

## Effect of Invariance Requirements on the Elastic Strain Energy of Crystals with Application to the Diamond Structure\*

P. N. KEATING†

*Tyco Laboratories, Incorporated, Waltham, Massachusetts*

(Received 17 December 1965)

After demonstrating inconsistencies in some of the better known elasticity calculations, an alternative method of imposing the necessary invariance conditions on the strain energy of a crystal is presented. The new method is equivalent to the Born-Huang procedure but, in addition to providing further insight, also offers one or two operational advantages. For example, it demonstrates that all purely first-neighbor interactions are central only. The method is applied to the calculation of the elasticity of a two-constant model of the diamond type of crystal, and this predicts the relation  $2c_{44}(c_{11}+c_{12})=(c_{11}-c_{12})(c_{11}+3c_{12})$ , which is very well satisfied by the experimental data for diamond, silicon, and germanium.

### I. INTRODUCTION

THE calculation of the elasticity, cohesive energy, and lattice dynamics of crystals by means of microscopic models has always formed an important part of solid-state physics. Since the earliest work by Born and co-workers,<sup>1,2</sup> there have, in fact, been so many such calculations that no attempt will be made here to give even a sketchy account of them; only specifically relevant work will be referred to in the course of the present contribution. The purpose of this contribution is to show several inconsistencies which have arisen in some of the better known calculations because of the use of unsatisfactory force-constant models, to present a method whereby a physically realistic set of force constants may be chosen, and to describe calculations of the elasticity of the diamond type of crystal based on a corrected force-constant scheme and using a relatively new and more convenient calculational method.

The conditions which are imposed by rotational invariance on the coefficients of an expansion of the potential energy in terms of the atomic displacement components were derived by Born and Huang,<sup>3</sup> and many of the more recent calculations satisfy this requirement. However, many calculations have been based on a force-constant set which has not been explicitly subjected to these conditions and the use of nearest-neighbor central and noncentral force constants has been fairly widespread.<sup>1,2,4</sup> On the other hand, the existence of noncentral first-neighbor interactions has recently been questioned by Lax,<sup>5</sup> who treated a one-

dimensional finite lattice and showed that, in this case, the first-neighbor noncentral constant vanishes. The extension of these arguments to finite three-dimensional structures is complex, and it is not certain that the Lax result can be extended to this general case; however, the implications in his result are obvious.

In the next section of this contribution, we show how some of the classic calculations of Born give unreasonable results, and thereby add emphasis to Lax's point<sup>5,6</sup> that rotational invariance must always be imposed. In the third section, we present an alternative method of imposing the necessary invariance conditions on the elastic strain energy which is essentially equivalent to the Born-Huang method<sup>3</sup> but offers several operational advantages, as well as providing additional insight. One such advantage allows us to show quite generally that the purely first-neighbor noncentral constants must always vanish. In the fourth section, we apply these formal results to the calculation of the elastic constants of the diamond type of crystal, using a method which is an extension of a method first used by Kittel<sup>7</sup> and which is simpler to apply than the Born long-waves method.<sup>3</sup>

### II. FAILURE OF THE EARLY BORN CALCULATIONS

The classic calculation of the elasticity and lattice dynamics of a crystal is that published by Born and von Kármán<sup>1</sup> for the simple cubic case and summarized by Kittel.<sup>7</sup> Two nearest-neighbor force constants  $\alpha'$  and  $\beta'$  (central and noncentral, respectively) and several second-neighbor constants are employed. The calculational method used is an early version of the long-waves method whereby the force on an atom is calculated in terms of the force constants and atomic displacements. The equations of motion of the atom in terms of the components of the strain gradients are then compared with the macroscopic equations of motion, and relations between the force constants and

\* Work supported in part by the Energetics Branch of the U. S. Air Force Cambridge Research Laboratories under contract AF19(628)2845.

† Present address: Bendix Research Laboratories, Southfield, Michigan.

<sup>1</sup> M. Born and T. von Kármán, *Physik. Z.* **13**, 297 (1912).

<sup>2</sup> M. Born, *Ann. Physik* **44**, 605 (1914).

<sup>3</sup> M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford University Press, New York, 1954), Chap. V.

<sup>4</sup> See also, for example, J. De Launey, *Solid State Phys.* **2**, 220 (1956); A. Segmüller, *Physik Kondensierten Materie* **3**, 18 (1964); J. Crobeau, *J. Phys. Radium* **25**, 925 (1964).

<sup>5</sup> M. Lax, in *Lattice Dynamics*, edited by R. F. Wallis (Pergamon Press Ltd., New York, 1965), p. 583.

