B}$ only, i.e., expandable in a power series in the matrix $\mathbb{B}$. Since $M(\mathbb{B})$ commutes with

$$\exp i \frac{2\pi}{g} \Lambda[\sigma_{1z} + \cdots\cdot + \sigma_{Nz}]$$

and the latter in unitary our form for Θ will be unitary if $M(\mathbb{B})$ is unitary. Furthermore Θ^{g} will be $\mathbb{B}$ if $M(\mathbb{B})^{g} = \mathbb{B}$. Also $\Theta^{\dagger}$ will be Θ^{*} if $M(\mathbb{B})^{\dagger} = M(\mathbb{B})^{*}$. It is clear that our form for Θ represents a generalization of that for the single molecule case as $M(\mathbb{B})$ must be identically the unit matrix for $N=1$ if the above conditions are satisfied. We therefore seek a matrix $M(\mathbb{B})$ with the properties listed above. The most general form for such a matrix may be shown to be

$$M(\mathbb{B}) = \frac{1}{N} \sum_{s=1}^{N} \sum_{n=1}^{N} a_s e^{i(2\pi sn/N)} \mathbb{B}^{n} \qquad (6.34)$$

with

$$a_s^{g} = e^{-i(2\pi s/N)} \tag{6.35a}$$

$$a_s^{*} = a_s^{-1} \tag{6.35b}$$

$$a_s^{*} = a_{N-s} \qquad s=1,\ldots\ldots,N-1 \tag{6.35c}$$

$$a_N^{*} = a_N \tag{6.35d}$$

the a_s being numbers.

Such a set of numbers will always exist - providing g is an odd integer as we have already specified. The numbers a_N and $a_{N/2}$ must be identically 1 and -1 as there are no other solutions. If s is not identically N or N/2 there is still some ambiguity in the a_s - the most general solution being

$$a_s = e^{-i\,\theta_s} = a_{N-s}^{*} \qquad 1 \leq s < N/2 \tag{6.36}$$

where

$$g\,\theta_s = \frac{2\pi}{N} \ (\mathrm{Mod}\ 2\pi)$$

The matrix $M(\mathbb{B})$ for the cases N=2 and N=3 is then

$$M(\mathbb{B}) = \mathbb{B} \qquad (N=2) \tag{6.37}$$

$$M(\mathbb{B}) = (1/3)\left\{[1 + 2\cos(\theta - \tfrac{2\pi}{3})]\,\mathbb{B} + [1+2\cos(\theta + \tfrac{2\pi}{3})]\,\mathbb{B}^2 + [1 + 2\cos\theta]\right\} \qquad (N=3) \tag{6.38}$$

where $g\theta = \frac{2\pi}{3}$ (Mod 2π).

Thus a matrix Θ satisfying all the desired requirements may always be found if g is odd. While it may be doubtful that our choice of Θ in the form (6.33) is unique there appears to be no alternative form of any reasonable simplicity.

E. The Quasi-Hamiltonian for the Case N=2.

Let us now determine the form of the quasi-Hamiltonian for two coupled super cells which has the invariants $T\Theta$, $C\mathbb{L}$, and $\boldsymbol{\Gamma}(\boldsymbol{\kappa})$ and which reduces to the quasi-Hamiltonian (3.101) when the coordinates of the two cells are formally set equal and the resultant is divided by two. The matrices Θ and $\boldsymbol{\Gamma}$ shall be taken to be in the forms exhibited in the preceding sections.

The general quasi-Hamiltonian satisfying these requirements will be somewhat complex and it is therefore convenient to make some additional assumptions and obtain the quasi-Hamiltonian in a simple form before considering more complicated terms. The assumptions we shall make are the following:

1) The quasi-Hamiltonian is of the general form

$$K + U(X,\vec{\sigma}_1,\vec{\sigma}_2)$$

where K is the kinetic energy operator and U does not depend on the nuclear momenta.

2) The coefficients of the quadratic terms in the nuclear coordinates are independent of the spin coordinates. This implies that the constant B in (3.101) is assumed zero.

The first assumption is equivalent to that of neglecting the matrix $\mathcal{T}$ in the single molecule case. The second assumes that the electron-lattice interaction is linear in the nuclear displacements and is fairly standard in most theories of electron-lattice interaction.

With these assumptions we may write down the quasi-Hamiltonian in the form

$$\begin{aligned}\mathcal{H} = \overline{K} &+ \sum_{j=2}^{g} \omega_o^2 (X_{1,j}-X_{1,j-1})^2 + \sum_{j=2}^{g} \omega_o^2 (X_{2,j}- X_{2,j-1})^2 \\ &+ \omega_o^2 (X_{1,1} - X_{2,g})^2 + \omega_o^2 (X_{2,1} - X_{1,g})^2 \\ &+ U^{(0)}(\vec{\sigma}_1, \vec{\sigma}_2) + \sum_{j=1}^{g} U_j^{(1)}(\vec{\sigma}_1, \vec{\sigma}_2)\, X_{1,j} \\ &+ \sum_{j=1}^{g} U_j^{(1)}(\vec{\sigma}_2, \vec{\sigma}_1)\, X_{2,j}\end{aligned}$$

Here we have already made use of the fact that the quadratic terms must commute with T and that they must reduce with

$$\sum_{j=1}^{g} \omega_o^2 \, (X_j - X_{j-1})^2$$

when the coordinates of corresponding nuclei in the cells are set equal and the resultant is divided by two. Also we have made use of the fact that the Hamiltonian must be invariant under interchanging coordinates of the two cells in stating that the coefficients of X_{1j} and X_{2j} must be related by an interchange of $\vec{\sigma}_1$ and $\vec{\sigma}_2$. For the same reasons we must require

$$1) \quad U^{(0)}(\vec{\sigma}_1,\vec{\sigma}_2) = U^{(0)}(\vec{\sigma}_2,\vec{\sigma}_1)$$

Requiring that the Hamiltonian commute with the invariants suggested in the previous sections gives us the following relations:

$$2) \quad \left[e^{i2\pi(\Lambda/g)(\sigma_{1z}+\sigma_{2z})} , U^{(0)}(\vec{\sigma}_1,\vec{\sigma}_2) \right] = 0$$

$$3) \quad U^{(1)}_{j-1}(\vec{\sigma}_1,\vec{\sigma}_2) = e^{i(2\pi\Lambda/g)(\sigma_{1z}+\sigma_{2z})} U_j^{(1)}(\vec{\sigma}_2,\vec{\sigma}_1) \cdot \cdot e^{-i(2\pi\Lambda/g)(\sigma_{1z}+\sigma_{2z})}$$

$$4) \quad \left[\sigma_{1x}\sigma_{2x} , U^{(0)}(\vec{\sigma}_1,\vec{\sigma}_2) \right] = 0$$

$$5) \quad \sigma_{1x}\sigma_{2x} U_j^{(1)}(\vec{\sigma}_1,\vec{\sigma}_2) = - U^{(1)}_{g-j+1}(\vec{\sigma}_1,\vec{\sigma}_2) \, \sigma_{1x}\sigma_{2x}$$

$$6) \quad \sigma_{1x}\sigma_{2x} U^{(0)}(\vec{\sigma}_1,\vec{\sigma}_2)^{*} \sigma_{1x}\sigma_{2x} = U^{(0)}(\vec{\sigma}_1,\vec{\sigma}_2)$$

$$7) \quad \sigma_{1x}\sigma_{2x} U_j^{(1)}(\vec{\sigma}_2,\vec{\sigma}_1)^{*} \sigma_{1x}\sigma_{2x} = U_j^{(1)}(\vec{\sigma}_1,\vec{\sigma}_2)$$

Additional requirements are

8) $U^{(0)}(\vec{\sigma}, \vec{\sigma}) = 2\,\mathcal{E}(X_o)$

9) $U_j^{(1)}(\vec{\sigma}, \vec{\sigma}) = A[\sigma_x \sin\frac{2\pi\Lambda}{g}(2j-1) + \sigma_y \cos\frac{2\pi\Lambda}{g}(2j-1)]$

as the Hamiltonian must reduce to twice the Hamiltonian for one cell when all corresponding coordinates of the two cells are formally set equal.

The most general form of these matrices $U^{(0)}$ and $U_j^{(1)}$ which satisfies these requirements listed above may be found after some labor to be

$$U^{(0)}(\vec{\sigma}_1, \vec{\sigma}_2) = 2\,\mathcal{E}(X_o) - 3a - b + a\vec{\sigma}_1\cdot\vec{\sigma}_2 + b\,\sigma_{1z}\sigma_{2z} + c(\sigma_{1x}\sigma_{2y} - \sigma_{1y}\sigma_{2x})$$

$$\begin{aligned} U_j^{(1)}(\vec{\sigma}_1, \vec{\sigma}_2) = \\ = \ & \tfrac{A}{2}[(\sigma_{1x} + \sigma_{2x})\sin\tfrac{2\pi\Lambda}{g}(2j-1) + (\sigma_{1y}+\sigma_{2y})\cos\tfrac{2\pi\Lambda}{g}(2j-1)] \\ & + (-1)^j d[(\sigma_{1x}-\sigma_{2x})\cos\tfrac{2\pi\Lambda}{g}(2j+1) + (\sigma_{1y}-\sigma_{2y})\sin\tfrac{2\pi\Lambda}{g}(2j+1)] \\ & + f\,[\,-(\sigma_{1x}\sigma_{2x} - \sigma_{1y}\sigma_{2y})\sin\tfrac{4\pi\Lambda}{g}(2j+1) + \\ & \quad + (\sigma_{1x}\sigma_{2y} + \sigma_{1y}\sigma_{2x})\cos\tfrac{4\pi\Lambda}{g}(2j+1)\,] \end{aligned}$$

where a, b, c, d, and f are constants which may not be determined by considering the single molecule case alone.

Thus our quasi-Hamiltonian for two coupled super cells is of the form:

$$
\begin{aligned}
\mathcal{H} = K &+ \sum_{j=2}^{g} \omega_o^2 (X_{1j} - X_{1,j-1})^2 + \sum_{j=2}^{g} \omega_o^2 (X_{2j} - X_{2,j-1})^2 \\
&+ \omega_o^2 (X_{1,1} - X_{2,g})^2 + \omega_o^2 (X_{2,1} - X_{1,g})^2 \\
&+ 2\mathcal{E}(X_o) - 3a - b + a\vec{\sigma}_1\vec{\sigma}_2 + b\sigma_{1z}\sigma_{2z} + c(\sigma_{1x}\sigma_{2y} - \sigma_{1y}\sigma_{2x}) \\
&+ \frac{A}{2} \sum_{j=1}^{g} \Big[(\sigma_{1x} + \sigma_{2x}) \sin\frac{2\pi\Lambda}{g}(2j-1) + (\sigma_{1y} + \sigma_{2y}) \cos\frac{2\pi\Lambda}{g}(2j-1) \Big] \\
&\qquad\qquad (X_{1j} + X_{2j}) \\
&+ d \sum_{j=1}^{g} (-1)^j \Big[(\sigma_{1x} - \sigma_{2x}) \cos\frac{2\pi\Lambda}{g}(2j+1) + (\sigma_{1y} - \sigma_{2y}) \sin\frac{2\pi\Lambda}{g}(2j+1) \\
&\qquad\qquad (X_{1j} - X_{2j}) \\
&+ f \sum_{j=1}^{g} \Big\{ -(\sigma_{1x}\sigma_{2x} - \sigma_{1y}\sigma_{2y}) \sin\frac{4\pi\Lambda}{g}(2j+1) \\
&\qquad + (\sigma_{1x}\sigma_{2y} + \sigma_{2x}\sigma_{1y}) \cos\frac{4\pi\Lambda}{g}(2j+1) \Big\} (X_{1j} + X_{2j})
\end{aligned}
\tag{6.39}
$$

We shall not give a detailed discussion of this quasi-Hamiltonian. Our primary purpose in obtaining it at the present was merely to show how one may apply the requirements we listed at the beginning of this chapter to find the quasi-Hamiltonian. Nevertheless, a few remarks about the term

$$a\vec{\sigma}_1\vec{\sigma}_2 + b\sigma_{1z}\sigma_{2z} + c(\sigma_{1x}\sigma_{2y} - \sigma_{1y}\sigma_{2x})$$

are appropriate for it represents a situation which may be expedted in more general cases. This term corresponds to the electron-electron interaction. Its eigenvalues are

$$a + b \text{ (two times)}$$
$$-(a+b) + 2(a^2+c^2)^{1/2}$$
$$-(a+b) - 2(a^2+c^2)^{1/2}$$

Thus, while we have assumed that the electronic state is two-fold degenerate for a single cell our method shows that in general the interaction between two cells will lead to a splitting of the original four-fold degeneracy (which would be the case if the cells were not coupled) into three levels - one of which is two-fold degenerate.

This feature is similar to that found in the Heisenberg model. The Hamiltonian for this model is derived by a Heitler-London method in which the electronic state of each atom is assumed to be two-fold degenerate (the spin degeneracy) - the interaction between atoms splitting this degeneracy.

The magnitude of the constant term

$$2\mathcal{E}(X_0) - 3a - b$$

in our quasi-Hamiltonian is irrelevant as was $\mathcal{E}(X_0)$ for the single molecule case. Since an absolute reference level for energy cannot be defined and since the properties of a solid depend only on the relative spacing of the energy levels the values of these constants are of no concern to us here and may be taken to be zero.

