

Vibrational Wave Functions for Anharmonic Crystals*

THOMAS H. WALNUT

Department of Chemistry, Syracuse University, Syracuse, New York

(Received July 27, 1959; revised manuscript received August 25, 1961)

An analysis is made of the way in which the behavior of various wave functions approximates that of exact vibrational eigenfunctions of an anharmonic crystal. The criterion used in determining the closeness of approximation is the order of magnitude of the difference between the second moment and the square of the first moment of the Hamiltonian operator. The functions chosen are products of harmonic oscillator wave functions, three different linear combinations of products of harmonic oscillator functions differing slightly in form from the exact eigenfunction, and a function which is the result of operating on an exact eigenfunction with a dipole moment operator.

The first of these functions gives a value of the criterion which is proportional to the number of unit cells in the crystal. For the

rest the criterion is independent of the size of the crystal but is still not zero.

It is shown that, despite the fact that product functions are very poor approximations to eigenfunctions, quantities such as vibrational absorption intensities and frequencies calculated with them should give approximate agreement with experiment and that perturbation theory calculations using them can also give approximately correct results.

It is also shown that the results obtained by using the product wave functions have intrinsic limitations on their accuracy that arise from the collective interaction of all the phonons or modes of vibration.

I. INTRODUCTION

IN most treatments of the vibrations of atoms in crystals, harmonic oscillator wave functions are used. This procedure has been successful in treating vibrational absorption, emission, or Raman scattering frequencies and intensities of anharmonic crystals. The basic technique has usually consisted of some form of perturbation theory carried out no further than first order. When higher order approximations are carried out by straightforward applications of the theory, divergences appear and hence the results of the first-order treatment cannot be rigorously correct. This observation has been discussed in a previous paper¹ (hereafter called I) by the writer and also has been well known for other systems from the work of many authors in the field of many-body perturbation theory.²

The fact that perturbation theory gives satisfactory results in some cases suggests that it should be worthwhile to investigate the reasons and to find the limitations on its applicability. In a sense, this investigation has already been made in the work on many-body perturbation theory, but for vibrational problems the basic question of convergence is unanswered. The convergence problem has been changed from one where there is clearly divergence to one where the convergence is in doubt.

In this paper an approach quite different from that of many-body perturbation theory is used. A simple criterion is used as a measure of the suitability of an approximating wave function. The investigation is restricted primarily to those properties which involve transitions from one state to another such as the intensities and frequencies of absorption and emission.

The criterion used is the value of

$$\langle H^2 \rangle_{av} - (\langle H \rangle_{av})^2, \quad (1)$$

* Supported by the National Science Foundation.

¹ T. H. Walnut, *J. Chem. Phys.* **31**, 361 (1959).

² See for example, J. Goldstone, *Proc. Roy. Soc. (London)* **A239**, 267 (1957).

where H is the Hamiltonian operator. For any exact eigenfunction of H the criterion is zero and for any function which is not an eigenfunction it is different from zero. This paper will be concerned with the gross order of magnitude of the criterion for various types of approximating wave functions. The term gross order of magnitude refers to the way in which the size varies with N , the number of unit cells in the Born-von Karman parallelepiped. For example, the orders of magnitude may be N , $N^{\frac{1}{2}}$, $N^{-\frac{1}{2}}$, etc.

Let us first apply the criterion to a product of harmonic oscillator wave functions designated by $\Phi(M)$. Here, M represents the set of all quantum numbers for the individual harmonic oscillator functions. A more detailed description is given in I.

With the criterion succinctly expressed as $\langle H_a^2 \rangle_{av}$, where

$$H_a \equiv H - \langle H \rangle_{av}, \quad (2)$$

it is found that

$$\begin{aligned} \langle H_a^2 \rangle_{av} &= \int \Phi^*(M) H_a^2 \Phi(M) d\tau \\ &= \sum_J \langle M | H_a | J \rangle \langle J | H_a | M \rangle. \end{aligned} \quad (3)$$

The symbol J , like M , represents a set of quantum numbers and the sum over J refers to the sum over all possible sets of values of J . From the same reasoning that was used to show that \sum_J of Eq. (31) of I was of order N , it follows that

$$\langle H_a \rangle_{av}^2 = O(N). \quad (4)$$

This equation means that there is a set of energies associated with the function $\Phi(M)$ and that the mean-square deviation of a measurement of energy from the average is of order N . The root-mean-square deviation of the energy is then of order $N^{\frac{1}{2}}$. Since the total vibrational energy is of order N , it would be reasonable to use the result of any measurement of the energy as the energy for the crystal in the state $\Phi(M)$. The fractional

deviation of any such measurement is very small (of order N^{-1}).

On the other hand, if an attempt is made to obtain the difference between the energies of two closely related states $\Phi(M)$ and $\Phi(J)$ as the difference between a measured energy of $\Phi(M)$ and a measured energy of $\Phi(J)$, meaningless results are obtained. For example, if the two states were the upper and lower states of an optical transition, i.e., where only one of the quantum numbers in J differed from the corresponding numbers in M , the difference in the measured values of the energy of the two states would lie in a range of order $N^{1/2}$ whereas the energy of observable optical transitions is of order 1. Therefore, use of harmonic oscillator functions is not satisfactory when spectroscopically observable energy differences are desired. It should be noted that these results follow for any local anharmonic interaction that is independent of the crystal size, when N is sufficiently large. It would, of course, be possible that in the case of extremely small anharmonicity the size of N for which these results would be so would correspond to a crystal that is much larger than those ordinarily treated.

The size of the criterion can also be used to give information on the relationship between the functions $\Phi(M)$ and the exact eigenfunctions $\Psi(L)$. In determining this relationship it is necessary to estimate roughly the number of eigenstates lying in an energy range of order N . The estimate can be made readily from the fact that the number is roughly the same as that for the unperturbed energies of the states described by the functions $\Phi(M)$, a quantity whose logarithm is of order N . This rough order will be designated by the symbol \mathfrak{N} .

The range of energies of a function $\Phi(M)$ is of order $N^{1/2}$. This range contains a number of eigenstates of rough order \mathfrak{N} (differing from the number in a range of order N by a negligible factor of order $N^{-1/2}$). Of these states only those with the proper symmetry can contribute to a function $\Phi(M)$, and essentially all of the contributing functions must lie in an energy range of order $N^{1/2}$. If we examine the eigenfunctions of the appropriate symmetry in this energy range, it is apparent that no one of them is more closely related in any discernible way to an unperturbed function $\Phi(M)$ than any other. Therefore, it can be reasonably assumed that the contributions of the eigenfunctions to $\Phi(M)$ are of the same gross order of magnitude. Hence, it can be concluded that the number of eigenfunctions $\Psi(L)$ that make collectively significant contributions to a product function $\Phi(M)$ is of order \mathfrak{N} and that the projection of any eigenfunction on a product function is of order \mathfrak{N}^{-1} or less. The effect of symmetry is to introduce a negligible factor of N^{-1} . It is, of course, possible that by chance there might be a projection of order 1 by one function $\Psi(L)$ on a particular product function $\Phi(M)$, but according to the basic assumption stated above the probability is negligibly small.