<sup>6</sup> Reference 5, p. 179.

<sup>7</sup> C. Kittel, *Introduction to Solid State Physics* (John Wiley & Sons, New York, 1956).

the elastic coefficients are thereby obtained. Born and von Kármán obtained<sup>1,7</sup>

$$c_{11}=\alpha'/a; \quad c_{12}=-\beta'/a; \quad c_{44}=\beta'/a$$

(where  $a$  is the lattice constant) in the limit of negligible second-neighbor interactions. Not only are these expressions inconsistent with the results of Kittel's calculations [Eq. (4.34) of Ref. 7], but they are also clearly unsatisfactory physically. For example, the bulk modulus is  $B=(c_{11}+2c_{12})/3=(\alpha'-2\beta')/3a$ , using the Born-von Kármán results. However, this is the modulus which determines volume changes under hydrostatic pressure and should be independent of  $\beta'$ , the first-neighbor noncentral constant, since only the interatomic distances change under this type of deformation and there are no angle changes. Furthermore, the negative value of  $c_{12}$  is clearly unsatisfactory for this model.

Another classic calculation of elasticity and lattice dynamics was also published by Born,<sup>2</sup> for the diamond type of crystal, and it was this work which caused the author to initiate the present study. During the course of a calculation<sup>8</sup> of the second-order electric moment of this type of crystal, it became necessary to estimate the relative magnitude of some of the atomic force constants via elastic-constant data. Important differences were noted between the results obtained by a static calculation and those obtained by Born.<sup>2</sup> Using the long-waves method, Born obtained<sup>2,4</sup>

$$\begin{aligned} c_{11} &= (\alpha'' + \beta'')/4a, \\ c_{12} &= (\alpha'' - 2\beta'')/4a, \\ c_{44} &= \frac{3\beta''(\alpha'' + \beta'')/4}{4a(\alpha'' + \beta'')}, \end{aligned}$$

where  $4a$  is the lattice constant,  $\alpha''$  is the central first-neighbor constant (proportional to Born's  $A'$ ), and  $\beta''$  is the noncentral first-neighbor constant (proportional to  $-B'$ ). These results are again unsatisfactory physically. The bulk modulus again depends on  $\beta''$ , whereas there are no angle changes under hydrostatic pressure. Furthermore, the compliance constant  $s_{11}$  calculated from these expressions has a very unphysical singularity at  $\beta'' = \alpha''$  and is *negative* in the range  $\alpha'' < \beta'' < 2\alpha''$ .

These difficulties led us at first to suspect that they were due to inadequacy in the Voigt elasticity formalism of the type suggested by Laval.<sup>9</sup> However, arguments put forward by Lax<sup>5,10</sup> show that the Voigt formalism is perfectly adequate for the purposes intended. This fact suggests that it may be the force-constant models which are invalid, since the Born long-waves method is satisfactory. This reasoning, together with the doubts

about the force constants already raised by Lax,<sup>5</sup> shows that a more careful study of the types of force constant which are acceptable is necessary. This is the subject of the next section.

### III. SUITABLY INVARIANT FORM FOR THE STRAIN ENERGY

The elastic strain energy of a crystal is subjected to various physical requirements and these may be divided into two classes: the general conditions, such as rotational and displacement invariance, and those imposed by the symmetry of the crystal structure. It is the primary purpose of this section to consider the effect of the general conditions on the form of the strain energy  $V$  of a large crystal and thereby on the types of force constant which are acceptable. We shall consider any general type of deformation and will assume that the elastic strain energy depends only on the positions of the nuclei.<sup>11</sup> However, the requirement that the energy be invariant under an arbitrary displacement of the lattice as a whole ensures that  $V$  can depend only on the *differences* between nuclear positions, i.e.,

$$V = V(\mathbf{x}_k - \mathbf{x}_l) = V(\mathbf{x}_{kl}),$$

where  $\mathbf{x}_{kl} = \mathbf{x}_k - \mathbf{x}_l$  and where  $\mathbf{x}_k$  is the position vector of the  $k$ th nucleus after deformation. But  $V$  must be invariant under a transformation in which the atoms are further displaced in such a way that a rigid rotation of the crystal is produced. The  $\mathbf{x}_{kl}$  are not invariant under such a transformation; they transform as vectors. The only invariants which can be formed from the  $\mathbf{x}_{kl}$  are the scalar products between them and functions of such products. Hence we define

$$\lambda_{klmn} = (\mathbf{x}_{kl} \cdot \mathbf{x}_{mn} - \mathbf{X}_{kl} \cdot \mathbf{X}_{mn})/2a, \quad (1)$$