F. A New Model for Electron-Lattice Interaction in Solids.

Our primary concern in this chapter is to construct a model quasi-Hamiltonian for a solid which circumvents the two difficulties listed at the beginning of the chapter. In order that the first difficulty be overcome it is clear that the form of our quasi-Hamiltonian should be as simple as possible- but yet containing features which make it a non-trivial model for the study of electron-lattice interaction in solids. It is for this reason we desire the quasi-Hamiltonian to be restricted to only nearest neighbor interactions. In the event that we should find that we cannot satisfy all the assumptions we have previously made concerning the quasi-Hamiltonian this assumption shall take precedence.

This requirement obviously had no meaning for the cases N=1 and N=2. It becomes pertinent only when $N \geqslant 3$. In this section we shall determine a quasi-Hamiltonian for an arbitrary number of cells which satisfies this requirement. For simplicity we shall assume that the number of atoms g in each cell is equal to three. The same methods which we will use here may be used in treating cases when g is taken to be greater than three but we choose $g = 3$ for definiteness. The possible choices of the integer Λ giving the symmetry of the electronic states in the single molecule case are then only 1 and -1 with this choice of g. Since interchanging Λ with its negative amounts only to a relabeling of the electronic states no generality will be lost if we take Λ to be identically 1.

Of the various assumptions we have made concerning the quasi-Hamiltonian the one in which we have the least confidence is that of the form of the matrix Θ given by (6.33). We do know, however, that such a matrix exists for we have exhibited one which satisfies all the requirements requested. Let us therefore first determine the general form of the quasi-Hamiltonian in the absence that $T\Theta$ be an invariant. We shall assume, however, that $T^{g}\mathbb{B}$, $C\mathbb{\Gamma}$, and $\mathbb{\Gamma}(\blacktriangleleft)$ be symmetry operations of the quasi-Hamiltonian. In addition, we shall make the two assumptions made in the previous section - that the quasi-potential does not depend on the nuclear momenta and that the electron-lattice interaction term is linear in the nuclear displacements.

With the assumption of only nearest neighbor interactions the matrix coefficient of X_{nj} in the electron-lattice interaction term may be written in the general form

$$U_j^{(1)}(\vec{\sigma}_{n-1};\ \vec{\sigma}_n,\ \vec{\sigma}_{n+1}) = \vec{A}_j\cdot\vec{\sigma}_{n-1} + \vec{B}_j\cdot\vec{\sigma}_n + \vec{C}_j\vec{\sigma}_{n+1}$$
$$+\ \vec{\sigma}_{n-1}\ \mathbb{C}_j\cdot\vec{\sigma}_n + \vec{\sigma}_{n+1}\cdot\ \mathbb{D}_j\vec{\sigma}_n \qquad (6.40)$$

where the coeffieients in the expansion are all real. These must be independent of n in order that the quasi-Hamiltonian commute with $T^g\mathbb{B}$. There are only three such matrices (6.40) which we need consider - $j=1, 2, 3$ as we have chosen g to be three. Requiring that the quasi-Hamiltonian commute with $C\,\Gamma$ and $\Gamma(\bigstar)$ gives us the following relations between the coefficients in the expansion of these three matrices:

$$(\vec{A}_j)_z = (\vec{B}_j)_z = (\vec{C}_j)_z = (\mathbb{C}_j)_{xz} = (\mathbb{D}_j)_{xz}$$
$$= (\mathbb{C}_j)_{zx} = (\mathbb{D}_j)_{zx} = (\mathbb{C}_j)_{yz} = (\mathbb{D}_j)_{zy} = 0$$

$$(\vec{A}_1)_x = -(\vec{C}_3)_x \qquad (\vec{A}_3)_x = -(\vec{C}_1)_x$$
$$(\vec{A}_1)_y = (\vec{C}_3)_y \qquad (\vec{A}_3)_y = (\vec{C}_1)_y$$
$$(\vec{A}_2)_x = -(\vec{C}_2)_x \qquad (\vec{A}_2)_y = (\vec{C}_2)_y$$
$$(\vec{B}_2)_x = 0 \qquad (\vec{B}_1)_x = -(\vec{B}_3)_x$$
$$(\vec{B}_1)_y = (\vec{B}_3)_y \qquad (\mathbb{C}_1)_{xx} = -(\mathbb{D}_3)_{xx}$$
$$(\mathbb{C}_1)_{yy} = -(\mathbb{D}_3)_{yy} \qquad (\mathbb{C}_1)_{xy} = (\mathbb{D}_3)_{xy}$$
$$(\mathbb{C}_1)_{yx} = (\mathbb{D}_3)_{yx} \qquad (\mathbb{C}_3)_{xx} = -(\mathbb{D}_1)_{xx}$$

etc. Additional requirements (not necessarily linearly independent of the above) come from requiring the quasi-Hamiltonian to be invariant under simultaneous translation of all the nuclei and requiring that it reduce to (3.101) in the sense described previously. Thus:

$$\sum(A_j + B_j + C_j) = 0$$
$$\sum(\mathbb{C}_j + \mathbb{D}_j^T) = 0$$
$$A_{1x} + B_{1x} + C_{1x} = - A_{3x} - B_{3x} - C_{3x} = \frac{\sqrt{3}}{2} A$$
$$A_{2x} + B_{2x} + C_{2x} = 0$$
$$A_{1y} + B_{1y} + C_{1y} = A_{3y} + B_{3y} + C_{3y} = - A/2$$
$$A_{2y} + B_{2y} + C_{2y} = A$$
$$(\mathbb{C}_j)_{xx} + (\mathbb{D}_j)_{xx} + (\mathbb{C}_j)_{yy} + (\mathbb{D}_j)_{yy} + (\mathbb{C}_j)_{zz} + (\mathbb{D}_j)_{zz} = 0$$

where A is the constant appearing in (3.101). Explicit values for $\sin\frac{2\pi\Lambda}{g}(2j-1)$ and $\cos\frac{2\pi\Lambda}{g}(2j-1)$ have been substituted into the above expression when applying (5.5).

It is now a simple matter to write down the most general solution of these equations in terms of a minimal number of undetermined constants. When this is done one finds the number of such constants to be of the order of ten. In principle, the next step, in order to further reduce the number of constants, would be that of requiring the quasi-Hamiltonian to be invariant under the operation $T\Theta$. However, when we attempt to impose this requirement we run into difficulties. We find that it is impossible to take Θ in the form (6.33) and yet have

$$\Theta U_3(\sigma_{n-1}, \sigma_n, \sigma_{n+1}) \Theta^{-1}$$

depend only on the spins of the n-th, (n-1)-th, and (n+1)-th cells. This difficulty could conceivably be due to the fact that we have made a poor choice in the form of Θ. To find a general 2^N-by-2^N matrix for arbitrary N which satisfies all the properties desired of Θ and which allows us to formulate a quasi-Hamiltonian with the property of nearest neighbor interactions only, however, appears to be a hopeless task. Indeed, it could very well be that such a matrix does not even exist. Thus if we are to proceed further we must abandon one of the two assumptions - that $T\Theta$ commute with the Hamiltonian and that the Hamiltonian be limited to interactions between neighboring cells only.

Since we must abandon one of these two assumptions we choose to abandon the requirement that $T\Theta$ be a symmetry operation. Doing this will enable us to find a quasi-Hamiltonian of sufficient simplicity that it may be studied and may show some promise as a useful model of electron-lattice interaction in solids. While it is true that this model may not have alltthe features which we have desired it is to be hoped that this defect will not seriously impair its usefullness.

In the absence of any further requirements, however, the number of undetermined constants in the expansion (6.40) is still too large. If we wish to have as simple a model as possible it is clear that we must reduce this number of constants. Just what assumptions should be made to accomplish this appears somewhat arbitrary but we choose the following as they have the greatest appeal to our intuition:

1) The only atoms which may interact with the vibronic spin of a cell are the atoms in the cell and the two atoms adjacent to the boundaries of the cell.

2) There are no terms in the quasi-Hamiltonian which correspond to a product of spins of different cells with a nuclear displacement.

The first assumption implies that $\vec{A}_1$, $\vec{C}_3$, $\vec{A}_2$, and $\vec{C}_2$ are zero while the second assumption implies that $\mathcal{C}_1$, $\mathcal{D}_1$, $\mathcal{C}_2$, $\mathcal{D}_2$, $\mathcal{C}_3$, and $\mathcal{D}_3$ are zero. The first assumption is equivalent to assuming the interaction of the vibronic spin with external atoms to be of very short range and seems to be a logical refinement of the requirement of nearest neighbor interactions only. The second assumes that the spin-spin-lattice terms are not as important as the spin-lattice terms and enter in a higher order.

With these two assumptions and the relations we have previously obtained among the coefficients in the expansion (6.40) we may now write down the following expressions for the matrices $U_j^{(1)}$:

$$U_1^{(1)} = s(\sigma_{n-1,x} - \sigma_{n,x}) + \frac{\sqrt{3}}{2} A\sigma_{nx} + t(\sigma_{n-1,y} - \sigma_{n,y}) - \frac{A}{2}\sigma_{ny} \tag{6.41a}$$

$$U_2^{(1)} = A\sigma_{ny} \tag{6.41b}$$

$$U_3^{(1)} = - s(\sigma_{n+1,x} - \sigma_{n,x}) - \frac{\sqrt{3}}{2} A\sigma_{n,x} + t(\sigma_{n+1,y} - \sigma_{n,y}) - \frac{A}{2}\sigma_{n,y} \tag{6.41c}$$

where s and t are two additional unspecified constants.

Our remaining task before writing down the quasi-Hamiltonian is to specify the form of the electron-electron interaction term $U^{(0)}$. Requiring this to exhibit nearest neighbor interactions and to commute with the operators $C\Gamma$, $\Gamma(\kappa)$, and $T^{g}\mathcal{B}$ gives

$$U^{(0)} = \sum_{n}\Big(u\sigma_{n}\sigma_{n+1} + v\sigma_{nz}\sigma_{n+1,z} + w\sigma_{n,y}\sigma_{n+1,y} + r[\sigma_{nx}\sigma_{n+1,y} - \sigma_{ny}\sigma_{n+1,y}]\Big) + q \tag{6.42}$$

where q, u, v, w, and r are real constants. The value of the constant term q -just as was the case for N=2 - is irrelevant and may be set to zero.

If we were to take $T\Theta$ as being an invariant with Θ in the form (6.33) the only effect it would have on the form above would be that of imposing the requirement $w = 0$. As it is desirable to reduce the number of constants in the Hamiltonian as much as possible we shall impose this requirement with no further justification. It must be borne in mind, however, that this is a weak assumption and that the neglected term may always be added to the quasi-Hamiltonian.

With our choices for the matrices $U^{(0)}$ and $U_j^{(1)}$ we may now write down our model Hamiltonian

$$\begin{aligned} H = K &+ \sum_{n=1}^{N}\sum_{j=2}^{j=3}\omega_o^2(X_{nj} - X_{n,j-1})^2 + \sum_{n=1}^{N-1}\omega_o^2(X_{n+1,1} - X_{n3})^2 \\ &+ \omega_o^2(X_{1,1} - X_{N,3})^2 + \sum_n\left(u\sigma_n\sigma_{n+1} + v\sigma_{nz}\sigma_{n+1,z}\right) \\ &+ r\sum_n(\sigma_{nx}\sigma_{n+1,y} - \sigma_{ny}\sigma_{n+1,x}) + s\sum_n\{(\sigma_{n-1,x} - \sigma_{nx})X_{n1} - (\sigma_{n+1x} - \sigma_{nx})X \\ &+ t\sum_n\{(\sigma_{n-1y} - \sigma_{ny})X_{n1} + (\sigma_{n+1,y} - \sigma_{n,y})X_{n3}\} + \frac{\sqrt{3}}{2}A\sum_n\sigma_{nx}(X_{n1} - X_{n3}) \\ &- \frac{A}{2}\sum_n\sigma_{ny}(X_{n1} + X_{n3}) + A\sum_n\sigma_{ny}X_{n,2} \qquad (6.43) \end{aligned}$$

Although this has been derived for a particular choice of g it is clear that the generalization to any value of g is straightforward and simple if one makes use of all the assumptions used in deriving (6.43).

If our model were to be considered as a model of any one type of metal it is clear that there should be an optimum choice of g which will yield the most useful model. The view has been put forward by Tisza[1] [Ref. 45] and by Slater [Ref. 46] that there may very well exist a superstructure in some metals in which the basic cell contains more than one unit cell of the crystal. Slater has also quoted some experimental evidence which seems to support this view. If such a viewpoint is correct a model of the type proposed here would take advantage of this superstructure with the proper choice of g. Such a circumstance could also validate our relaxing of the requirement that $T\Theta$ be an invariant of the quasi-Hamiltonian..

G. Discussion.

We have thus exhibited a possible new model for electron-lattice interactions in solids which is a generalization of the molecular models based on the Born-Renner approximation studied in the previous chapters of this thesis. This model certainly over-comes the second difficulty listed at the beginning of this chapter and it is sufficiently simple to show promise in overcoming the first difficulty.

The one basic assumption upon which this model quasi-Hamiltonian rests is that given by (6.4), where we have taken the point of view that a solid may be considered as an aggregate of molecules - each molecule having a non-classical degree of freedom which we have called vibronic spin. The remainder of the assumptions have either appeared as a consequence of this viewpoint or else have been based on other intuitive notions with the primary desire of making our quasi-Hamiltonian as simple as possible and yet still have a non-trivial model electron lattice interactions in solids. While we have tried to make the latter set of assumptions be the most appealing choice by using our intuition based on a knowledge of other models of a solid and the experience gained by a study of the Moffitt-Liehr quasi-Hamiltonians in the molecular case, it is clear that alternate choices could be made and that further study is necessary to ascertain the optimum choice of assumptions.