The result obtained above means that in the ordinary sense $\Phi(M)$ is no approximation at all to $\Psi(L)$. This

statement can be understood by considering what is meant by an approximation. In a crude sense, two functions approximate each other when at each point of the configuration space the values of the two functions are approximately equal. In a more formal way, the functions $\Phi(M)$ and $\Psi(L)$ could be said to approximate each other if the integral

$$\int |\Phi(M) - \Psi(L)|^2 d\tau \quad (5)$$

were small. For normalized functions with arbitrary phase factors, the worst possible approximation of one function for another is one where the functions are orthogonal. Since the projection of $\Phi(M)$ on any function $\Psi(L)$ is extremely small, the functions are almost completely orthogonal and, hence, it can be stated that the two functions do not approximate each other. Thus there is essentially no correlation between the values of $\Phi(M)$ and of any eigenfunction $\Psi(L)$ for all of the points in the configuration space of the system. Presumably the same result would be obtained for any other reasonable definition of approximation for either the wave functions themselves or the probability densities. However, there is no intrinsic justification for obtaining a function which gives a good approximation to an exact eigenfunction since such a function would give probabilities of finding particular sets of simultaneous positions of all of the atoms in a crystal, information which is far too detailed to be useful.

This analysis suggests that obtaining an exact eigenfunction for a complex problem is a procedure that is not useful in itself and that attempts should be made to obtain observable quantities without obtaining the wave functions. A possible approach would be to find functions which would behave in a manner similar to that of the exact eigenfunctions with respect to various operators. A function $\Phi(M)$, for example, when operated on by H yields a set of energies whose fractional deviation is small and, hence, it can be used to represent an exact eigenfunction in a measurement of the total vibrational energy of the crystal despite the fact that it is a very poor approximation.

In the parts of this paper that follow an investigation is made of the ways in which the functions $\Phi(M)$ represent the functions $\Psi(L)$ in the determination of intensities and energies of transitions. In the investigation slightly modified exact eigenfunctions are formally expressed and the criterion of (1) is used to determine the way in which they can be used to represent the behavior of true eigenfunctions.

II. ONE MODIFICATION OF AN EXACT EIGENFUNCTION

Any of the exact eigenfunctions can be written as

$$\Psi(L) = \sum_M a(L, M) \Phi(M). \quad (6)$$

Since such an enormous number of $\Phi(M)$'s contribute

to a given $\Psi(L)$, it is both a hopeless and a useless task to express the coefficients $a(L, M)$ in any explicit way. What can be done is to express an analog of an excited state of a particular eigenfunction in terms of that eigenfunction. This analogous function can be formally expressed as

$$\Psi_b(L+S) \equiv \sum_M a(L, M) \Phi(M+S). \quad (7)$$

The coefficients $a(L, M)$ are identical with those of $\Psi(L)$ and S represents a set of additions to the quantum numbers in L and M . In the approximation the number of nonzero elements of S is of order 1 (i.e., is small compared to $N^{\frac{1}{2}}$) and their magnitudes are of order 1. In the notation of I these conditions are expressed as $[S] = O(1)$. In addition, all of the nonzero elements of S are assumed to be positive.

The validity of $\Psi_b(L+S)$ as a representation of an excited state can be checked with the criterion of (1). The first step in its evaluation is the comparison of $H\Psi_b(L+S)$ with $H\Psi(L)$. Since $\Psi(L)$ is an exact eigenfunction it follows that

$$H\Psi(L) = \sum_{MJ} a(L, M) \langle J | H | M \rangle \Phi(J) = E(L) \Psi(L), \quad (8)$$

and hence, that

$$\sum_M a(L, M) \langle J | H | M \rangle = E(L) a(L, J). \quad (9)$$

When H is applied to $\Psi_b(L+S)$, the equation

$$H\Psi_b(L+S) = \sum_{MJ} a(L, M) \times \langle J+S | H | M+S \rangle \Phi(J+S) \quad (10)$$

is obtained. To compare this equation with Eqs. (8) and (9), the difference

$$\langle J | \Delta H_S | M \rangle \equiv \langle J+S | H | M+S \rangle - \langle J | H | M \rangle \quad (11)$$

can be defined. In terms of it, Eq. (10) becomes

$$H\Psi_b(L+S) = \sum_{MJ} a(L, M) \langle J | H | M \rangle \Phi(J+S) + \sum_{MJ} a(L, M) \langle J | \Delta H_S | M \rangle \Phi(J+S), \quad (12a)$$

or

$$H\Psi_b(L+S) = E(L) \Psi_b(L+S) + \sum_{MJ} a(L, M) \langle J | \Delta H_S | M \rangle \Phi(J+S). \quad (12b)$$

To find the magnitude of the difference between the values of $\langle H \rangle_{av}$ for the two functions the numbers of matrix elements of different orders of magnitude must be determined. The simplest set of matrix elements to consider are those where the sum of the absolute values of the differences of the corresponding elements of M and J is equal to 3. In the notation of I these are matrix elements where $[\Delta M] = 3$.

In I it was shown that the terms in the potential function that can contribute to an element $\langle J | H | M \rangle$ where $[\Delta M] = 3$ has the form

$$\text{const} q(f_1) q(f_2) q(f_3), \quad (13a)$$

or

$$\text{const} q(f_1) q(f_2) q(f_3) \prod_i |q(g_i)|^2, \quad (13b)$$

where $q(f)$ and $q(g)$ are normal coordinates. The three quantum numbers of M that differ from the corresponding numbers of J are those for $q(f_1)$, $q(f_2)$, and $q(f_3)$.³ Consequently, terms of the forms shown above can contribute to only one element having a given set of values for J when $[\Delta M] = 3$.

The only quantum numbers on which the contribution of a term of the form shown in (13a) depends are those for $q(f_1)$, $q(f_2)$, and $q(f_3)$. Therefore the contribution to $\langle J+S | H | M+S \rangle$ and to $\langle J | H | M \rangle$ must be the same for all terms where none of the f 's are in S , the set of values of f which specify the nonzero elements of S . Hence only those cubic terms where at least one of the f 's is in S can contribute to $\langle J | \Delta H_S | M \rangle$. Since, as was found in I, the magnitude of the contribution of the appropriate cubic term to an element $\langle J | H | M \rangle$ is of order $N^{-\frac{1}{2}}$, the nonvanishing contributions to $\langle J | \Delta H_S | M \rangle$ are also of order $N^{-\frac{1}{2}}$.

Similar statements can be made about the contributions of the terms shown in (13b). Here, the total contribution to a given element $\langle J | \Delta H_S | M \rangle$ of terms where the f 's are not in S is not zero, but is still negligible because of its small magnitude, i.e., of order $N^{-\frac{1}{2}}$.