where  $a$  is a lattice constant,  $\mathbf{X}_{kl}$  is  $\mathbf{X}_k - \mathbf{X}_l$ , and  $\mathbf{X}_k$  is the position vector of the  $k$ th nucleus in the undeformed crystal. The final term is included so that  $\lambda_{klmn}$  vanishes when the deformation is removed; this does not affect the invariance properties of  $\lambda$ , since the material vector  $\mathbf{X}_k$  does not change under these transformations. Thus the strain energy is a function  $V(\lambda_{klmn})$  of the large number of  $\lambda_{klmn}$ . The similarity between the above discussion and the arguments presented by Lax<sup>5</sup> to establish a form for the *macroscopic* energy is not deceptive. The  $\lambda_{klmn}$  are the microscopic analogs of the  $E_{LM}$  of Sec. 2 of Ref. 5.

Since the  $\lambda_{klmn}$  are small, we assume that they can be used as a basis for a series expansion of  $V$ . The constant term in such an expansion is clearly unimportant, and

<sup>8</sup> P. N. Keating, Phys. Rev. **140**, A369 (1965); P. N. Keating and G. Rupprecht, *ibid.* **138**, A866 (1965).

<sup>9</sup> J. Laval, Compt. Rend. **232**, 1947 (1951); J. Phys. Radium **18**, 289, 369 (1957).

<sup>10</sup> M. Lax (private communication).

<sup>11</sup> This is valid in nonmetallic crystals even when the electrons are treated in detail, since the Born-Oppenheimer approximation ensures that the electrons follow the nuclei in full. Thus a shell-model strain energy must have the form of Eqs. (3) and (4) when expressed in terms of only the nuclear positions after imposing the condition (see Refs. 8, 19, for example) that the forces on the electrons are always negligible.

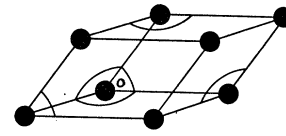
the linear terms must vanish so that the potential energy is an extremum at equilibrium. Because the change in potential energy under any infinitesimal uniform strain from equilibrium must be zero, the terms which are linear in the components of the displacements  $\mathbf{u}_{k\ell} = \mathbf{x}_{k\ell} - \mathbf{X}_{k\ell}$  must vanish. This, in turn, causes the terms which are linear in the  $\lambda_{klmn}$  to vanish and hence, using the usual summation convention,

$$V = \frac{1}{2} B_{klmn} \rho_{qrst} \lambda_{klmn} \lambda_{qrst} + O(\lambda^3) \tag{2}$$

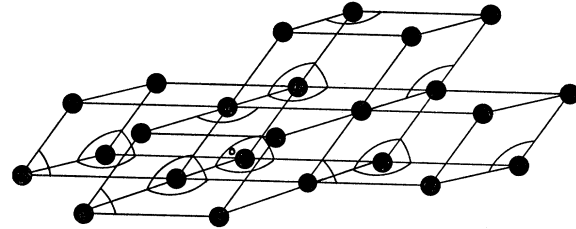
for small strains. It is interesting to note that this condition is more stringent than the familiar requirement that the forces on every atom vanish at equilibrium, since the latter is frequently satisfied for non-vanishing linear terms merely by the cancellation of contributions from different atoms due to the coordination symmetry which often occurs in crystals. The  $\{B_{klmn} \rho_{qrst}\}$  must be positive definite to ensure that  $V$  is not only an extremum at equilibrium but also a definite minimum.

However, the expansion (2) is not yet the required expression for the strain energy since there is a large amount of redundancy in the  $\lambda_{klmn}$ , and this is undesirable since it means that most of the coefficients  $B$  are then not independent. This redundancy arises from two sources. Firstly, Lax has pointed out<sup>10</sup> that only three lengths are necessary to locate an extra point with respect to an assembly of other points, provided we are considering only small displacements and the equilibrium arrangement of all the points is known. In fact, only  $3N - 6$  invariants are necessary to specify an arrangement of  $N$  points in three-dimensional space, and this is considerably smaller than the number of  $\lambda_{klmn}$  which can be defined from Eq. (1). Whilst the invariance properties of the equilibrium crystal under the operations of the space group of the crystal also causes redundancy, the main source of difficulty lies in the choice of only  $3N - 6$  of the  $1/8N(N + 1) \times [N(N + 1) + 2]$   $\lambda$ 's defined by Eq. (1). This problem is the subject of the following paragraphs.