In constructing this model, we chose a phenomenological approach and asked ourselves just what properties should such a quasi-Hamiltonian have, and proceeded, as best we could, to find such a quasi-Hamiltonian. The alternative method would have been to derive it starting from the exact particle-Hamiltonian. We chose the former because it was the simpler - the method of accomplishing the latter not being readily apparent. However, if our principle tenet, (6.4), is to be correct, there should exist a series of mathematical approximations which, in principal, will allow us to accomplish this. A determination of the form of the quasi-Hamiltonian starting from the particle-Hamiltonian will be desired if we wish to examine the correctness of our assumptions and to find whether an alternate model is needed.

We shall not seek to determine the eigenfunctions and eigenvalues of the model Hamiltonian obtained in this chapter - this remains for later investigations. However, it should be pointed out that we already know a considerable amount concerning the properties of this model from our considerations of the molecular case. If this model were to be solved subject to periodic boundary conditions where the "periodicity volume" were taken to be that of a single super cell, we would be solving a molecular vibronic equation similar to those studied in the previous chapters. While the validity of the method of periodic boundary conditions can be expected to "diminish" with the size of the "periodicity volume" we can, at least, expect the solutions in this limit to give us a "hint" about the nature of the solutions in more general cases. In particular, since one would expect the energy level density, roughly, to vary exponentially with the size of the crystal, the energy level density for the single cell case should give us an indication of the energy level density for a periodicity volume of arbitrary length.

Chapter 7

SOME CONCLUDING REMARKS

A. An Appraisal of the Theory.

The bulk of this thesis has been concerned with the generalization of the Born-Oppenheimer approximation to deal with problems in molecular physics. However, our prime motivation in writing the thesis was the fact that these ideas and methods may have some application in the study of solid-state problems. Thus our primary emphasis was concerned with the formalism of the theory and of the nature of the resulting quasi-Hamiltonians rather than in the detailed interpretation of anomalies in molecular spectra on the basis of the theory. When we proceeded to consider the nature of the eigenfunctions and eigenvalues of these quasi-Hamiltonians we emphasized the general features which may be expected for this general class of quasi-Hamiltonians rather than numerically tabulating the energy spectrum of any one given model. We were primarily interested in the properties of the ground state - whether or not it was degenerate and, if degenerate, whether or not the two degenerate states could be chosen such that they carried currents in opposite directions. While it is clear that a single molecule would not serve as an accurate model of a superconductor, it was nevertheless interesting to find that these simple molecular models of electron-vibration interaction exhibited some features which one would expect in a model of a superconductor.

Chapter 6 stands apart from the rest of the thesis for it is not part of our exposition on the formalism of the Born-Renner approximation. Rather, it represents a first attempt at accomplishing that objective of applying the concepts based on a generalization of the Born-Oppenheimer approximation to solids. The theory we present there is basically guesswork and phenomenology as compared to the comparatively rigorous treatment given in the remainder of the thesis. Naturally a theory obtained directly from the exact particle Hamiltonian is to be preferred, but obtaining such a theory appears to be a much more difficult problem and we have excluded its consideration from this thesis. Nevertheless, the basic tenet of the phenomenological model - that the crystal may be considered as an aggregate of coupled molecules in degenerate states appears attractive and it was interesting to explore the consequences of such a viewpoint.

We now give a brief summary of the dissertation.

First,we have reviewed the Born-Oppenheimer approximation - its justifications and limitations - and have defined a method of generalizing this approximation. We have considered the application of the approximation to a general model of the interaction of a two-fold degenerate level with a two-fold degenerate vibration and have shown how symmetry considerations may be utilized in obtaining the quasi-Hamiltonian for this model and have shown that earlier models (Renner and dynamic Jahn-Teller) may be considered as special cases of our general model and, in so doing, have clarified the relation between these two general classes of models. In addition, we have extended Born and Oppenheimer's original justification to include the degenerate case. We have studied the solutions of the vibronic equations resulting from this approximation and have made use of symmetry considerations to deduce certain general properties of these solutions. We have also examined the nature of the solutions in two limiting cases and have deduced the circumstances in which one may expect the ground state to be degenerate. We have discussed the current carrying properties of the eigenfunctions and have shown that a linear combination of degenerate eigenfunctions may in general always be taken such that the expectation value of the current is non-zero. Finally, we have made an attempt to generalize these considerations to solids and have devised a phenomenological model of a solid based on the conceptual view of a crystal as an aggregate of coupled molecules in degenerate electronic states.

B. Suggestions for Further Study.

As evidenced by the number of recent papers on the subject, the study of electron-vibration interactions by use of the generalized Born-Oppenheimer approximation is presently an active research topic in quantum chemistry. However, we shall confine our remarks solely to the problems raised by the application of the method to solids.

The possibility of the formulation of new models of a solid based on the methods and the concepts developed in this thesis merits further study. The problem of establishing the relation of these models to the exact particle Hamiltonian is suggested and, indeed, these two problems should be considered simultaneously. The establishment of such a model will naturally lay open a wealth of problems - one principal objective being a theory of superconductivity complementary to the BCS theory which gives a more detailed insight into the nature (wavefunction, structure, and symmetry) of the ground state of a superconductor. In the course of the study of such a model one would hopefully determine the nature of the non-classical order-disorder parameters whose existence was inferred in the macroscopic thermodynamics of equilibrium.

In the absence of a model whose relationship to the exact particle Hamiltonian is firmly established we suggest the study of the phenomenological model which is presented in this thesis. Some further thought, however, might be given to the assumptions made in formulating the model. One possible realm of application of such a model, only briefly mentioned in this thesis, is that of molecular crystals in which the molecules comprising the crystal are not strongly coupled and may be considered in first approximation as existing in degenerate electronic states. We have not investigated the question of whether there are any molecular crystals for which these requirements are satisfied, but it does appear that considerations such as given in Chapter 6 of this thesis could lead to a succesful explanation of electron-vibration interaction in such crystals.

Appendix A

PROOF OF THE MOFFITT-LIEHR THEOREM

Our manner of proof[1] will consist in explicitly exhibiting the functions $\overline{\Phi}_{\Lambda}$, $U_{\Lambda\Lambda}$, etc. as determined by a perturbation-like calculation. Let us assume that

$$\overline{\Phi}_{\Lambda} = \overline{\Phi}_{\Lambda}^{(o)} + \sum_{n} C_{n\Lambda}\overline{\Phi}_{n}^{(o)} + \sum_{\Lambda'} C_{\Lambda\Lambda'}\overline{\Phi}_{\Lambda'}^{(o)} \qquad (A.1)$$

where the $C_{n\Lambda}$ and $C_{\Lambda\Lambda'}$ are functions of the vibrational coordinates only. The letters Λ, Λ' will be used to designate the two eigenfunctions (which we defined as $\overline{\Phi}_{\Lambda}^{(o)}$ and $\overline{\Phi}_{-\Lambda}^{(o)}$ in Chapter 3) of H_{X_o} associated with the eigenvalue $\mathcal{E}(X_o)$. The letter n will be used to label eigenfunctions of H_{X_o} which are not degenerate with these two eigenfunctions, i.e.,

$$H_{X_o}\,\overline{\Phi}_{n}^{(o)} = E_{n}^{(o)}\,\overline{\Phi}_{n}^{(o)} \qquad (A.2)$$

where

$$E_{n}^{(o)} \neq \mathcal{E}(X_o)\,. \qquad (A.3)$$

The functions $\overline{\Phi}_{n}^{(o)}$, $\overline{\Phi}_{n'}^{(o)}$, $\overline{\Phi}_{\Lambda}^{(o)}$, etc. are taken to be orthonormal.

---

[1] We refer to Chapter 3 for a statement of the theorem.

If we now substitute the expression (A.1) into the equation

$$H_X \, \bar{\Phi}_{\Lambda} \quad = \quad \sum_{\Lambda''} U_{\Lambda''\Lambda} \, \bar{\Phi}_{\Lambda''} \tag{A.4}$$

and multiply by $\bar{\Phi}^{(0)*}_{\Lambda'}$ and integrate over the electron coordinates we obtain the equation

$$(H_X)_{\Lambda'\Lambda} \quad + \quad \sum_{n} (H_X)_{\Lambda' n} C_{n\Lambda} \quad + \quad \sum_{\Lambda''} (H_X)_{\Lambda'\Lambda''} C_{\Lambda''\Lambda}$$

$$= \quad U_{\Lambda'\Lambda} \quad + \quad \sum_{\Lambda''} C_{\Lambda'\Lambda''} U_{\Lambda''\Lambda} \quad . \tag{A.5}$$

Similarly, multiplying by $\bar{\Phi}^{(0)*}_{n}$ and integrating yields

$$(H_X)_{n\Lambda} \quad + \quad \sum_{n'} (H_X)_{nn'} C_{n'\Lambda} \quad + \quad \sum_{\Lambda'} (H_X)_{n\Lambda'} C_{\Lambda'\Lambda}$$

$$= \quad \sum_{\Lambda'} C_{n\Lambda'} \, U_{\Lambda'\Lambda} \quad . \tag{A.6}$$

Requiring the eigenfunctions of H_X to be orthonormal gives in addition the equation

$$\sum_{n} C^{*}_{n\Lambda} \; C_{n\Lambda'} \quad + \quad \sum_{\Lambda''} C_{\Lambda''\Lambda} \quad C_{\Lambda''\Lambda'}$$

$$+ \quad C_{\Lambda\Lambda'} \quad + \quad C_{\Lambda'\Lambda} \quad = 0. \tag{A.7}$$

If we can now find solutions of (A.5), (A.6), and (A.7) satisfying the assertions stated in Chapter 3 the theorem will be proven. Let us expand H_X in a Taylor series in Q_1 and Q_2

$$H_X = H_{X_0} + H^{(1)}_{XX} + H^{(2)}_X + \cdots \qquad (A.8)$$

where $H^{(1)}_X$ is of first order, etc. (It is assumed that this expansion is possible. For real molecules this can always in principle be done.) The $C_{n\Lambda'}$, $C_{\Lambda\Lambda'}$, $U_{\Lambda\Lambda'}$ are initially also assumed to be expandable in a power series:

$$C_{n\Lambda} = C^{(0)}_{n\Lambda} + C^{(1)}_{n\Lambda} + \cdots \qquad (A.9)$$

$$U_{\Lambda\Lambda'} = U^{(0)}_{\Lambda\Lambda'} + U^{(1)}_{\Lambda\Lambda'} + \cdots \qquad (A.10)$$

etc. (The theorem will be proven if we can explicitly find the terms in these expansions.) The expansions (A.8), (A.9), (A.10) are now substituted into equations (A.5), (A.6), and (A.7) and the various powers of Q_1 and Q_2 are equated. The zeroth order equations are

$$\mathcal{E}(x_0)\,\delta_{\Lambda'\Lambda} + \mathcal{E}(x_0)\,C^{(0)}_{\Lambda'\Lambda} = U^{(0)}_{\Lambda'\Lambda} + \sum_{\Lambda''} C^{(0)}_{\Lambda'\Lambda''}\,U^{(0)}_{\Lambda''\Lambda} \qquad (A.11)$$

$$E^{(0)}_n C^{(0)}_{n\Lambda} = \sum_{\Lambda'} C^{(0)}_{n\Lambda'}\,U^{(0)}_{\Lambda'\Lambda} \qquad (A.12)$$

$$\sum_n C^{(0)*}_{n\Lambda} C^{(0)}_{n\Lambda'} + \sum_{\Lambda''} C^{(0)*}_{\Lambda''\Lambda}\,C^{(0)}_{\Lambda''\Lambda'} + C^{(0)}_{\Lambda\Lambda'} + C^{(0)*}_{\Lambda'\Lambda} = 0 \qquad (A.13)$$

These equations are satisfied by taking

$$U^{(0)}_{\Lambda\Lambda} = \mathcal{E}(X_0) \qquad (A.14a)$$

$$U^{(0)}_{\Lambda\Lambda'} = 0 \qquad (A.14b)$$

$$C^{(0)}_{n\Lambda} = C^{(0)}_{\Lambda\Lambda'} = 0 \qquad (A.14c)$$

Thus we have satisfied equations (3.14) and (3.15).