The number of cubic terms that can contribute to the elements $\langle J | \Delta H_S | M \rangle$ is of order N , for there are of order one possible choices of f_1 , such that it is in S , of order N possible choices of f_2 for a given f_1 , and, from the translational symmetry of the crystal, of order 1, choices of f_3 for a given f_1 and f_2 . Consequently, the number of elements with a given J to which the cubic terms contribute is of order N . By similar reasoning, analogous to that in I, it is found that the number of elements $\langle J | \Delta H_S | M \rangle$ for $[\Delta M] = 3$ and for a given J which are not negligible is of order N and their magnitude is of order $N^{-\frac{1}{2}}$.

In a similar way it can be shown that for $[\Delta M] = n > 2$ the number of matrix elements is of order N^{n-2} and their magnitude is of order $N^{(1-\frac{1}{2}n)}$.

When $[\Delta M] = 0$, one element of order one is obtained. The proof of this statement is the same as that given in I for the magnitude of the denominator of Eq. (31).

When $[\Delta M] = 1$ and $[\Delta M] = 2$ the problem is more complex. The difference arises from using self-consistent field normal coordinates in expressing the Hamiltonian. These coordinates are discussed in another paper⁴ (hereafter known as II). They are necessary to reduce the magnitudes of the elements $\langle J | H | M \rangle$ when $[\Delta M] = 1$ and $[\Delta M] = 2$ from order $N^{\frac{1}{2}}$ to order $N^{-\frac{1}{2}}$ and from order 1 to order N^{-1} , respectively.

These coordinates are defined for a particular set of values for M (i.e., M_0). The matrix elements $\langle M | H | J \rangle$ remain reduced to the same orders of magnitude for

³ Strictly speaking there are six quantum numbers associated with these coordinates, but from the more cumbersome but more exact treatment of I it is apparent that the correct results will be obtained if three effective quantum numbers are considered.

⁴ T. H. Walnut, J. Chem. Phys. 31, 1468 (1959).

all cases where

$$[M - M_0] = O(1). \quad (14)$$

However, the range of sets of values of M is larger than this. It is such that the average energies for the functions $\Phi(M)$ extend over a range of $N^{\frac{1}{2}}$. The existence of such a spread in energy is consistent with a situation where a large fraction of the quantum numbers of M differ from the corresponding numbers of M_0 , but where the differences consist of a random set of positive and negative values. Thus, while the actual value of $[M - M_0]$ may be of order N , it behaves as if it were of order $N^{\frac{1}{2}}$ because of the cancellation of the effects of the differences in quantum numbers.

The sum of terms which contribute to a matrix element $\langle J | H | M \rangle$ where $[\Delta M] = 1$ has the form

$$(a + \sum_g b_g |q(g)|^2 + \sum_{gg'} c_{gg'} |q(g)|^2 |q(g')|^2 \cdots) q(f). \quad (15)$$

The self-consistent field coordinates are so chosen that the average value of the sum which is the coefficient of $q(f)$ in this equation is of order $N^{-\frac{1}{2}}$ for the set of quantum numbers M_0 . For other sets of elements that differ from it in effectively of order $N^{\frac{1}{2}}$ elements, the average value is larger.

The constant term a does not contribute to the difference in average value. The sum $\sum_g b_g |q(g)|^2$ contributes of order $N^{\frac{1}{2}}$ terms of order $N^{-\frac{1}{2}}$ or a total of order 1. Similarly the total contribution of the higher terms is of the same order. Therefore, for most of the allowed sets of values of M , the average value of the sum is of order 1.

By the arguments of II it can be shown that the nonvanishing elements $\langle J | H | M \rangle$ where $[\Delta M] = 1$ and M is in the range described above are also of order 1. By reasoning that was used in the previous cases it is found that $\langle J | \Delta H_S | M \rangle$ is non-negligible only when f is in \mathcal{S} . Consequently, for certain sets of values for S there are of order 1 elements $\langle J | \Delta H_S | M \rangle$ of order 1 with a given J , and for other sets of values for S there are no elements of significant size.

By similar arguments it can be shown that for $[\Delta M] = 2$ there are, for a given J , of order 1 elements of order $N^{-\frac{1}{2}}$.

The next step in calculating the magnitude of $\langle H \rangle_{av}$ is to determine the size of the coefficients in the sum in Eq. (12b). For convenience the definition

$$A(L, J) \equiv \sum_M a(L, M) \langle J | \Delta H_S | M \rangle \quad (16)$$

can be made. The pattern of magnitudes of the elements can be used to find the orders of magnitudes of the coefficients $A(L, J)$. In making these calculations the contributions of the elements having different values of $[\Delta M]$ are considered separately.

The first elements treated are those where $[\Delta M] = 0$. They are of order 1 and there is one of them for each set of values for J . The coefficients $a(L, M)$ must be of order $\mathfrak{N}^{-\frac{1}{2}}$ to normalize $\Psi(L)$. For given sets of values

for L and J there is only one term in the sum in Eq. (16) which describes the contribution of the elements where $[\Delta M] = 0$. Since the magnitude of $a(L, M)$ is of order $\mathfrak{N}^{-\frac{1}{2}}$ and that of $\langle J | \Delta H_S | M \rangle$ is of order 1, the contribution to $A(L, J)$ is of order $\mathfrak{N}^{-\frac{1}{2}}$.

For certain sets of values of S , elements where $[\Delta M] = 1$ are of order 1 since the number and magnitudes of the elements are both of order 1. Consequently, the contribution of these elements to $A(L, J)$ is again of order $\mathfrak{N}^{-\frac{1}{2}}$.

Similarly for $[\Delta M] = 2$ the contribution to $A(L, J)$ is of order $N^{-\frac{1}{2}} \mathfrak{N}^{-\frac{1}{2}}$.

For the case where $[\Delta M] = 3$, the elements contribute of order N terms to the sum in Eq. (16) because the number of elements for a given J is of order N . The individual elements are of order $N^{-\frac{1}{2}}$ and so the maximum contribution to $A(L, J)$ could be of order $N^{\frac{1}{2}} \mathfrak{N}^{-\frac{1}{2}}$. This would be the case if each of the products $a(L, M) \langle J | \Delta H_S | M \rangle$ were real and had the same sign or more generally were complex and had the same phase angle.

If the terms have incoherent phase angles then the contribution is less by a factor of order $N^{-\frac{1}{2}}$ and, hence, is of order $\mathfrak{N}^{-\frac{1}{2}}$.

By the same reasoning it is found that for $[\Delta M] = n \geq 2$ the coherent contribution is of order $N^{\frac{1}{2}(n-1)} \mathfrak{N}^{-\frac{1}{2}}$ and the incoherent is of order $\mathfrak{N}^{-\frac{1}{2}}$.

Only a very small fraction of the coefficients $A(L, J)$ can receive coherent contributions. This statement can be demonstrated by considering the contribution of the elements where $[\Delta M] = 3$ as an example. This contribution can be designated as $A_3(L, J)$. The set of these elements for a given J form a vector, the square of the magnitude of which is

$$\sum_M \langle J | \Delta H_S | M \rangle \langle M | \Delta H_S | J \rangle = O(1). \quad (17)$$

It is of order 1 because the number of terms is of order N and the magnitude of each term is of order N^{-1} (i.e., the square of a term of order $N^{-\frac{1}{2}}$).