Similarly we can write down the first, second, third, etc., order equations. If this is done we find the resulting equations are satisfied by taking

$$U^{(1)}_{\Lambda\Lambda'} = (H^{(1)}_X)_{\Lambda\Lambda'} \qquad (A.15a)$$

$$C^{(1)}_{n\Lambda} = - \frac{(H^{(1)}_X)_{n\Lambda}}{E^{(0)}_n - \mathcal{E}(X_0)} \qquad (A.15b)$$

$$C^{(1)}_{\Lambda\Lambda'} = 0 \qquad (A.15c)$$

$$U^{(2)}_{\Lambda\Lambda'} = (H^{(2)}_X)_{\Lambda\Lambda'} - \sum_n \frac{(H^{(1)}_X)_{\Lambda n}(H^{(1)}_X)_{n\Lambda'}}{E^{(0)}_n - \mathcal{E}(X_0)} \qquad (A.16a)$$

$$C^{(2)}_{n\Lambda} = \sum_{n'} \frac{(H^{(1)}_X)_{nn'}(H^{(1)}_X)_{n'\Lambda}}{(E^{(0)}_n - \mathcal{E}(X_0))(E^{(0)}_{n'} - \mathcal{E}(X_0))} - \frac{(H^{(2)}_X)_{n\Lambda}}{E^{(0)}_n - \mathcal{E}(X_0)} - \sum_{\Lambda'} \frac{(H^{(1)}_X)_{n\Lambda'}(H^{(1)}_X)_{\Lambda'\Lambda}}{(E^{(0)}_n - \mathcal{E}(X_0))^2} \qquad (A.16b)$$

$$C^{(2)}_{\lambda\lambda'} = - \sum_{n} \frac{(H_X^{(1)})_{\lambda n}(H_X^{(1)})_{n\lambda'}}{2(E_n^{(0)} - \mathcal{E}(X_0))^2} \qquad \text{(A.16c)}$$

$$\begin{aligned} U^{(3)} = (H_X^{(3)})_{\lambda\lambda'} &- \sum_{n} \frac{(H_X^{(1)})_{\lambda n}(H_X^{(2)})_{n\lambda'}}{E_n^{(0)} - \mathcal{E}(X_0)} \\ &- \sum_{n} \frac{(H_X^{(2)})_{\lambda n}(H_X^{(1)})_{n\lambda'}}{E_n^{(0)} - \mathcal{E}(X_0)} \\ &+ \sum_{n,n'} \frac{(H_X^{(1)})_{\lambda n}(H_X^{(1)})_{n'n}(H_X^{(1)})_{n\lambda'}}{(E_n^{(0)} - \mathcal{E}(X_0))(E_{n'}^{(0)} - \mathcal{E}(X_0))} \\ &- \sum_{n,\lambda''} \frac{(H_X^{(1)})_{\lambda n}(H_X^{(1)})_{n\lambda''}(H_X^{(1)})_{\lambda''\lambda'}}{2(E_n^{(0)} - \mathcal{E}(X_0))^2} \\ &- \sum_{\lambda'',n} \frac{(H_X^{(1)})_{\lambda\lambda''}(H_X^{(1)})_{\lambda'' n}(H_X^{(1)})_{n\lambda'}}{2(E_n^{(0)} - \mathcal{E}(X_0))^2} \qquad \text{(A.17)} \end{aligned}$$

The process can be continued indefinitely. The only possible flaw in this proof is that some of the sums over n might not converge. We expect no trouble however as long as the $E_n^{(0)}$ are sufficiently far removed from the energy level $\mathcal{E}(X_0)$.

Let us now tackle Theorem II. We first note that expressions of the form

$$G(x', x) = \sum_{n} \frac{\overline{\Phi}_{n}^{(0)*}(x')\ \overline{\Phi}_{n}^{(0)}(x)}{E_n^{(0)} - \mathcal{E}(X_0)} \tag{A.18}$$

are invariant under all the symmetry operations of H_X, i. e.,

$$G(Rx', Rx) = G(x', x) \tag{A.19}$$

$$G^{*}(x', x) = G(x', x) \tag{A.20}$$

These two relations follow as any set of ortho-normal wavefunctions associated with a single level must be a basis for a unitary representation of the symmetry operations.

From the above it follows that expressions such as

$$\sum_{n} c_{n\Lambda}^{(1)}(r, \varphi)\ \overline{\Phi}^{(0)}_{n}(x)$$

behave just as $\overline{\Phi}_{\Lambda}^{(0)}$ under symmetry operations. Consider for example the operation R_α where

$$\overline{\Phi}_{\Lambda}^{(0)}(R_\alpha x) = e^{i\Lambda\alpha}\ \overline{\Phi}_{\Lambda}^{(0)}(x)$$

We note that

$$\begin{aligned} \sum_{n} c_{n}^{(1)}(r,\varphi)\ \overline{\Phi}_{n}^{(0)}(x) &= \int G(x', x) H_X^{(1)}(r, \varphi, x')\overline{\Phi}_{\Lambda}^{(0)}(x')\,d\tau_e' \\ &= \int G(R_\alpha x', x)\ H_X^{(1)}(r, \varphi, R_\alpha x')\ \overline{\Phi}_{\Lambda}^{(0)}(R_\alpha x')\,d\tau_e' \end{aligned} \tag{A.21}$$

since replacing x' by $R_\alpha x'$ does not affect the value of the integral as we are integrating over all "space". Therefore

$$\begin{aligned}
&\sum_n c_n^{(1)}(r,\ R_\alpha\varphi)\ \bar{\Phi}_n^{(0)}(R_\alpha x) \\
&= \int G(R_\alpha x',\ R_\alpha x)\ H_X^{(1)}(r,\ R_\alpha\ ,\ R_\alpha x')\ \bar{\Phi}_\Lambda^{(0)}(R_\alpha x')\,d\tau_e' \\
&= e^{i\Lambda\alpha}\int G(x',\ x)\ H_X^{(1)}(r,\varphi,\ x')\ \bar{\Phi}_\Lambda^{(0)}(x')\ d\tau_e' \\
&= e^{i\Lambda\alpha}\ \sum_n c_n^{(1)}(r,\varphi)\ \bar{\Phi}_n^{(0)}(x) \qquad (A.22)
\end{aligned}$$

as

$$H_X^{(1)}(r,\ R_\alpha\varphi,\ R_\alpha x') \quad = \quad H_X^{(1)}(r,\varphi,\ x') \qquad (A.23)$$

Similarly

$$\sum_n c_{n\Lambda}^{(1)*}\ \bar{\Phi}_n^{(0)*} \quad = \quad \sum_n c_{n,-\Lambda}^{(1)}\,\bar{\Phi}_n^{(0)} \qquad (A.24)$$

and

$$\sum_n c_{n\Lambda}^{(1)}(r,\ C\varphi)\ \bar{\Phi}_n^{(0)}(Cx) = \sum_n c_{n,-\Lambda}^{(1)}(r,\varphi)\ \bar{\Phi}_n^{(0)}(x) \qquad (A.25)$$

if the operation C exists. Such relations also hold for

$$\sum_n c_{n\Lambda}^{(2)}(r,\varphi)\ \bar{\Phi}_n^{(0)}(x)$$

and so on to all orders. Thus

$$\bar{\Phi}_{\Lambda}(r, \varphi, x)^{*} = \bar{\Phi}_{-\Lambda}(r, \varphi, x) \qquad (A.26a)$$

$$\bar{\Phi}_{\Lambda}(r, R_{\alpha}\varphi, R_{\alpha}x) = e^{i\Lambda\alpha}\, \bar{\Phi}_{\Lambda}(r, \varphi, x) \qquad (A.26b)$$

$$\bar{\Phi}_{\Lambda}(r, C\varphi, C x) = \bar{\Phi}_{-\Lambda}(r, \varphi, x) \qquad (A.26c)$$

as stated in the theorem.

The properties of the $U_{\Lambda\Lambda'}$ follow from equations (A.26) and the fact that

$$U_{\Lambda\Lambda'} = \int \bar{\Phi}_{\Lambda}\, H_X\, \bar{\Phi}_{\Lambda'}\, d\tau_e \qquad (A.27)$$

Thus

$$\begin{aligned} U^{*}_{\Lambda\Lambda'} &= \int \bar{\Phi}_{\Lambda} H_X\, \bar{\Phi}^{*}_{\Lambda'}\, d\tau_e \\ &= \int \bar{\Phi}^{*}_{-\Lambda}\, H_X\, \bar{\Phi}_{-\Lambda'}\, d\tau_e \\ &= U_{-\Lambda,-\Lambda'} \qquad (A.28) \end{aligned}$$

and

$$\begin{aligned} U_{\Lambda\Lambda'}(r, R_{\alpha}\varphi) &= \\ &= \int \bar{\Phi}^{*}_{\Lambda}(R_{\alpha}x, r, R_{\alpha}\varphi)\, H_X(R_{\alpha}x, r, R_{\alpha}\varphi)\, \bar{\Phi}_{\Lambda'}(R_{\alpha}x, r, R_{\alpha}\varphi)\, d\tau_e \\ &= e^{i\alpha(\Lambda'-\Lambda)} \int \bar{\Phi}^{*}_{\Lambda}(x, r, \varphi)\, H_X(x, r, \varphi)\, \bar{\Phi}_{\Lambda'}(x, r, \varphi)\, d\tau_e \\ &= e^{i\alpha(\Lambda'-\Lambda)}\, U_{\Lambda\Lambda'}(r, \varphi). \qquad (A.29) \end{aligned}$$

In a similar fashion

$$U_{\lambda\lambda'}(r,\ C\varphi) = U_{\lambda\lambda'}(r,\ -\varphi) = U_{-\lambda,-\lambda'}(r,\varphi) \qquad (A.30)$$

Equations (3.19), (3.21), and (3.23) then follow in a trivial manner.

Appendix B

TWO-DIMENSIONAL REPRESENTATIONS OF MOLECULAR SYMMETRY GROUPS

The only type of representations that we need consider here are those which are equivalent to real orthogonal representations. This follows since the molecular vibrational coordinates may always be chosen to be linear combinations of the nuclear coordinates with real coefficients. Also any two-fold degenerate electronic level whose degeneracy is not the Kramer's degeneracy due to time reversal symmetry may always be associated with a representation of this type.

Any two-by-two real orthogonal matrix may be written in one of the two following forms

$$\Gamma'_{I,\alpha} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \qquad (B.1)$$

or

$$\Gamma'_{II,\alpha} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} \qquad (B.2)$$

depending on whether the determinant of the matrix is 1 or -1. Here α is a real angle and characteristic of the particular matrix. Under a similarity transformation of the form

$$\Gamma = \mathcal{A}\,\Gamma'\mathcal{A}^{\dagger} \tag{B.3}$$

where

$$\mathcal{A} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \tag{B.4}$$

we find

$$\Gamma_{I,\alpha} = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix} \tag{B.5}$$

$$\Gamma_{II,\alpha} = \begin{pmatrix} 0 & e^{-i\alpha} \\ e^{i\alpha} & 0 \end{pmatrix} \tag{B.6}$$

Thus our representation may always be chosen such that any matrix in the representation is either of the form (B.5) or (B.6) for some α.

We distinguish two possibilities: either every matrix in the representation may be taken in the form (B.5) or else there are some matrices of the form (B.6). (The latter possibility may be correlated with the existence of a reflection plane containing the symmetry axis or a two-fold rotation axis perpendicular to the symmetry axis.) These are mutually exclusive possibilities as the determinant of a matrix is invariant under similarity transformations.

If the second possibility is true then it is always possible to take the matrix

$$\Gamma_{II,0} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

as being in the representation. This is easily seen from the following considerations. Suppose the matrix

$$\Gamma_{II,\beta} = \begin{pmatrix} 0 & e^{-i\beta} \\ e^{i\beta} & 0 \end{pmatrix}$$

appears in the representation. Then if we subject all of the matrices in the representation to the similarity transformation

$$\Gamma' = \Gamma_{I,\beta/2}\,\Gamma\,\Gamma_{I,\beta/2}^{\dagger} \tag{B.7}$$

we find that $\Gamma_{II,\beta}$ is transformed into $\Gamma_{II,0}$ and that all the $\Gamma_{I,\alpha}$ remain invariant.

We may therefore speak of a standard form for a two-dimensional representation in which all of the matrices are either in the form (B.5) or (B.6) and in which the matrix $\Gamma_{II,0}$ is included if there are any matrices of the form (B.6) in the representation.

Since the matrices in a representation form a group we can make the following conclusions concerning just what matrices are included in a representation:

1) If $\Gamma_{I,\alpha}$ and $\Gamma_{I,\beta}$ appear in the representation then $\Gamma_{I,\alpha+\beta}$ appears in the representation. In particular this implies that $\Gamma_{I,n\alpha}$ (for any integer n) appears in the representation.

2) If $\Gamma_{I,\alpha}$ appears in the representation then $\Gamma_{I,-\alpha}$ also appears in the representation.

3) If $\Gamma_{I,\alpha}$ and $\Gamma_{II,\beta}$ appear in the representation then $\Gamma_{II,\alpha+\beta}$ appears in the representation.

These follow from considering that the product of any two matrices appearing in the representation as well as the inverse of any matrix appearing in the representation must also appear in the representation.

It also follows that if $\Gamma_{II,\beta}$ is in the representation, then $\Gamma_{I,\beta}$ is in the representation, provided we have taken our representation in the standard form which we previously defined, for since $\Gamma_{II,\beta}$ and $\Gamma_{II,\alpha}$ are in the representation

$$\Gamma_{II,0}\Gamma_{II,\beta} = \Gamma_{I,\beta}$$

is also in the representation. Conversely, it follows that if $\Gamma_{I,\alpha}$ and $\Gamma_{II,0}$ are both in the representation then $\Gamma_{II,\alpha}$ is in the representation.

If our representation is, moreover, a finite representation, i. e., there exists a finite number of distinct elements, we can show that there exists an $\bar{\alpha}$ such that any matrix in the representation is of the form $\Gamma_{I,n\bar{\alpha}}$ or $\Gamma_{II,n\bar{\alpha}}$ where n is an integer.

In order that the degeneracy (vibrational or electronic) associated with a two-fold representation not be an <u>accidental</u> degeneracy the representation must not be reducible into two real one-dimensional representation. It is easily seen that a necessary and sufficient condition for this is that there be matrices $\Gamma_{I,\alpha}$ in the representation for which neither

$$\alpha = 0 \ (\text{Mod } 2\pi)$$

or

$$\alpha = \pi \ (\text{Mod } 2\pi)$$

This, in conjunction with our previous remarks, implies that any molecule having a symmetry group with two-dimensional representations must have at least a three-fold rotation axis or at least a four-fold rotation-reflection axis.

If there is only one such axis, say, a P-fold axis, and we let R_γ denote a rotation (or rotation-reflection) through an angle $2\pi/P$ about this axis the matrix $\Gamma(R_\gamma)$ representing R_γ in a two-dimensional representation (taken in our standard form) must be of the form $\Gamma_{I\alpha}$ (as contrasted to $\Gamma_{I\beta}$) in order that our representation be irreducible. Furthermore, the associated α must be of the form $\ell\gamma$ where ℓ is an integer characteristic of the representation[1]. This follows since

$$(\Gamma_{I\alpha})^P$$

must be the unit matrix which implies that

$$e^{iP\alpha} = 1$$

$$P\alpha = 2\pi\ell$$

$$\alpha = (2\pi/P)\ell = \gamma\ell$$

The integer ℓ is related to the $\bar{\alpha}$ defined previously in that we may take

$$\bar{\alpha} = (2\pi\ell/P)$$

For molecules having cylindrical symmetry the associated group is infinite and it is meaningless to speak of an $\bar{\alpha}$ for any two-dimensional representation. We may, however, define an ℓ for every representation

---

[1] The letter Λ would be used instead of ℓ if we were speaking of the representation associated with two degenerate *electronic* states.

as being the ratio of the associated α with the angle of rotation about the molecular axis.

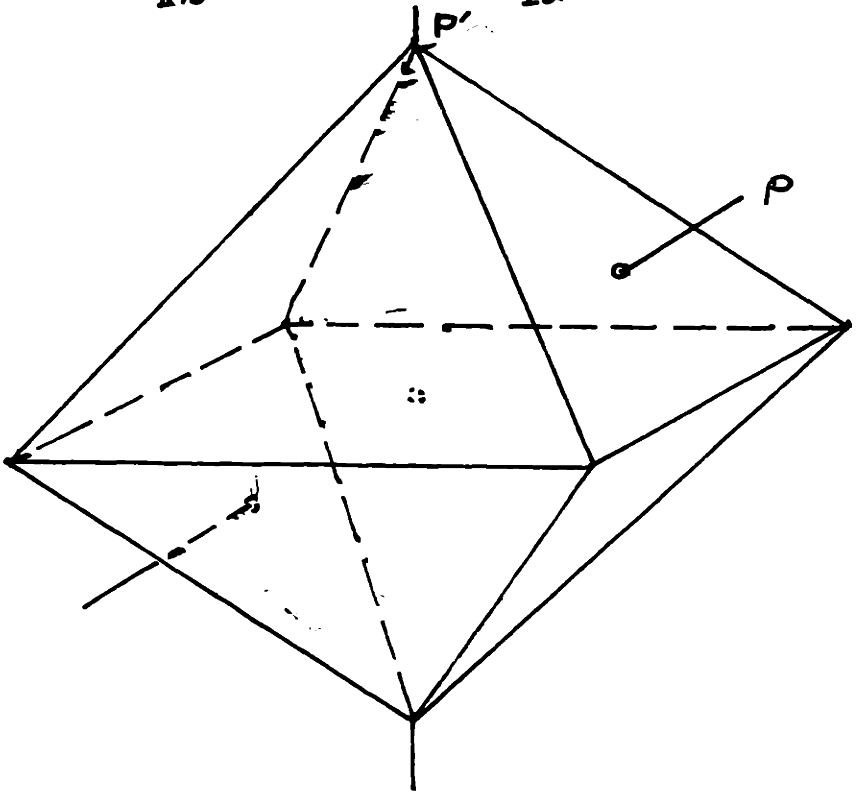
If we have more than one multiple axis, however, and merely pick out one axis at random we have no guarantee that the matrices representing a rotation (or rotation-reflection) about this axis are of the form $\Gamma_{I\alpha}$ rather than $\Gamma_{II\alpha}$. Nevertheless, an inspection of the character tables for all the molecular groups[1] reveals that we can always choose an axis such that any matrix representing a rotation (or rotation-reflection) about this axis in a two-dimensional irreducible representation is of the form $\Gamma_{I\alpha}$ regardless of the particular representation - as long as each representation is taken in our standard form.

Consider, for example, a molecule described by the group O_h. This molecule then has the same symmetry as a regular octahedron and the associated group d has two two-dimensional representations. The characters for a rotation through $\pi/2$ (C_4) about one of the four-

---

[1] See, for example, Herzberg, G., Infra-Red and Raman Spectra, First Ed. (1945) Princeton.

fold axes (such as P' in the figure) for both of these representations are zero while the characters for a rotation through π about the same axis is two. This indicates that the matrix representing the C_4 operationsmust be of the form $\Gamma_{II\beta}$ rather than $\Gamma_{I\alpha}$.



For a rotation-reflection (S_6) about one of the six-fold rotation-reflection axes (such as P in the figure) the characters are -1 and 1 respectively for the two representations - indicating that the matrices representing this operation should be of the form $\Gamma_{I\alpha}$ where α is neither o or π. Indeed, α should be 2π/3 for the first representation and π/3 for the second. Thus ℓ (or Λ) is 2 and 1, respectively, for these two representations.

If $\Phi(x)$ and $\Phi'(x)$ are any eigenfunctions transforming according to a two-dimensional representation we may now show that it is possible to take

$$\Phi(x) = \Phi'^{*}(x)$$

and still take our representation in the standard form[1]. Let R be an element of the molecular group represented by the matrix $\Gamma_{I\alpha}$ where α is neither π or 0. Then we may choose Φ and Φ' such that

$$\begin{pmatrix} \Phi(Rx) \\ \Phi'(Rx) \end{pmatrix} = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix} \begin{pmatrix} \Phi(x) \\ \Phi'(x) \end{pmatrix} \qquad (B.8)$$

Furthermore, Φ^{*} must be a linear combination of Φ and Φ' as they are both eigenfunctions of a real Hamiltonian and we are assuming that our level is only two-fold degenerate. Therefore

$$\Phi^{*}(Rx) = a\,\Phi(Rx) + b\,\Phi'(Rx)$$

$$e^{-i\alpha}\,\Phi^{*}(x) = a e^{i\alpha}\,\Phi(x) + b\,e^{-i\alpha}\,\Phi'(x)$$

$$\Phi^{*}(x) = a\,e^{2i\alpha}\,\Phi(x) + b\,\Phi'(x) \qquad (B.9)$$

---

[1] Strictly speaking, we should say that it is possible to take one function to be the result of the time reversal operator operating on the other. The consideration of time reversal (including spin) is an unnecessary complication which will not change our results.

Since $e^{2i\alpha}$ is not one we are forced to conclude that $a=0$ and

$$\varphi^{*}(x) = b\ \varphi'(x) \tag{B.10}$$

Since we are at liberty to choose $b = 1$ our assertion is proven.

It similarly follows that we may take our vibrational coordinates in the form $re^{i\varphi}$ and $re^{-i\varphi}$ such that they will transform under the operations of the molecular group according to a two-dimensional representation in the standard form. In particular, this implies that

$$R\,\varphi = \varphi + \alpha \tag{B.11}$$

if

$$\Gamma(R) = \Gamma_{I,\alpha}$$

or

$$R\,\varphi = -\varphi - \alpha \tag{B.12}$$

if

$$\Gamma(R) = \Gamma_{II,\alpha}$$

The radial coordinate r will remain invariant under all symmetry operations.

Suppose now that the molecule has either a reflection plane containing the symmetry axis or a two-fold (or even-fold) rotation axis perpendicular to the symmetry axis. Let R_γ denote a rotation (or rotation-reflection) of angle γ about the symmetry axis and C denote either a reflection through the plane or a rotation through π about the two-fold axis. Then it is readily verified that

$$R_{-\gamma}C = CR_\gamma \tag{B.13}$$

This must hold also for the matrices representing R_γ and C. If R_γ is represented by $\Gamma_{I,\alpha}$ then

$$\Gamma_{I,-\alpha}\Gamma(C) = \Gamma(C)\,\Gamma_{I,\alpha} \tag{B.14}$$

Taking α to be neither π or 0 and using the fact that $\Gamma(C)^2$ must be the identity matrix we find that $\Gamma(C)$ must be of the form

$$\begin{pmatrix} 0 & e^{-i\beta} \\ e^{i\beta} & 0 \end{pmatrix}$$

for some β. Since we are at liberty to apply a transformation of the form (B.7) to our representation we may always take $\beta = 0$ for any given C. Doing this, furthermore, does not invalidate the steps leading to (B.10). We therefore see that we can always choose our vibrational coordinates to be $re^{i\varphi}$ and $re^{-i\varphi}$ where

$$R_\gamma\varphi = \varphi + \ell\gamma \tag{B.15}$$

$$C\varphi = -\varphi \tag{B.16}$$

and our wavefunctions such that

$$\Phi'(x)^{*} = \Phi(x) \tag{B.17}$$

$$\Phi(R_{\gamma}x) = e^{i\Lambda\gamma}\,\Phi(x) \tag{B.18}$$

$$\Phi(Cx) = \Phi'(x) \tag{B.19}$$

for some integers ℓ and Λ thus proving the statements made in Chapter 3.

Appendix C

EXTENSION OF THE BORN-OPPENHEIMER METHOD TO THE DEGENERATE CASE

It is possible to extend Born and Oppenheimer's method to include the case of electronic degeneracies - thus giving in a certain sense a justification of the Born-Renner approximation. The development we give here will be very similar to that given by Born and Huang (Ref.) for the non-degenerate case.

We consider the two equations

$$H_X \overline{\Phi}_{\Lambda} = \sum_{\Lambda'} U_{\Lambda'\Lambda} \overline{\Phi}_{\Lambda'} \qquad (C.1)$$

and

$$(K + H_X) \Psi = E \Psi \qquad (C.2)$$

where the functions $\overline{\Phi}_{\Lambda}(x, X)$, $U_{\Lambda\Lambda'}(X)$, etc., are chosen to satisfy the conditions of the Moffitt-Liehr theorem. Both of these equations shall now be formally solved by a perturbation procedure, the solutions of (C.2) being expressed in terms of those of (C.1).

The usual manner of introducing such a perturbation scheme is to break up the respective Hamiltonians, H_X and $(K + H_X)$, for these two equations, defining zeroth, first, etc., order Hamiltonians. It is just such a perturbation scheme that shall be employed here.

It is desirable to choose the decomposition of our Hamiltonians in a manner as close as possible to that chosen by Born and Oppenheimer. In their scheme they formally expanded H_X in a Taylor's series in the vibrational coordinates

$$H_X = H_{X_0} + H_X^{(1)} + H_X^{(2)} + \cdots \qquad (C.3)$$

taking H_{X_0} as the zeroth order Hamiltonian, $H_X^{(1)}$ as a first order perturbation, and $H_X^{(k)}$ as a k-th order perturbation. In (c.2) the nuclear kinetic energy K would be taken as a second order "perturbation". If we were to carry out the solution of (C.2) following the usual perturbation methods we would find the zeroth order eigenvalue to be $\mathcal{E}(X_0)$. The first order equation would then demand that the eigenvalues of the matrix

$$\begin{pmatrix} (H_X^{(1)})_{\Lambda\Lambda} & (H_X^{(1)})_{\Lambda,-\Lambda} \\ (H_X^{(1)})_{-\Lambda,\Lambda} & (H_X^{(1)})_{-\Lambda,-\Lambda} \end{pmatrix}$$

be independent of the vibrational coordinates. In general this condition would only be satisfied if the elements of the matrix vanish identically. In some cases, i. e., linear molecules, this will be true if our equilibrium configuration is chosen properly (i. e., to be a linear configuration). In other cases, however, (namely those affected by the Jahn-Teller theorem) we cannot choose the configuration X_0 such that this will be true. As we wish to include such cases in our formalism it is clear that we must slightly modify Born and Oppenheimer's ordering scheme.

It is tacitly assumed that the displacements of the nuclei which alter the assumed equilibrium configuration are small. To maintain this assumption it is clear that we must assume the

$$(H_X^{(1)})_{\lambda\lambda'}$$

to be small, if not zero. This fact will serve as a guide in our decomposition of H_X.

Let P be a projection operator which projects out of any wave function the contribution from the eigenfunctions of H_{X_0} associated with the eigenvalue $\mathcal{E}(X_0)$. Thus

$$P\,\bar{\Phi}_{\lambda}^{(0)}(x) = \bar{\Phi}_{\lambda}^{(0)}(x) \qquad \text{(C.4a)}$$

$$P\,\bar{\Phi}_{-\lambda}^{(0)}(x) = \bar{\Phi}_{-\lambda}^{(0)}(x) \qquad \text{(C.4b)}$$

$$P\,\bar{\Phi}_{n}^{(0)}(x) = 0 \quad \text{if } E_n^{(0)} \neq \mathcal{E}(X_0) \qquad \text{(C.4c)}$$

We now define

$$\hat{H}_X^{(1)} = H_X^{(1)} - P\, H_X^{(1)} P \qquad \text{(C.5a)}$$

$$\hat{H}_X^{(2)} = H_X^{(2)} + P\, H_X^{(1)} P \qquad \text{(C.5b)}$$

$$\hat{H}_X^{(3)} = H_X^{(3)} \qquad \text{(C.5c)}$$

etc. Thus

$$H_X = H_{X_0} + \hat{H}_X^{(1)} + \hat{H}_X^{(2)} + \ldots \qquad \text{(C.6)}$$

and

$$(\hat{H}_X^{(1)})_{\Lambda\Lambda'} = 0 \qquad \text{(C.7)}$$

We shall treat $\hat{H}_X^{(1)}$ as a first order perturbation, $\hat{H}_X^{(2)}$ as a second order perturbation, etc. The nuclear kinetic energy, K, will, as before, be treated as a second order perturbation.