The elements of this vector can be expressed as linear combinations of the coefficients $a^*(L, M)$, that is,

$$\langle J | \Delta H_S | M \rangle = \sum_L A_3(L, M) a^*(L, J), \quad (18a)$$

and

$$\langle M | \Delta H_S | J \rangle = \sum_L A_3^*(L, M) a(L, J). \quad (18b)$$

By using these equations, the square of the magnitude of the vector can be expressed as

$$\sum_M \sum_{L'L'} A_3^*(L, M) a(L, J) A_3(L', M) a^*(L', J). \quad (19)$$

From this expression, Eq. (17), and the fact that the true wave functions $\Psi(L)$ are orthogonal, it follows that

$$\sum_M |\langle J | \Delta H_S | M \rangle|^2 = \sum_L A_3^*(L, J) A_3(L, J) = O(1). \quad (20)$$

Consequently, in the sum the order of magnitude of the number of terms containing coefficients $A_3(L, J)$ of order $N^{\frac{1}{2}} \mathfrak{N}^{-\frac{1}{2}}$ cannot be greater than $N^{-1} \mathfrak{N}$. Since the

number of possible sets of values of L that can contribute to $\langle J|\Delta H_S|M\rangle$ is of order \mathfrak{N} , only a fraction of order N^{-1} of the coefficients $A_3(L,J)$ can be of order $N^{1/2}\mathfrak{N}^{-1/2}$. This fraction is an upper limit since there does not need to be any coefficients of this size. By extension of this argument to elements of higher $[\Delta M]$ it is found that at most only a very small fraction of the coefficients $A(L,J)$ can be coherently derived and consequently the vast majority of them are of order \mathfrak{N}^{-1} . Presumably the coherently derived coefficients are negligible.

The fact that the coefficients $A(L,J)$ are of order $\mathfrak{N}^{-1/2}$ can be applied to determining the magnitude of $\langle H\rangle_{av}$. From Eqs. (12b) and (16) the average can be expressed as

$$\langle H\rangle_{av} = \int \Psi_b^*(L+S)H\Psi_b(L+S)d\tau \\ = E(L) + \sum_J a^*(L,J)A(L,J). \quad (21)$$

The magnitude of the sum in this equation is of order 1 because the number of terms is of order \mathfrak{N} and their magnitude is of order \mathfrak{N}^{-1} and at least part of the terms are coherent. Coherence must arise from the contribution of the terms where $[\Delta M]=0$. These terms contribute

$$\sum_M \langle M|\Delta H_S|M\rangle a^*(L,M)a(L,M), \quad (22)$$

which must be of order 1 because $a^*(L,M)a(L,M)$ must be positive and $\langle M|\Delta H_S|M\rangle$ must be either preponderantly positive or preponderantly negative, depending on S .

Therefore we can conclude that

$$\langle H-E(L)\rangle_{av} = O(1). \quad (23)$$

When H_a [as defined in Eq. (2)] is used instead of H , only the diagonal elements of $\langle M|\Delta H_S|J\rangle$ are changed and these by not more than a quantity of order 1. Consequently, the order of magnitude of $\langle H_a^2\rangle_{av}$ is the same as that of $\langle [H-E(L)]^2\rangle_{av}$.

From Eqs. (12b) and (16) and the fact that

$$\langle [H-E(L)]^2\rangle_{av} = \int \{[H-E(L)]\Psi_b(L+S)\}^* \\ \times [H-E(L)]\Psi_b(L+S)d\tau, \quad (24)$$

it follows that

$$\langle [H-E(L)]^2\rangle_{av} = \sum_J A^*(L,J)A(L,J). \quad (25)$$

Each of the coefficients $A(L,J)$ is of order $\mathfrak{N}^{-1/2}$ and the number of sets of values for J is of order \mathfrak{N} . Therefore, it follows that

$$\langle [H-E(L)]^2\rangle_{av} = O(1), \quad (26a)$$

and that

$$\langle H_a^2\rangle_{av} = O(1). \quad (26b)$$

From these results it is clear that the function $\Psi_b(L+S)$ is much more nearly an approximate eigenfunction than any of the product functions $\Phi(M)$.

In addition to information on orders of magnitude, it is possible to obtain further qualitative information on the energy distribution for a particular function $\Psi_b(L+S)$. By considering some of the elements $\langle J|\Delta H_S|M\rangle$ in more detail it is possible to estimate the size of $\langle H-E(L)\rangle_{av}$.

The largest of the elements is $\langle M|\Delta H_S|M\rangle$ because it is the only term of appreciable size that is derived from the quadratic terms, which are the dominant terms in the potential function. From the derivation in I of the magnitudes of terms of various orders, it can be seen that the contribution of a quadratic term to individual matrix elements is in general larger than the $N^{1/2}$ times the contribution of a cubic term⁵ and than $N^{1/2}$ times the contribution of a term where the sum of the exponents is $n+2$.

Because the contribution of the quadratic terms is greater than that of any other set of terms, the expression shown in (22) is approximately equal to $\langle H-E(L)\rangle_{av}$. It is also approximately equal to the difference between the energies of $\Phi(M+S)$ and $\Phi(M)$ when the anharmonic part of the potential function is left out.

When the operator H_a is used instead of H , as pointed out previously, only the diagonal elements are affected. Since the major contribution both to the diagonal elements and to $\langle H-E(L)\rangle_{av}$ comes from the quadratic terms, the diagonal elements for H_a are greatly reduced. As these were the most important elements, $\langle (H_a^2)_{av}\rangle^{1/2}$ should be considerably less than $\langle H-E(L)\rangle_{av}$. Consequently, the function $\Psi_b(L+S)$ has a range of energy that has a root-mean-square deviation from its average energy of $\langle (H_a^2)_{av}\rangle^{1/2}$ and has a mean energy which differs from that of $\Psi(L)$ by $\langle H-E(L)\rangle_{av}$.

The fact that $\langle H-E(L)\rangle_{av}$ is considerably larger than $\langle (H_a^2)_{av}\rangle^{1/2}$ means that the projection of $\Psi_b(L+S)$ on $\Psi(L)$ or on the set of all functions with energy close to $E(L)$ must be small. This is an important conclusion because it shows that $\Psi_b(L+S)$ does not approximate $\Psi(L)$ or any linear combination of eigenfunctions with energies close to $E(L)$.

There are two other things that should be noted about the results of this discussion. The first is that the criterion of (1) is less stringent than the one involving perturbation theory as given in Eq. (31) of I. A function made up of a linear combination of near-degenerate wave functions could satisfy the first but not the second criterion.

The second thing is that, included within the averages, are the anharmonic effects that are normally described as giving rise to overtones and combination bands. Consequently, it should be expected that the energy distribution for the function $\Psi_b(L+S)$ should have a complex structure with a number of peaks. These considerations suggest that it should be possible

⁵ This conjecture has been confirmed in another paper [J. Chem. Phys. **36**, 158 (1962)] by the writer by estimates made from experimental data for some molecular crystals.

to simplify the structure and reduce $\langle H_a^2 \rangle_{av}$ by a further modification of $\Psi_b(L+S)$.