The wavefunctions $\bar{\Phi}_\Lambda$, $\bar{\Phi}_{-\Lambda}$, and Ψ are now expanded in a formal manner

$$\bar{\Phi}_\Lambda = \bar{\Phi}_\Lambda^{(0)} + \bar{\Phi}_\Lambda^{(1)} + \ldots \qquad \text{(C.8a)}$$

$$\Psi = \Psi^{(0)} + \Psi^{(1)} + \ldots \qquad \text{(C.8b)}$$

as well as the U and the eigenvalue E of K + H_X

$$U_{\Lambda\Lambda'} = U_{\Lambda\Lambda'}^{(0)} + U_{\Lambda\Lambda'}^{(1)} + \ldots \qquad \text{(C.9a)}$$

$$E = E^{(0)} + E^{(1)} + \ldots \qquad \text{(C.9b)}$$

and zeroth, first, second, etc., order equations are obtained in the well-known manner for the step-wise solution of (C.1) and (C.2). The solution of (C.1) proceeds just as outlined in Appendix A - the $H_X^{(k)}$ there being replaced by the $\hat{H}_{\underline{X}}^{(k)}$.

We now proceed to solve equation (C.2) expressing the solutions in terms of the $\bar{\Phi}_\Lambda^{(k)}$. The zeroth order equation

$$(H_{X_0} - E^{(0)})\,\Psi^{(0)} = 0 \tag{C.10}$$

is satisfied by taking

$$E^{(0)} = \mathcal{E}(X_0) \tag{C.11}$$

$$\Psi^{(0)} = \sum \bar{\Psi}_\Lambda^{(0)} \quad \bar{\Phi}_\Lambda^{(0)} \tag{C.12}$$

where the $\bar{\Psi}_\Lambda^{(0)}$ are functions of the vibrational coordinates only and as yet arbitrary.

The first order equation with the substitutions (C.11) and (C.12) and the use of the corresponding first order equation for the solution of (C.1) is found to be

$$\begin{aligned}[H_{X_0} - \mathcal{E}(X_0)][\Psi^{(1)} - \quad \bar{\Phi}_\Lambda^{(1)}\psi^{(0)}]\\ = \sum_\Lambda \bar{\Phi}_\Lambda^{(0)}[E^{(1)}\bar{\Psi}_\Lambda^{(0)} - \sum_{\Lambda'} U_{\Lambda\Lambda'}^{(1)}\bar{\Psi}_{\Lambda'}^{(0)}]\end{aligned} \tag{C.13}$$

In order that this equation be integrable we must require that

$$E^{(1)} \bar{\Psi}_{\Lambda}^{(0)} - \sum U_{\Lambda\Lambda'}^{(1)} \bar{\Psi}_{\Lambda'}^{(0)} = 0 \qquad (C.14)$$

The

$$U_{\Lambda\Lambda'}^{(1)} = (\hat{H}_X^{(1)})^{\psi}_{\Lambda\Lambda'}$$

are zero and we therefore must demand that

$$E^{(1)} = 0 \qquad (C.15)$$

in order that the $\bar{\Psi}_{\Lambda}^{(0)}$ be non-zero. The solution of (C.13) for $\Psi^{(1)}$ is then

$$\underline{\Psi}^{(1)} = \sum_{\Lambda} \bar{\Phi}_{\Lambda}^{(1)} \bar{\Psi}_{\Lambda}^{(0)} + \sum_{\Lambda} \bar{\Phi}_{\Lambda}^{(0)} \bar{\Psi}_{\Lambda}^{(1)} \qquad (C.16)$$

where again the $\bar{\Psi}_{\Lambda}^{(1)}$ are functions of the vibrational coordinates only and yet to be determined.

The solutions of the first and zeroth order equations are now substituted into the second order equation and again use is made of the perturbation equations for the solution of (C.1). This gives us the following equation for the determination of $\Psi^{(2)}$ and $E^{(2)}$:

$$[H_{X_0} - \mathcal{E}(X_0)][\underline{\Psi}^{(2)} - \sum \bar{\Psi}_{\Lambda}^{(1)} \bar{\Phi}_{\Lambda}^{(1)} - \sum \bar{\Psi}_{\Lambda}^{(0)} \bar{\Phi}_{\Lambda}^{(2)}]$$

$$= - \sum_{\Lambda} \bar{\Phi}_{\Lambda}^{(0)} [K\bar{\Psi}_{\Lambda}^{(0)} + \sum_{\Lambda'} U_{\Lambda\Lambda'}^{(2)} \bar{\Psi}_{\Lambda'}^{(0)} - E^{(2)} \bar{\Psi}_{\Lambda}^{(0)}] \qquad (C.17)$$

The integrability condition (found by multiplying by each of the $\bar{\Phi}_{\Lambda}^{(0)*}$ and integrating over the electronic coordinates) for this equation is

$$K\,\bar{\Psi}_{\Lambda}^{(0)} + \sum U_{\Lambda\Lambda'}^{(2)}\,\bar{\Psi}_{\Lambda'}^{(0)} = E^{(2)}\,\bar{\Psi}_{\Lambda}^{(0)} \qquad \text{(C.18)}$$

This is a coupled set of differential equations whose solutions give the $\bar{\Psi}_{\Lambda}^{(0)}$ and $E^{(2)}$. If equation (C.18) is substituted into (C.17) we can immediately solve for $\Psi^{(2)}$:

$$\Psi^{(2)} = \sum \bar{\Psi}_{\Lambda}^{(2)}\,\bar{\Phi}_{\Lambda}^{(0)} + \sum \bar{\Psi}_{\Lambda}^{(1)}\,\bar{\Phi}_{\Lambda}^{(1)} + \sum \bar{\Psi}_{\Lambda}^{(0)}\,\bar{\Phi}_{\Lambda}^{(2)} \qquad \text{(C.19)}$$

Again the $\bar{\Psi}_{\Lambda}^{(2)}$ are functions of the vibrational coordinates yet to be determined.

Let us now consider the third order equation. In terms of the solutions of the zeroth, first, and second order equations this equation may be written in the form

$$\begin{aligned}
&[H_X - \mathcal{E}(X_0)][\Psi^{(3)} - \sum \bar{\Psi}_{\Lambda}^{(0)}\,\bar{\Phi}_{\Lambda}^{(3)} - \sum \bar{\Psi}_{\Lambda}^{(1)}\,\bar{\Phi}_{\Lambda}^{(2)} - \sum \bar{\Psi}_{\Lambda}^{(2)}\,\bar{\Phi}_{\Lambda}^{(1)}] \\
&= - \sum_{\Lambda} \bar{\Phi}_{\Lambda}^{(0)}[K\bar{\Psi}_{\Lambda}^{(1)} + \sum U_{\Lambda\Lambda'}^{(2)}\bar{\Psi}_{\Lambda'}^{(1)} - E^{(2)}\bar{\Psi}_{\Lambda}^{(0)} \\
&\qquad + \sum_{\Lambda'} U_{\Lambda\Lambda'}^{(3)}\bar{\Psi}_{\Lambda'}^{(0)} - E^{(3)}\bar{\Psi}_{\Lambda}^{(0)}] \\
&\qquad - \sum_{\Lambda} (K, \bar{\Phi}_{\Lambda}^{(1)})\,\bar{\Psi}_{\Lambda}^{(0)} \qquad \text{(C.20)}
\end{aligned}$$

In order that this equation be integrable we require

$$K \bar{\Psi}_{\Lambda}^{(1)} + \sum U_{\Lambda\Lambda'}^{(2)} \bar{\Psi}_{\Lambda'}^{(1)} - E^{(2)} \bar{\Psi}_{\Lambda}^{(1)}$$

$$= - \sum U_{\Lambda\Lambda'}^{(3)} \bar{\Psi}_{\Lambda'}^{(0)} + E^{(3)} \bar{\Psi}_{\Lambda}^{(0)} - \sum_{\Lambda'} \int \bar{\Phi}_{\Lambda}^{(0)*} (K, \bar{\Phi}_{\Lambda'}^{(1)}) \, d\tau_e \quad \bar{\Psi}_{\Lambda'}^{(0)} \tag{C.21}$$

The last term on the right hand side is zero (this follows from equation (A.15c)):

$$\int \bar{\Phi}_{\Lambda}^{(0)} (K, \bar{\Phi}_{\Lambda'}^{(1)}) \, d\tau_e = 0 \tag{C.22}$$

The solution of (C.20) for $\underline{\Psi}^{(3)}$ is then

$$\Psi^{(3)} = \sum \bar{\Psi}_{\Lambda}^{(3)} \bar{\Phi}_{\Lambda}^{(0)} + \sum \bar{\Psi}_{\Lambda}^{(2)} \bar{\Phi}_{\Lambda}^{(1)} + \sum \bar{\Psi}_{\Lambda}^{(1)} \bar{\Phi}_{\Lambda}^{(2)} + \sum \bar{\Psi}_{\Lambda}^{(0)} \bar{\Phi}_{\Lambda}^{(3)} + F(x, X) \tag{C.23}$$

where the $\psi^{(3)}$ are as yet undetermined functions of the vibrational coordinates. The function F(x, X) is the particular solution of

$$[\, H_{X_0} - \mathcal{E}(X_0) \,]\, F = - \sum (K, \bar{\Phi}_{\Lambda}^{(1)}) \, \bar{\Psi}_{\Lambda}^{(0)} \tag{C.24}$$

which is orthogonal to all the $\bar{\Phi}_{\Lambda}^{(0)}$ with respect to integration over the electronic coordinates. This term cannot be expressed in terms of the $\bar{\Phi}_{\Lambda}^{(k)}$ and therefore represents the lowest order correction to the wavefunction determined by the Born-Renner approximation.

Let us now proceed to study the fourth order equation. When the solutions of the zeroth, first, second, and third order equations are substituted into this equation and use is made of the equations for the step-wise perturbation solution of equation (C.1) the fourth order equation can be cast in the following form:

$$
\begin{aligned}
(H_{X_0} - \mathcal{E}(X_0)) \Big[\psi_{\Lambda}^{(4)} &- \sum \bar{\psi}_{\Lambda}^{(0)} \; \bar{\Phi}_{\Lambda}^{(4)} - \sum \bar{\psi}_{\Lambda}^{(1)} \; \bar{\Phi}_{\Lambda}^{(3)} \\
&- \sum \bar{\psi}_{\Lambda}^{(2)} \; \bar{\Phi}_{\Lambda}^{(2)} - \sum \bar{\psi}_{\Lambda}^{(3)} \; \bar{\Phi}_{\Lambda}^{(1)} \Big] \\
= - \sum \bar{\Phi}_{\Lambda}^{(0)} \Big[K \bar{\psi}_{\Lambda}^{(2)} &+ \sum U_{\Lambda\Lambda'}^{(2)} \psi_{\Lambda'}^{(2)} - E^{(2)} \bar{\psi}_{\Lambda}^{(2)} \\
&\sum U_{\Lambda\Lambda'}^{(3)} \bar{\psi}_{\Lambda'}^{(1)} - E^{(3)} \bar{\psi}_{\Lambda}^{(1)} + \sum U_{\Lambda\Lambda'}^{(4)} \bar{\psi}_{\Lambda'}^{(0)} \\
&- E^{(4)} \bar{\psi}_{\Lambda}^{(0)} \Big] \\
- H_X^{(1)} F &- \sum (K, \bar{\Phi}_{\Lambda}^{(2)}) \bar{\psi}_{\Lambda}^{(0)} - \sum (K, \bar{\Phi}_{\Lambda}^{(1)}) \bar{\psi}_{\Lambda}^{(1)}
\end{aligned}
\tag{C.25}
$$

Multiplying this equation by one of the $\bar{\Phi}_{\Lambda}^{(0)*}$ and integrating over the electronic coordinates yields the following equation

$$
\begin{aligned}
K \bar{\psi}_{\Lambda}^{(2)} &+ \sum U_{\Lambda\Lambda'}^{(2)} \bar{\psi}_{\Lambda'}^{(2)} - E^{(2)} \bar{\psi}_{\Lambda}^{(2)} + \sum U_{\Lambda\Lambda'}^{(3)} \psi_{\Lambda'}^{(1)} \\
- E^{(3)} \bar{\psi}_{\Lambda}^{(1)} &+ \sum U_{\Lambda\Lambda'}^{(4)} \bar{\psi}_{\Lambda'}^{(0)} - E^{(4)} \bar{\psi}_{\Lambda}^{(0)} \\
+ \sum_{\Lambda'} \int \bar{\Phi}_{\Lambda}^{(0)*} (K, \bar{\Phi}_{\Lambda'}^{(2)}) \, d\tau_e \, \bar{\psi}_{\Lambda'}^{(0)} &+ \int \bar{\Phi}_{\Lambda}^{(0)} H_X^{(1)} F \, d\tau_e \\
&= 0
\end{aligned}
\tag{C.26}
$$

as an integrability condition. (In obtaining the above we have made use of (C.22)). This equation allows a determination of $\bar{\Psi}_{\Lambda}^{(2)}$ and $E^{(4)}$. Its solution is required in order that we know the wavefunction correct to second order (i. e., that we know $\Psi^{(2)}$). A certain amount of simplification of this equation can be achieved. We note that

$$\begin{aligned}\int \bar{\Phi}_{\Lambda}^{(0)*} H_X^{(1)} F \, d\tau_e &= \int H_X^{(1)} \bar{\Phi}_{\Lambda}^{(0)*} F \, d\tau_e \\ &= -\int ([H_X - \mathcal{E}(X_o)] \bar{\Phi}_{\Lambda}^{(1)}) F \, d\tau_e \\ &= -\int \bar{\Phi}_{\Lambda}^{(1)*} [H_X - \mathcal{E}(X_o)] F \, d\tau_e \\ &= \sum \int \bar{\Phi}_{\Lambda}^{(1)} (K, \bar{\Phi}_{\Lambda'}^{(1)}) \, d\tau_e \, \bar{\Psi}_{\Lambda'}^{(0)}\end{aligned} \tag{C.27}$$

from (C.24). Therefore

$$\begin{aligned}&\int \bar{\Phi}_{\Lambda}^{(0)*} H_X^{(1)} F \, d\tau_e + \sum \int \bar{\Phi}_{\Lambda}^{(0)*} (K, \bar{\Phi}_{\Lambda'}^{(2)}) \, d\tau_e \, \bar{\Psi}_{\Lambda'}^{(0)} \\ &= \sum \int \bar{\Phi}_{\Lambda}^{(0)*} [K, (\bar{\Phi}_{\Lambda'}^{(0)} + \bar{\Phi}_{\Lambda'}^{(1)} + \bar{\Phi}_{\Lambda'}^{(2)})] \, d\tau_e \, \bar{\Psi}_{\Lambda'}^{(0)} \\ &\quad + \sum \int \bar{\Phi}_{\Lambda}^{(1)*} [K, (\bar{\Phi}_{\Lambda'}^{(0)} + \bar{\Phi}_{\Lambda'}^{(1)})] \, d\tau_e \, \bar{\Psi}_{\Lambda'}^{(0)} \\ &= \sum_{\Lambda'} \int \bar{\Phi}_{\Lambda}^{*} (K, \bar{\Phi}_{\Lambda'}) \, d\tau_e \, \bar{\Psi}_{\Lambda'}^{(0)}\end{aligned} \tag{C.28}$$

to the same order of approximation.as that entailed in neglecting fifth and higher order perturbations inn our ordering scheme. The expressions

$$\int \bar{\Phi}^{*}_{\Lambda}(K, \bar{\Phi}_{\Lambda'})\, d\tau_e \quad ,$$

however, are the $\mathcal{T}_{\Lambda\Lambda'}$ defined in (3.33). With this identification we replace (C.26) by the "equivalent" equation

$$K\, \bar{\psi}^{(2)}_{\Lambda} + \sum U^{(2)}_{\Lambda\Lambda'}\, \psi^{(2)}_{\Lambda'} - E^{(2)}\, \bar{\psi}^{(2)}_{\Lambda} + \sum U^{(3)}_{\Lambda\Lambda'}\, \bar{\psi}^{(1)}_{\Lambda'}$$
$$- E^{(3)}\, \bar{\psi}^{(1)}_{\Lambda} + \sum_{\Lambda'} [U^{(4)}_{\Lambda\Lambda'} + \mathcal{T}_{\Lambda\Lambda'}]\, \bar{\psi}^{(0)}_{\Lambda'} - E^{(4)}\, \bar{\psi}^{(0)}_{\Lambda}$$
$$= 0 \qquad \text{(C.29)}$$

This is permissible as long as we do not proceed any higher than fourth order.

Let us now examine the results of our perturbation solution of (C.2). The wavefunction to third order is

$$\Psi^{(0)} + \Psi^{(1)} + \Psi^{(2)} + \Psi^{(3)} = \sum \bar{\psi}^{(0)}_{\Lambda} [\, \bar{\Phi}^{(0)}_{\Lambda} + \bar{\Phi}^{(1)}_{\Lambda} +$$
$$+ \bar{\Phi}^{(2)}_{\Lambda} + \bar{\Phi}^{(3)}_{\Lambda}] +$$
$$+ \sum \bar{\psi}^{(1)}_{\Lambda} [\, \bar{\Phi}^{(0)}_{\Lambda} + \bar{\Phi}^{(1)}_{\Lambda} + \bar{\Phi}^{(2)}_{\Lambda}]$$
$$+ \sum \bar{\psi}^{(2)}_{\Lambda} [\, \bar{\Phi}^{(0)}_{\Lambda} + \bar{\Phi}^{(1)}_{\Lambda}] + \sum \bar{\psi}^{(3)}_{\Lambda}\, \bar{\Phi}^{(0)}_{\Lambda}$$
$$+ F(x, X) \qquad \text{(C.30)}$$

This, however, is readily seen to be equivalent to

$$\sum_{\Lambda} [\, \bar{\psi}^{(0)}_{\Lambda} + \bar{\psi}^{(2)}_{\Lambda} + \bar{\psi}^{(3)}_{\Lambda} + \bar{\psi}^{(4)}_{\Lambda}]\, \bar{\Phi}_{\Lambda} + F((x, X) \qquad \text{(C.31)}$$

within the same order of approximation. If we define

$$\bar{\Psi}_{\Lambda} = \bar{\Psi}_{\Lambda}^{(0)} + \bar{\Psi}_{\Lambda}^{(1)} + \bar{\Psi}_{\Lambda}^{(2)} . \ . \ . \ . \qquad (C.32)$$

we can replace the quantity in parentheses in (C.31) by $\bar{\Psi}_{\Lambda}$ to the same order of approximation. Thus

$$\underline{\Psi} = \sum \bar{\Psi}_{\Lambda} \bar{\Phi}_{\Lambda} + F(x, X)$$
$$+ \text{ (fourth order terms)} \qquad (C.33)$$

If we were content to take the wavefunction to second order only we would have

$$\underline{\Psi} = \sum \bar{\Psi}_{\Lambda} \bar{\Phi}_{\Lambda} + \text{(third order terms)} \qquad (C.34)$$

as F(x, X) did not appear until third order. This indicates, in particular that the Born-Renner approximation correctly gives our wavefunctions to second order in our ordering scheme.

The equations (C.14), (C.18), (C.21), and (C.29) which we obtained as "integrability conditions" can easily be seen to be equivalent to the single equation

$$(K + \mathcal{E}(x_0))(\bar{\Psi}_{\Lambda}^{(0)} + \bar{\Psi}_{\Lambda}^{(1)} + \bar{\Psi}_{\Lambda}^{(2)} + . \ . \ . \ .)$$
$$+ [U_{\Lambda\Lambda'}^{(1)} + U_{\Lambda\Lambda'}^{(2)} + U_{\Lambda\Lambda'}^{(3)} + (T_{\Lambda\Lambda'} + U_{\Lambda\Lambda'}^{(4)}) + U_{\Lambda\Lambda'}^{(5)} +$$
$$+ . \ . \ . \ ](\bar{\Psi}_{\Lambda'}^{(0)} + \bar{\Psi}_{\Lambda'}^{(1)} + \bar{\Psi}_{\Lambda'}^{(2)} + . \ . \ . \ .)$$
$$= (E^{(0)} + E^{(1)} + E^{(2)} + . \ . \ .)(\bar{\Psi}_{\Lambda}^{(0)} + \bar{\Psi}_{\Lambda}^{(1)} + \bar{\Psi}_{\Lambda}^{(2)} + . \ . \ .) \qquad (C.35)$$

if we had treated it by perturbation methods taking $\mathcal{U}^{(3)}$ as a first order perturbation and $\mathcal{T} + \mathcal{U}^{(4)}$ as a second order perturbation. Thus the functions $\bar{\Psi}_{\Lambda}$ are the solutions of the equations

$$K \bar{\Psi}_{\Lambda} + \sum U_{\Lambda\Lambda'} \bar{\Psi}_{\Lambda'} + \sum \mathcal{T}_{\Lambda\Lambda'} \bar{\Psi}_{\Lambda'} = E \bar{\Psi}_{\Lambda} \qquad (C.36)$$

where

$$U_{\Lambda\Lambda'} = \mathcal{E}(X_0)\, \delta_{\Lambda,\Lambda'} + U^{(1)}_{\Lambda\Lambda'} + U^{(2)}_{\Lambda\Lambda'} + \cdot\ \cdot\ \cdot \qquad (C.37)$$

Our development shows that the eigenvalues of $K + H_X$ are given correct to fourth order by the eigenvalues of equation (C.36). This equation is simply the Moffitt-Liehr vibronic equation. Since the eigenvalues E of the Moffitt-Liehr quasi-Hamiltonian are identical to those of the corresponding Born-Renner quasi-Hamiltonian and

$$\sum \bar{\Psi}_{\Lambda} \bar{\Phi}_{\Lambda} = \psi^{I} \Phi^{I} + \psi^{II} \Phi^{II}$$

where Φ^{I} and Φ^{II} are eigenstates of H_X and ψ^{I} and ψ^{II} are components of the eigenvectors of the Born-Renner quasi-Hamiltonian we have shown that the Born-Renner approximation gives the eigenfunctions to second order and the energy to fourth order in our ordering scheme. Furthermore, if we were to neglect the term

$$\sum \mathcal{T}_{\Lambda\Lambda'} \bar{\Psi}_{\Lambda'}$$

in (C.36) we would still obtain the energy correctly to third order.

One appealing feature of the method utilized here is that we have now a method for estimating the error incurred by making the Born-Renner approximation. The percentage of error in our wavefunction is essentially

$$\int |F(x, X)|^2 \, d\tau_e \, d\tau_N$$

where the integration is carried out over all the electronic and nuclear coordinates. Thus we can test the validity or invalidity of the Born-Renner approximation by considering the solution of (C.24).

Appendix D

DYNAMICAL JAHN-TELLER EFFECT IN FROHLICH'S ONE-DIMENSIONAL MODEL OF A SUPERCONDUCTOR

Here we shall treat a one-dimensional model of a superconductor originally studied by Frohlich [Ref. 9] along lines similar to those used in treating the molecular cases in Chapters 4 and 5. We shall show that this model exhibits a Jahn-Teller effect in that the minimum of the lowest potential surface occurs when the atoms are not equally spaced. A consideration of the dynamical aspects of the model will show that the motion of the system on the lowest quasi-potential surface will induce an electric current.

1. The Lowest Quasi-Potential Surface.

We consider a one-dimensional model of a crystal - the nuclei being initially assumed to be spaced a distance a apart (this corresponds to the configuration X_o). We take the Hamiltonian for the crystal to be of the form

$$H = - \sum_n \frac{\hbar^2}{2m} \frac{d^2}{dx_n^2} + \sum_n V(x_n)$$
$$+(1/2) \sum_\kappa (P_\kappa^* P_\kappa + \omega^2 Q_\kappa^* Q_\kappa)$$
$$+ \frac{1}{\sqrt{NM}} \frac{2}{3}C \sum_{\kappa,n} e^{i\kappa x_n} Q_\kappa \qquad (D.1)$$

Here M is the mass of a lattice ion, N is the number of ions in the lattice, x_n denotes the coordinate of the n-th electron. The Q_κ are the normal vibrational coordinates and are related to the coordinates of the ions as follows

$$X_n - X_{n,o} = \frac{1}{\sqrt{MN}} \sum_\kappa Q_\kappa e^{i\kappa na} \qquad (D.2)$$

Here $X_{n,o}$ denotes the equilibrium position of the n-th ion when the configuration X_o is assumed. The wave number κ ranges over the values

$$-\pi\frac{(N-2)}{Na}, -\pi\frac{(N-4)}{Na}, 1 \ldots ., 0, \ldots \pi\frac{N}{Na}$$

if N is even and over the values

$$-\pi\frac{(N-1)}{Na}, -\pi\frac{(N-3)}{Na}, \ldots ., 0, \ldots \pi\frac{(N-1)}{Na}$$

if N is odd. The P_κ are the momentum operators corresponding to the Q_κ and are

$$P_\kappa = \frac{\hbar}{i} \frac{d}{dQ_\kappa} \qquad (D.3)$$

The ω_κ are the vibrational frequencies in the absence of the electron-lattice interaction. The function $V(x_n)$ is periodic in <u>a</u>:

$$V(x_n) = V(x_n+a) \tag{D.4}$$

The constant C is the Bethe-Sommerfeld [Ref. 47] constant giving the strength of the electron-lattice interaction.

This is essentially the model proposed by Frohlich - although he chose to write it in a second-quantized form rather than as presented here. It is assumed that the Hamiltonian only explicitly contains the coordinates of the conduction electrons - the remaining electrons being considered bound to the lattice ions and contributing to the Hamiltonian through the parameters M, ω_κ , and E and through the form of the periodic potential $V(x_n)$.

Let us now seek to determine the lowest eigenvalue for arbitrary Q_κ of the electronic Hamiltonian H_X (the Hamiltonian (D.1) minus the nuclear kinetic energy operator). We shall then examine the energy surface represented by this eigenvalue as a function of the vibrational coordinates and show that it attains its minimum at a configuration other than that where all the Q_κ are identically zero -

regardless of the magnitude of the electron-lattice coupling.

Let us first make a change of variables - letting

$$Q_\kappa = \left(\frac{\hbar}{2\omega_\kappa}\right)^{(1/2)} r_\kappa \, e^{i\varphi_\kappa} \tag{D.5}$$

if Q_κ corresponds to a degenerate vibration and

$$Q_\kappa = \left(\frac{\hbar}{\omega_\kappa}\right)^{(1/2)} \tag{D.6}$$

if Q_κ corresponds to a non-degenerate vibration. Rather than examine the lowest eigenvalue of H_X as a function of all the vibrational coordinates simultaneously let us instead examine it as a function of only one pair of degenerate vibrations (coordinates r_κ and φ_κ). If the energy surface has a minimum when r_κ is non-zero along the line where all the remaining vibrational coordinates are zero our assertion will be verified and we will need consider no other vibrational modes. A judicious choice of this vibrational mode is that corresponding to

$$\kappa = 2k_F \quad (\text{Mod } 2\pi/a)$$

A vibrational distortion corresponding to such a vibrational mode will tend to split the energy band at the Fermi level.