III. A SECOND MODIFICATION OF AN EXACT EIGENFUNCTION

In the previous section an analog of an excited state was formed by modifying all of the component harmonic oscillator wave functions in the same way. In this section, the exact solutions to a very similar problem are used as analogs of excited states. The problem used is one where all the nonquadratic terms in the potential function which contain the coordinates $q(f)$, $q(f')$, etc. are deleted. The indices f , f' , etc. form a set \mathfrak{F} and the number of indices in the set is assumed to be of order 1. The Hamiltonian for this problem can be represented as H_c and the eigenfunctions and eigenvalues as $\Psi_c(L_c, s)$ and $E_c(L_c, s)$, respectively. The symbol s represents the quantum numbers for those indices, f , which are in \mathfrak{F} , and L_c is a set of quantum numbers that completes the description of the state.

Since $\Psi_c(L_c, s)$ is an eigenfunction of H_c , it follows that

$$H_c \Psi_c(L_c, s) = E_c(L_c, s) \Psi_c(L_c, s). \quad (27)$$

The function $\Psi_c(L_c, s)$ can be expressed as

$$\Psi_c(L_c, s) = \sum_{M_c} c(L_c, M_c) \Phi(M_c, s), \quad (28)$$

where M_c and s together comprise the set M ; and the true Hamiltonian can be expressed as

$$H = H_c + \Delta H_c. \quad (29)$$

Consequently,

$$H \Psi_c(L_c, s) = E_c(L_c, s) \Psi_c(L_c, s) + \sum_{M_c, J_c, t} c(L_c, M_c) \langle J_c, t | \Delta H_c | M_c, s \rangle \Phi(J_c, t), \quad (30)$$

where t represents a set of quantum numbers whose indices are in \mathfrak{F} . By arguments analogous to those used in the previous section, it can be shown that, for all but a negligibly small number of cases,

$$C(L_c, s; J_c, t) = O(\mathfrak{N}^{-\frac{1}{2}}), \quad (31)$$

where

$$C(L_c, s; J_c, t) = \sum_{M_c} c(L_c, M_c) \langle J_c, t | \Delta H_c | M_c, s \rangle. \quad (32)$$

Therefore, we find that for $\Psi_c(L_c, s)$

$$\langle H - E_c(L_c, s) \rangle_{av} = O(1), \quad (33a)$$

and

$$\langle H_a^2 \rangle_{av} = O(1). \quad (33b)$$

These results show that the functions $\Psi_c(L_c, s)$ as well as $\Psi_b(L+S)$ approximate the behavior of eigenfunctions much more closely than the functions $\Phi(M)$. There are some slight differences in the approximations by the two functions. The function $\Psi_b(L+S)$ is compared with the exact wave function whereas there is no direct comparison for the function $\Psi_c(L_c, s)$. Also $\langle H - E_c(L_c, s) \rangle_{av}$ should be much less than $\langle H - E(L) \rangle_{av}$

because $E_c(L_c, s)$ includes the effects of the large quadratic terms that were present in the elements $\langle M | \Delta H_s | M \rangle$.

IV. USE OF MODIFIED WAVE FUNCTIONS TO CALCULATE DIPOLE MOMENT MATRIX ELEMENTS

In the last two sections wave functions that could be considered analogs of excited states were considered. In this section, as an example, some of the same concepts will be applied to the absorption and emission of radiation.

The treatment is based on the behavior of the function $\mu \Psi(L)$ where μ is the dipole moment operator. This function is not in general an eigenfunction and, consequently, it has a distribution of energies. It is readily demonstrable that for energies greater than $E(L)$ the distribution function differs only by a constant factor from the induced absorption divided by the energy of transition as a function of energy. The corresponding function for induced emission is obtained for energies less than $E(L)$.

In principle the distribution function and hence the absorption and emission spectra could be obtained from all of the moments of the distribution but in order to achieve any detail in a complex spectrum a large number of moments would have to be found. In this discussion the problem will be restricted mainly to a single band. This restriction can be made if the operator μ is limited to that part of the contribution of a single linear term that increases one of the quantum numbers in a product function by one unit. The product wave functions are designated by $\Phi(M_c, \rho)$, where ρ is the quantum number of the coordinate present in μ [i.e., $q(f)$] and M_c stands for the rest of the quantum numbers.

From these definitions, it follows that

$$\mu \Phi(M_c, \rho) = \mu_\rho \Phi(M_c, \rho+1), \quad (34)$$

where μ_ρ is an abbreviation for $\langle \rho+1 | \mu | \rho \rangle$; and hence, that

$$\mu \Psi(L) = \sum_{M_c, \rho} a(L; M_c, \rho) \mu_\rho \Phi(M_c, \rho+1). \quad (35)$$

The function $\mu \Psi(L)$ can be normalized and examined according to the criterion of (1) to determine whether it behaves as an approximate eigenfunction. Thus, the normalized function

$$\Theta(L) \equiv \mu \Psi(L) \left[\sum_{M_c, \rho} |\mu_\rho|^2 |a(L; M_c, \rho)|^2 \right]^{-\frac{1}{2}} \quad (36)$$

can be defined.

The first step in setting up the criterion is to write the equation

$$H \Theta(L) = \sum_{M_c, \rho} \sum_{J_c, \rho'} a(L; M_c, \rho) \mu'_{\rho} \times \langle J_c, \rho'+1 | H | M_c, \rho+1 \rangle \Phi(J_c, \rho'+1), \quad (37)$$

where μ_{ρ}' is the normalized value of μ_{ρ} , i.e.,

$$\mu_{\rho}' = \mu_{\rho} \left(\sum_{M_{c\rho}} |\mu_{\rho}|^2 |a(L; M_{c\rho})|^2 \right)^{-\frac{1}{2}}. \quad (38)$$

To determine the orders of magnitude of the various terms in Eq. (37) it is convenient to make the definition

$$\Delta\mu'_{\rho'\rho} \equiv \mu_{\rho}' - \mu_{\rho'}'. \quad (39)$$

With this expression, Eq. (37) becomes

$$\begin{aligned} H\Theta(L) = & \sum_{M_{c\rho}} \sum_{J_{c\rho'}} a(L; M_{c\rho}) \\ & \times \langle J_{c\rho'} + 1 | H | M_{c\rho} \rangle \mu_{\rho'}' \Phi(J_{c\rho'} + 1) \\ & + \sum_{M_{c\rho}} \sum_{J_{c\rho'}} a(L; M_{c\rho}) \langle J_{c\rho'} + 1 | H | M_{c\rho} \rangle \\ & \times \Delta\mu'_{\rho'\rho} \Phi(J_{c\rho'} + 1). \end{aligned} \quad (40)$$

From Eq. (11), $\langle J_{c\rho'} + 1 | H | M_{c\rho} \rangle$ becomes

$$\begin{aligned} \langle J_{c\rho'} + 1 | H | M_{c\rho} \rangle = & \langle J_{c\rho'} | H | M_{c\rho} \rangle \\ & + \langle J_{c\rho'} | \Delta H_S | M_{c\rho} \rangle. \end{aligned} \quad (41)$$

Here the elements of S are all zero except for the one corresponding to ρ , which is 1. From this equation and Eqs. (12b) and (16) the first sum in Eq. (40) can be expressed as

$$E(L)\Theta(L) + \sum_{J_{c\rho'}} \mu_{\rho'}' A(L; J_{c\rho'}) \Phi(J_{c\rho'} + 1). \quad (42)$$

Since the magnitudes of μ_{ρ}' and $A(L; J_{c\rho'})$ are respectively of orders 1 and $\mathfrak{N}^{-\frac{1}{2}}$, the coefficients $\mu_{\rho'}' A(L; J_{c\rho'})$ are also of order $\mathfrak{N}^{-\frac{1}{2}}$.

The magnitudes of the contributions of the coefficients $\Delta\mu'_{\rho'\rho} \langle J_{c\rho'} + 1 | H | M_{c\rho} \rangle$ can be obtained by the method used in the last two sections. Since $\Delta\mu'_{\rho'\rho} = 0$ when $\rho' = \rho$, the elements $\langle J_{c\rho'} + 1 | H | M_{c\rho} \rangle$ make no contribution to the second sum in Eq. (40). When $\rho' \neq \rho$, the pattern of orders of magnitudes for $\Delta\mu'_{\rho'\rho} \langle J_{c\rho'} + 1 | H | M_{c\rho} \rangle$ is the same as that found for $\langle J | \Delta H_S | M \rangle$ or $\langle J_{c\rho'} | \Delta H_c | M_{c\rho} \rangle$. Thus it is clear that

$$\begin{aligned} B(L; J_{c\rho'}) = & \sum_{M_{c\rho}} \Delta\mu'_{\rho'\rho} \langle J_{c\rho'} + 1 | H | M_{c\rho} \rangle \\ & \times a(L; M_{c\rho}) = O(\mathfrak{N}^{-\frac{1}{2}}), \end{aligned} \quad (43)$$

and also that

$$\begin{aligned} D(L; J_{c\rho'}) = & \mu_{\rho'}' A(L; J_{c\rho'}) \\ & + B(L; J_{c\rho'}) = O(\mathfrak{N}^{-\frac{1}{2}}). \end{aligned} \quad (44)$$

Therefore, the quantities $\langle H - E(L) \rangle_{\text{av}}$ and $\langle H_a^2 \rangle_{\text{av}}$ are both of order 1 and, hence, the behavior of $\Theta(L)$ approximates that of an exact eigenfunction.

The value of $\langle \langle H_a^2 \rangle_{\text{av}} \rangle^{\frac{1}{2}}$ should be small compared to that of $\langle H - E(L) \rangle_{\text{av}}$ for the same reason that it is small for $\Psi_b(L+S)$. The value of the criterion imposes some restriction on the shape of the induced absorption band governed by the operator μ . Its moderately small size with respect to $\langle H - E(L) \rangle_{\text{av}}$ limits the possible

structures to the following: a single broad band, a narrow band with a weak distant satellite, a set of closely spaced bands, a narrow band with an extensive foot of low intensity, or some mixture of these. In all of the cases the principal part of the spectrum consists of a single band (under moderate resolution) close to $\langle H - E(L) \rangle_{\text{av}}$.

Integration of the intensity divided by the frequency of transition yields

$$(2\pi/3\hbar^2) \sum_{M_{c\rho}} |\mu_{\rho}|^2 |a(L; M_{c\rho})|^2. \quad (45)$$

This quantity can be considered to be the average value of $(2\pi/3\hbar^2) |\mu_{\rho}|^2$ over all of the component functions $\Phi(M_{c\rho})$. Since the quantities $|a(L; M_{c\rho})|^2$ are roughly the same size as are $|\mu_{\rho}|^2$, this integrated intensity is at worst a moderately weighted average of the quantities $(2\pi/3\hbar^2) |\mu_{\rho}|^2$ for each of the functions $\Phi(M_{c\rho})$ having the proper symmetry and lying in the energy range. Thus the integrated intensity is about the same as that for essentially any of the functions $\Phi(M_{c\rho})$ in the energy range.

In addition the difference between the energies of $\Phi(M_{c\rho})$ and $\Phi(M_{c\rho} + 1)$ in the harmonic approximation is very nearly equal to $\langle H - E(L) \rangle_{\text{av}}$, the average transition energy. This result follows from the discussion in Sec. II of the contribution of quadratic terms to $\langle M | \Delta H_S | M \rangle$.

From this analysis it is apparent that absorption intensities and transition energies calculated for single product functions $\Phi(M)$ or $\Phi(M_{c\rho})$ are approximately correct despite the fact that these functions are very poor approximations to the exact eigenfunctions.

The treatment described above can readily be extended to include the complete dipole moment expansion. Since the distance between bands is usually large compared with their widths as described by $\langle \langle H_a^2 \rangle_{\text{av}} \rangle^{\frac{1}{2}}$, it is a good approximation to consider the spectrum as being the sum of the contributions of the individual bands.

The functions used in Sec. III can also be used in the calculation of the dipole moment matrix elements. Again the same sort of conclusions can be reached. For example, the operator μ gives

$$\mu \Psi_c(L_{c\rho}) = \sum_{M_c} c(L_{c\rho}, M_c) \mu_{\rho} \Phi(M_{c\rho} + 1), \quad (46)$$

where the single quantum number ρ constitutes the set s . The transition moment for the two states is then μ_{ρ} .

The formal simplicity of this result is offset by the fact that there is a range of energies associated with each of the functions $\Psi_c(L_{c\rho} + 1)$ and $\Psi_c(L_{c\rho})$, and, hence, the energy distribution for the transition cannot readily be determined.

In both of the treatments of this section the basic technique can be considered to be the isolation of the behavior of a single mode of vibration undergoing a

transition. The transition can be thought of as taking place in a vast number of different vibrational environments described by the sets of quantum numbers M_c . The fact that it is not possible to isolate the behavior of a single mode completely is illustrated by the distributions of energies for at least some of the functions involved in either of the treatments.

V. CALCULATION OF HAMILTONIAN MATRIX ELEMENTS

The partial success of the treatment of dipole-moment matrix elements suggests that it should be possible to treat other elements in the same way. Since perturbation theory has been used successfully in the past to treat anharmonicity of crystals, particularly in describing overtone and combination bands, it should be possible to calculate appropriate matrix elements for the anharmonic part of the Hamiltonian operator H .

If true wave functions were used as basis functions there would be no off-diagonal elements for H , and, hence, there would be no elements that are analogous to those that symbolize the effects of anharmonicity in simple systems. However, such off-diagonal elements do appear when wave functions which are nearly complete solutions are used as basis functions. Functions of this sort are

$$\Psi_c(L_c, \rho) = \sum_{M_c} c(L_c, M_c) \Phi(M_c, \rho), \quad (47)$$

and

$$\Psi_{bc}(L_c + S_c, \rho') = \sum_{M_c} c(L_c, M_c) \Phi(M_c + S_c, \rho'). \quad (48)$$

The symbols ρ and ρ' represent values of a single quantum number, which are assumed to be different in most of the following discussion.