The Hamiltonian H_X with all the other vibrational coordinates being zero is then

$$H_X = -\sum_n \frac{\hbar^2}{2m}\frac{d^2}{dx_n^2} + \sum_n V(x_n) + \frac{1}{2}\hbar\omega\, r^2 + 2Dr\sum_n \cos(2k_F x_n + \varphi) \qquad (D.7)$$

Here we have, for convenience, abbreviated r, φ, and ω for r_{2k_F}, φ_{2k_F}, and ω_{2k_F} and have defined

$$D = \left(\frac{\hbar}{2\,\omega_{2k_F}}\right)^{1/2}\frac{1}{\sqrt{NM}}\;\frac{2}{3}C\;2k_F \qquad (D.8)$$

To find the eigenvalues of (D.7) for fixed r and it is clear that we must first solve the single particle equation

$$\left\{-\frac{\hbar^2}{2m}\frac{d^2}{dx_n^2} + V(x_n) + 2Dr\cos(2k_F x_n + \varphi)\right\}\Phi = E\,\Phi \qquad (D.9)$$

Schroedinger equations of the above form have studied frequently in connection with solid state problems and, in particular, Frohlich has essentially solved this equation by perturbation methods.

Following Frohlich, we find a energy gap 2Dr at $k = k_F$ and we find that the energy levels below this gap are approximately given by

$$E_k = \xi(1-|y|)^2 - 2\xi\left[(1-|y|)^2 + \left(\frac{Dr}{2\xi}\right)^2\right]^{1/2} - \frac{(Dr)^2}{4\xi}\,\frac{1}{1+|y|} \qquad (D.10)$$

where we have abbreviated y for k/k_F. Here k denotes the wavenumber of the unperturbed single-particle state, while ζ and k_F denote the energy and wavenumber, respectively, of an electron at the Fermi level.

The lowest eigenvalue of H_X is then given by

$$\begin{aligned}\mathcal{E}(r) &= 4\sum_{k=0}^{k_F} E_k + \frac{1}{2}\hbar\omega\, r^2 \\ &\doteq -\frac{2}{3}N_e + \frac{1}{2}\hbar\hat{\omega}\, r^2 + N_e\zeta\left(\frac{Dr}{2\zeta}\right)^2 \ln\left(\frac{Dr}{2\zeta}\right)\end{aligned} \qquad (D.11)$$

where

$$\frac{1}{2}\hbar\hat{\omega} - \frac{1}{2}\hbar\omega = \frac{1}{2}N_e\zeta\frac{D^2}{(2\zeta)^2} - 2N_e\zeta\frac{D^2}{(2\zeta)^2}\ln 2 \qquad (D.12)$$

The quantity $\hat{\omega}$ is a "renormalized" vibrational frequency. It is to be expected that it will be very nearly ω as generally

$$\hbar\omega \gg (N_e D^2/\zeta) = (16/9)(N_e/N)(m/M)(C^2/\hbar\omega) \qquad (D.13)$$

In any event we shall assume ω to be positive as otherwise the lattice would be essentially unstable.

The nature of the logarithmic term is such that for sufficiently small values of r it completely dominates the term $(1/2)\hbar\hat{\omega}r^2$. In general

$$-x^2 \ln x > Ax^2 \qquad (D.14)$$

for any given A if x is small enough. This fact indicates that (D.11) will always attain its minimum at a non-zero value of r. This confirms our original assertion that there is a Jahn-Teller effect exhibited by this model.

The value of r for which (D.11) is a minimum may be found by equating its first derivative with zero. Doing this yields

$$r[\hbar\hat{\omega} + (N_e D^2/2\rho) \ln \frac{e^{1/2} Dr}{2\rho}] = 0 \tag{D.15}$$

The solution which corresponds to the minimum is

$$r_o = \frac{2\rho}{D} e^{-1/2} e^{-(2\rho\hbar\hat{\omega}/N_e D^2)} \tag{D.16}$$

with the corresponding electronic energy of

$$E_{min} = -(2/3)N_e\rho - N_e e^{-4\rho\hbar\hat{\omega}/N_e D^2} \tag{D.17}$$

This is always less than the "normal electronic state" energy (that when r=0) which is $-(2/3)N_e\rho$. The displacement of the n-th lattice ion when the vibrational coordinate r has the value r_o is

$$X_n - X_{n,o} = \frac{3\rho}{Ck_F} e^{-1} e^{-4\rho\hbar\omega/N_e D^2} \cos(2k_F Na + \varphi)$$

The coordinate φ is arbitrary in the approximation giving (D.10).

2. Current Carried by Low Energy States.

The preceding discussion indicates that the $k = 2k_F$ vibrational modes will be strongly excited. In accordance with this we shall make the assumption that the contribution to the wavefunction for the lattice as a whole from the other vibrational degrees of freedom may be factored out and that the interaction of the electronic motions with these modes may be ignored. Thus, as in the molecular case, we shall consider a truncated Hamiltonian which depends on only one two-fold degenerate vibrational mode. This approximation was also made by Frohlich in his discussion of the model.

At first, we shall assume that the Born-Oppenheimer approximation applies rigorously and therefore assume that the nuclear part of the wavefunction for low lying energy states is found by taking the nuclei as moving in a quasi-potential which is the lowest eigenvalue of H_X:

$$\mathcal{E}(r) = (1/2)\hbar\,\hat{\omega}\,r^2 \quad + N_e\,\zeta\left(\frac{Dr}{2\zeta}\right)^2 \ln\left(\frac{Dr}{2\zeta}\right)) \tag{D.18}$$

The constant $-(2/3)N_e\zeta$ has been ignored as our reference point of energy is arbitrary.

Schroedinger's equation for the motion of the system in this quasi-potential is

$$[- \frac{\hbar\omega}{2} \left(\frac{d^2}{dr^2} + \frac{1}{r} \frac{d}{dr} + \frac{1}{r^2} \frac{d^2}{d\varphi^2} \right) + \mathcal{E}(r)] \, \psi(r, \varphi)$$

$$= E \, \psi(r, \varphi) . \qquad (D.19)$$

Itiis readily seen that the solutions of this equation are of the form

$$\psi = \frac{1}{\sqrt{2\pi}} \quad e^{ij\varphi} \; R(r)$$

The radial part of the wavefunction may be found using approximations similar to those used in the latter part of Section D in Chapter 4. For low energy states it should be appreciably different from zero only in the region where $r = r_o$.

If $j \neq 0$ we may picture thessystem as revolving around the trough of the quasi-potential surface $\mathcal{E}(r)$ with angular momentum $\hbar j$. From considering the analogy with the molecular case we expect that this motion will induce an electric current. This turns out to indeed be the case. While the wavefunction given by the Born-Oppenheimer approximation

$$\psi(r, \varphi) \;\; \Phi_o(x, r, \varphi)$$

(Φ_o denoting the ground electronic eigenstate of H_X)

is such that the expectation value of the current operator

$$\langle I \rangle_{op} = \frac{\hbar e}{imK} \sum \frac{d}{dx}$$

vanishes, it must be remembered that the Born-Oppenheimer state is not a true eigenstate of the Hamiltonian and that there will always be nonvanishing terms corresponding to the system being in electronic states other than Φ_0. We may find these terms by solving Born's equations (see Section D of Chapter 2). Doing this and making approximations analogous to (4.73), (5.24), and (5.25) gives the following expression for the expectation value of the current

$$I = \frac{-2}{\Delta} \mathrm{Re} \int \psi^* \Phi_0 \langle I \rangle_{op} (K, \Phi_0) \, \psi \, d\tau_e \, r dr \, d \qquad (D.20)$$
$$(\Delta = 2Dr)$$

The wavefunction Φ_0 may be written as product of the eigenfunctions of (D.9). In turn, the single-particle eigenfunctions may be calculated by perturbation theory. When this is done and the integrations over the electronic coordinates in the above expression are carried out we obtain the following expression:

$$I = \frac{-2\hbar^2}{M} \frac{v_F \, \bar{e} \, N_e}{\Delta L} \frac{D}{2} \frac{\pi}{4} j \int \psi^* \frac{1}{r} \psi \, r dr \, d\varphi \qquad (D.21)$$

where $v_F = (\hbar k_F/m^*)$ is the velocity of an electron at the Fermi level. Since, for low energy states, ψ is only appreciably diffenent from zero when $r = r_0$ we may take (as we did in (5.28))

$$\int \psi^* \frac{1}{r} \psi \quad r dr \, d\varphi \doteq \frac{1}{r_0}$$

Thus

$$I = \frac{\hbar^2 j}{M \Delta r_0^2} I_0 \qquad \text{(D.22)}$$

just as in (5.30) - providing we identify

$$I_0 = - \frac{\pi}{2} \frac{v_F \; e \; N_e}{L} \frac{D r_0}{2\rho} \qquad \text{(D.23)}$$

The current (D.22) is of appreciable magnitude only if j is of the order of N, the total number of ions in the lattice.

BIBLIOGRAPHY

1. Bardeen, J., Cooper, L., and Schrieffer, J.R., *Phys. Rev.*, *108* (1957), p. 1175

2. Bogoliubov, N.N., Tolmachev, V.V., and Shirkov, D.V., *A New Method in the Theory of Superconductivity*, (New York, 1959), translated from the Russian text (Moscow, 1958).

3. Renner, R., *Z. Physik*, *92*, (1934) p. 172.

4. Jahn, H.A. and Teller, E., *Proc. Roy. Soc.*, *A161*, (1937) p. 220.

5. Moffitt, W., and Liehr, A.D., *Phys. Rev.*, *106* (1937) p. 1195.

6. Longuet-Higgins, R.C., Opik, U., Pryce, M.H.L., and Sack, R.A., *Proc. Roy. Soc.*, *A224*, (1958) p. 1.

7. Liehr, A.D., *Interaction of the Vibrational and Electronic Motions in Some Simple Conjugated Hydrocarbons: Part III. A Simple Formulation.* (Unpublished.)

8. Tisza, L., *Annals of Physics*, *13* (1961) p. 1.

9. Frohlich, H., *Proc. Roy. Soc.*, *A223*, (1954), p. 296.

10. Born, M. and Oppenheimer, J.R., *Ann. Phys.*, *Lpz.*, *84*, (1927), p. 457.

11. Nielson, H.H., *Rev. Mod. Phys.*, *23*, (1951), p. 90.

12. Born, M., and Huang, K., *Dynamical Theory of Crystal Lattices*, (Oxford, 1954), Chapter IV and Appendix VIII.

13. Haug, A. and Sauermann, G., *Z. Phys.*, *153*, (1956), p. 269.

14. Schultz, T.D., (Unpublished).

15. Sponer, H. and Teller, E., *Rev. Mod. Phys.*, *13*, (1941), p. 75.

16. El'yashevich, M.A., *Soviet Physics*, *Uspekhi*, *3*, (1960), p. 440.

17. Markham, J.J., Phys. Rev., 103, (1956), p. 588.

18. Newton, R.R. and Thomas, L.H., J. Chem. Phys., 16, (1948), p. 310.

19. Seitz, F., Modern Theory of Solids, (New York, 1940), p. 470.

20. Pelzer, H., and Wigner, E., Z. Phys. Chem., 15B, (1932), p. 445.

21. Jahn, H.A., Proc. Roy. Soc., A164, (1938), p. 117.

22. Herzberg, G., Spectra of Diatomic Molecules, 2nd Ed., (Princeton, 1950).

23. Longuet-Higgins, H.C. and McEwen, K.L., J. Chem. Phys., 26, (1957), p. 719.

24. Bloch, F., Z. Phys., 52, (1928), p. 555; 59, (1930), p. 208.

25. Ziman, J.M., Proc. Camb. Phil. Soc., 51, (1955), p. 707.

26. Haug, A., Z. Phys., 146, (1956), p. 75.

27. Goodman, B., Phys. Rev., 110, (1958), p. 888.

28. Chester, G.V., and Houghton, A., Proc. Phys. Soc., 73, (1959), p. 609.

29. Frohlich, H., Proc. Roy. Soc., 215A, (1952), p. 291.

30. Schultz, T.D., Electron-Lattice Interactions in Polar Crystals, Solid State and Molecular Theory Group, Massachusetts Institute of Technology, Technical Report No. 9, (1956). (This is a comprehensive review of the properties of the polaron.)

31. Born, M., Gottingen Nachrichten, II. Math.-Phys. Chem. Abteilung, Nr. 6, 1, (1951).

32. Moffitt, W. and Thorson, W., Phys. Rev., 108, (1957), p. 1251.

33. Hobey, W.D.and McLachan, A.D., J. Chem. Phys., 33, (1960), p. 1695.

34. Pople, J.A. and Longuet-Higgins, H.C., Molecular Physics, 1, (1958), p. 372.

35. Liehr, A.D., "*An Essay on Higher Order Vibronic Interactions*." (Unpublished.)

36. Klein, M.J., *Am. J. of Phys.*, *20*, (1952), p. 65.

37. Van Vleck, J., *J. Chem. Phys.*, *7*, (1939), p. 72.

38. Opik, U. and Pryce, M.H.L., *Proc. Roy. Soc.*, *A238*, (1957), p. 425.

39. Teller, E., *Annals of the New York Academy of Sciences*, *16*, (1941), p. 173.

40. Liehr, A.D., *Z. Physik. Chem.*, *9*, (1956), p. 338.

41. Snyder, L.C., *J. Chem. Phys.*, *33*, (1960), p. 619.

42. Moffitt, W. and Thorson, W.R., *Calcul des Fonctions d'Onde Moleculaire*, Receuil de Memoires, Centre Nationale de la Recherche Scientifique, (1958), p. 141.

43. McConnell, H.M. and McLachlan, A.D., *J. Chem. Phys.*, *34*, (1961), p. 1.

44. Shafroth, M. R., *Solid State Physics*, Vol. 10, p. 404, (New York, 1960).

45. Tisza, L., *Proc. Nat. Bur. Stds. Circular 519, Symposium on Low Temperature Physics*, (1952), p. 11.

46. Slater, J.C., *Phys. Rev.*, *84*, (1951), p. 179.

47. Bethe, H., and Sommerfeld, A., *Handbuch der Physik*, *24/2*, (1933).