It should be noted that the function $\Psi_c(L_c, \rho)$ is the same as that used in Eq. (46) and is a special case of the function $\Psi_c(L_c, s)$ of Sec. III. The function $\Psi_{bc}(L_c + S_c, \rho')$ is a composite of the types of functions used in Secs. II and III.

The Hamiltonian matrix elements considered are

$$\begin{aligned} \langle L_c + S_c, \rho' | H | L_c, \rho \rangle \\ = \int \Psi_{bc}^*(L_c + S_c, \rho') H \Psi_c(L_c, \rho) d\tau. \end{aligned} \quad (49)$$

From Eq. (30) and the fact that $\Psi_{bc}(L_c + S_c, \rho')$ and $\Psi_c(L_c, \rho)$ are orthogonal when $\rho \neq \rho'$, it follows that

$$\begin{aligned} \langle L_c + S_c, \rho' | H | L_c, \rho \rangle &= \sum_{M_c, J_c} c(L_c, M_c) \\ &\times \langle J_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle c^*(L_c, J_c). \end{aligned} \quad (50)$$

The problem that must be solved to give useful results is whether a more tractable relation between the elements $\langle L_c + S_c, \rho' | H | L_c, \rho \rangle$ and $\langle J_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle$

can be found. As it stands no further simplification is possible since it is not practicable to express the coefficients $c(L_c, M_c)$ explicitly.

It is helpful in this analysis to base further treatment on some idealizations. The first of these is that for given sets of values for ΔM_c and S_c the element $\langle J_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle$ is independent of M_c . The second is that

$$\sum_{M_c} c^*(L_c, M_c + \Delta M_c) c(L_c, M_c) = 0, \quad (51)$$

when $[\Delta M_c] \neq 0$.

Under these idealized conditions we would find that

$$\langle L_c + S_c, \rho' | H | L_c, \rho \rangle = \langle M_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle. \quad (52)$$

If a further idealization is made that similar relationships hold for $\langle L_c + S_c', \rho'' | H | L_c + S_c, \rho' \rangle$, then the Hamiltonian matrix with $\Psi_c(L_c, \rho)$ and $\Psi_{bc}(L_c + S_c, \rho')$ as basis functions would be a matrix with elements $\langle M_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle$. Consequently, the true wave functions could be found by obtaining the eigenvectors of this matrix.

From Eq. (33b) it is apparent that

$$\sum_{s_c} |\langle M_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle|^2 = O(1), \quad (53)$$

and hence, that it is not unreasonable to expect that the methods of perturbation theory would be valid in obtaining the solutions to this idealized problem.

In the actual problem the idealizations are not valid. That the first idealization is not valid can be seen by examining the dependence of the matrix elements $\langle M_c + \Delta M_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle$ on the elements of M_c , ΔM_c , and S_c . The symbols ΔM_c and S_c represent sets of differences of quantum numbers all but a number of order 1 of which are zero. From the results of I and the discussion in Secs. II and III it can be shown that these matrix elements depend to an appreciable extent only on the quantum numbers in M_c that correspond to the nonzero elements of $\Delta M_c + S_c$.

To illustrate this statement, let us consider a case where $[\Delta M_c + S_c] = 2$ and $\Delta \rho = 1$. The principal term of ΔH_c which contributes to the element has the form shown in (13a). Here, $q(f_1)$ is the coordinate for which ρ is the quantum number and f_2 and f_3 are the indices of the two nonzero differences in quantum numbers in $\Delta M_c + S_c$. It is clear that the contribution of the cubic term to the matrix element depends on the quantum numbers for $q(f_1)$, $q(f_2)$, and $q(f_3)$, two of which [those for $q(f_2)$ and $q(f_3)$] are in the set M_c . For the same reason the rest of the contributing terms [of the form shown in (13b)] give rise to the same sort of dependence on M_c . From the generalization of this illustration it is readily seen that all of the matrix elements $\langle M_c + \Delta M_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle$ depend on M_c even when $\Delta M_c + S_c$ is fixed.

So far it has been demonstrated that the elements depend on M_c . The next thing that needs to be shown is that a significant fraction of the elements with a given $\Delta M_c + S_c$ differ in value from the rest of these elements.

This is done best by returning to the illustration with the cubic terms. It is clear that the contribution of one of these to a matrix element where $[\Delta M_c + S_c] = 2$ is the same for all sets of quantum numbers M_c which have the same pair of values for the quantum numbers for $q(f_2)$ and $q(f_3)$. To be consistent with the fact that the number of different sets of quantum numbers M_c is roughly of order 2^N , the fraction of these sets which have pairs of quantum numbers for $q(f_2)$ and $q(f_3)$ with values different from those originally cited is of order 1 (as opposed to N^{-1}).

By an extension of this argument, it can be shown that for any set of matrix elements $\langle M_c + \Delta M_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle$ for a particular set of values of the numbers in $\Delta M_c + S_c$, the fraction of elements that differ appreciably in value from the value of any one element is of order 1.

These conclusions can be utilized by rewriting Eq. (50) as

$$\begin{aligned} \langle L_c + S_c, \rho' | H | L_c, \rho \rangle &= \sum_{M_c} |c(L_c, M_c)|^2 \\ &\times \langle M_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle \\ &+ \sum_{M_c} \sum_{\Delta M_c'} c^*(L_c, M_c + \Delta M_c) c(L_c, M_c) \\ &\times \langle M_c + \Delta M_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle, \end{aligned} \quad (54)$$

where $\sum_{\Delta M_c'}$ means the sum over all sets of values for M_c except the one where $[\Delta M_c] = 0$. The first term on the right of this equation is the average value of $\langle M_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle$. It can also be considered as the projection of a vector \mathfrak{C} , whose elements are

$$\mathfrak{C}(L_c, \rho; M_c, \rho') \equiv c(L_c, M_c) \langle M_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle, \quad (55)$$

on the normalized vector \mathbf{c} whose elements are $c(L_c, M_c)$. The square of the magnitude of this projection, $|P|^2$, on the rest of the space is given by

$$|P|^2 = \sum_{M_c} |\mathfrak{C}(L_c, \rho; M_c, \rho')|^2 - \left| \sum_{M_c} c(L_c, M_c) \right|^2 \times \langle M_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle^2, \quad (56)$$

or

$$\begin{aligned} |P|^2 &= \sum_{M_c} \left(\langle M_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle \right. \\ &\quad \left. - |\langle M_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle_{\text{av}}| \right)^2 \\ &\quad \times |c(L_c, M_c)|^2. \end{aligned} \quad (57)$$

Since the fraction of the matrix elements which differ appreciably in magnitude from any one element $\langle M_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle$ is of order 1, then

$$\begin{aligned} &|\langle M_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle \\ &\quad - |\langle M_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle_{\text{av}}| \end{aligned} \quad (58)$$

must have a magnitude of the same order as $\langle M_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle$ and, therefore, $|P|$ must also have a magnitude of this order. The effect of this projection appears as a contribution to the second term on the right-hand side of Eq. (54).

The result of the fact that Eq. (51) is true only for the idealized case also causes a contribution to the second term on the right-hand side of Eq. (54). This

contribution should, however, be small as can be seen from the following discussion. From Eqs. (48a) and (48b) it is apparent that

$$\begin{aligned} &\sum_{M_c} c^*(L_c, M_c + \Delta M_c) c(L_c, M_c) \\ &= \int \Psi_{bc}^*(L_c + M_c, \rho) \Psi_c(L_c, \rho) d\tau. \end{aligned} \quad (59)$$

The integral can be shown to be small by arguments identical with those used in the demonstration in Sec. II that $\Psi_b(L+S)$ and $\Psi(L)$ are nearly orthogonal. Consequently, it can be presumed that the dominant contribution to the second term on the right of Eq. (54) comes from $|P|$.

Despite the fact that the idealizations are not correct, they may be used as the basis for approximations. A rational method of approximation would be to use the quantities

$$\langle \langle M_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle \rangle_{\text{av}} \quad \text{for} \quad \langle L_c + S_c, \rho' | H | L_c, \rho \rangle$$

particularly in cases where the value of most of the elements $\langle M_c + S_c, \rho' | \Delta H_c | M_c, \rho \rangle$ are close to the average. However, in every case there are still contributions that come from the difference in direction of \mathfrak{C} and \mathbf{c} . These quantities depend on the pattern of coefficients $c(L_c, M_c)$, which is extremely complex and essentially indeterminate.

From this discussion it is apparent that the use of single-product wave functions $\Phi(M)$ [or $\Phi(M_c, \rho)$] to calculate matrix elements for a part ΔH_c of the Hamiltonian operator is a useful procedure. The elements so calculated are the same as those calculated in the idealized case described above. Thus, it is seen that the mixing of the functions $\Phi(M_c, \rho)$ and $\Phi(M_c + S_c, \rho')$ found when the perturbing operator is added to the unperturbed Hamiltonian operator containing only quadratic terms is the same as the mixing of $\Psi_c(L_c, \rho)$ and $\Psi_{bc}(L_c + S_c, \rho')$ under the action of H in the idealized case and is approximately the same in the true case. This situation can be described by saying that the behavior of the functions $\Psi_c(L_c, \rho)$ and $\Psi_{bc}(L_c + S_c, \rho')$ is approximately represented by the behavior of the product functions. For example, the intensity of an infrared binary combination band caused by mechanical anharmonicity could be determined approximately by treating the upper and lower states as if they were the product functions $\Phi(M_c + S_c, \rho)$ and $\Phi(M_c, \rho)$ slightly altered by small perturbations.

From this analysis it is apparent why it is possible for perturbation calculations for anharmonic crystals using product wave functions to give approximately correct results at least when the treatment is restricted to a particular set of product functions, and also why perturbation calculations on anharmonic crystals have given satisfactory results in the past.

It should be noted, however, that the representation of the behavior of the functions $\Psi_c(L_c, \rho)$ and

$\Psi_{bc}(L_c+S_c, \rho')$ is not exact. The inexactness arises from the essentially indeterminate matrix elements that are caused by the variation in the elements $\langle M_c+\Delta M_c+S_c, \rho' | \Delta H_c | M_c, \rho \rangle$ with M_c for a given ΔM_c and S_c . Since the difficulty in approximation lies in the determination of the matrix elements themselves and not in the handling of the matrix, it is apparent that the use of higher order perturbation theory will still not lead to a correct solution. Thus, we see that while perturbation theory using product wave functions can give approximate results there is an intrinsic limit to its accuracy even when higher order perturbation theory is used.

At this point it is helpful to describe in a less formal way the basis of the work of this section. In expressing the functions $\Psi_c(L_c, \rho)$ and $\Psi_{bc}(L_c+S_c, \rho')$, attention is centered on the behavior of the phonons for the mode corresponding to the quantum number ρ . The mixing of these functions by the perturbing operator can be described in terms of the annihilation of sets of phonons that necessarily include one or more phonons corresponding to ρ .

When perturbation theory can be applied, the more complete description of a phonon corresponding to ρ includes small contributions of virtual states involving the annihilation of this phonon and the creation of others. The phonons for ρ exist in a vibrational environment described by $c(L_c, M_c)$ and this environment affects the size of the contributions of the virtual states.

The averaged matrix elements

$$\langle \langle M_c+S_c, \rho' | \Delta H_c | M_c, \rho \rangle \rangle_{av}$$

can be used in an attempt to cope with the effect of the vibrational environment, but it is not completely successful because of the specific nature of the interaction. In any attempt to treat the specific interaction further, the centering of attention on a particular phonon or mode must be abandoned.

In this section as in the last the basic method is the attempted isolation of a part of the motion of the crystal. In both sections the method is partially successful.

VI. CONCLUSIONS

From the work of Sec. I it was concluded that with possible rare exceptions when any product of harmonic oscillator wave functions is used as an approximation to the wave function of an anharmonic crystal, then the energy spectrum covers a range which is very large compared to ordinarily observed energy changes in crystals, e.g., infrared transitions. It is also found (again with possible rare exceptions) that the product

functions are very poor approximations to the true eigenfunctions in the sense that the projection of a product on any eigenfunction is extremely small. This conclusion is subject to the limitation that the number of unit cells in the crystal be sufficiently large.

In Secs. II and III it was found that it is possible to express formally functions for which the energy range is of the same order of magnitude as ordinarily observed energy changes. Two slightly different types of functions were investigated. They were both derived from the true eigenfunctions. It was found that both of the functions had at most small projections on the eigenfunctions from which they were derived.

In Sec. IV it was found that, with rare exceptions, the square of the dipole moment matrix element for a single product function is a reasonable approximation to that for a true wave function, despite the fact that product functions are very poor approximations in the sense described above.

In Sec. V the two types of functions derived in Secs. II and III were combined and used to construct matrix elements analogous to those for simple anharmonic molecular problems. It was found that these elements are approximately equal to corresponding elements calculated for single product functions. It was also concluded that perturbation theory calculations using matrix elements derived from single-product wave functions can give approximately correct results in calculating such quantities as absorption intensities and frequencies.

The basic conclusion that can be drawn from the work of this paper is that it is valid to use wave functions which are products of harmonic oscillator functions to approximate various vibrational properties of anharmonic crystals such as intensities and frequencies of fundamental absorption bands. It is also shown that there is an intrinsic limit or a residual uncertainty in such an approximation. This uncertainty appears in various ways: as ranges in energy of wave functions, as ranges in absorption frequency, and in the presence of essentially indeterminate matrix elements.

In effect all of the approximations used involve attempts to isolate the behavior of particular modes or phonons. The limitation of the approximations lies in the presence of an underlying specific and collective interaction of all the modes and phonons making complete isolation impossible.

The problem treated is in many ways similar to the problem of many-body interaction for other particles such as photons, electrons, or nucleons. Consequently, similar conclusions may apply to the treatment of these particles.