Initially, we consider the description of a slightly deformed primitive structure. All primitive structures can be considered as consisting of a large number of parallelepipeds with atoms at each of the eight corners. In the undistorted case, all parallelepipeds in the bulk of the crystal are identical and each is a unit cell, of course. We shall proceed to find a set of scalar products, both necessary and sufficient to describe the positions of the atoms in the bulk, by building up distorted cells around an initial cell in order to form the crystal. The arrangement of the eight atoms on the corners of this first cell is determined by 18 scalar products. A convenient, although not unique, set is obtained by taking the squares of the lengths of the 12 edges of the cell (i.e., 12 diagonal products) and the 6 off-diagonal products (essentially the angles between vectors since the diagonal products give the magnitudes of the vectors)



(a)



(b)

FIG. 1. (a) The first cell around which the distorted lattice is to be built up. (b) The first cell with four of its six neighboring cells.

represented by arcs in Fig. 1(a). The four atoms in an adjacent cell which are not already fixed by the above scalar products are determined by the 8 remaining edge lengths of this cell and four more angles, as shown in Fig. 1(b). The rest of the crystal is treated by adding additional cells and using only those scalar products (the lengths of all edges and some angles) necessary to define the positions of the atoms. When this is done, the lattice points in the bulk can be divided into three types according to the number of necessary scalar products associated with them. First, there are the points lying along three lines passing through the reference point 0 in the initial cell and parallel to the three basis vectors of the undistorted lattice. Associated with this type of point are three diagonal products (edge lengths) and three off-diagonal products (angles, essentially), as shown in Fig. 2(a). Secondly, there are the points which make up the remainder of the three planes in which the three reference lines lie. Each of these points is associated with three diagonal products but only one off-diagonal product, as shown in Fig. 2(b). The third, general, class of lattice points are those which do not lie in the above-mentioned planes, and these are associated with the three diagonal scalar products but no off-diagonal products [see Fig. 2(c)].

This nonuniformity in the distribution of the scalar products through the crystal arises with every necessary and sufficient set of scalar products one chooses to describe the arrangement of atoms and is clearly undesirable. We remove this difficulty by invoking (a) the invariance of a crystal under the operations of the relevant translation subgroup, and (b) the assumption that interactions over distances of the order of the crystal dimensions are negligible. This latter assumption

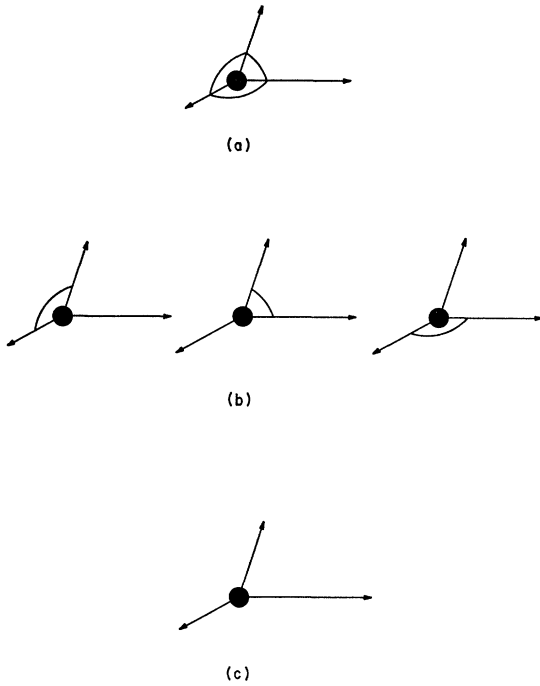


FIG. 2. Types of lattice point according to the number of scalar products associated with them: (a) points on the reference lines, (b) points in the reference planes but off the reference lines, (c) general points.

is, of course, necessary anyway if the concept of "bulk properties" is to have any meaning. In this case, we note that no extra distinctly different scalar products, or products of such products, arise in the bulk if we consider all the lattice points to be of the type shown in Fig. 2(a), i.e., associated with six scalars. In other words, we shall retain an unnecessarily large set of invariants which, however, does not give rise to any new, and therefore dependent, coefficients. This set of invariants is not an independent set; however, it is sufficiently restricted so that it is associated with an independent set of coefficients because of the lattice invariance properties. It is a complete set for the bulk of the crystal if the strain energy is, in fact, determined by the nuclear positions only (i.e., for nonmetallic crystals). In this case, even the smaller, independent set of  $3N-6$  invariants is a complete set, since this is sufficient to completely specify the positions of all the points of the distorted lattice for small displacements.

We write  $\mathbf{x}_1(l)$ ,  $\mathbf{x}_2(l)$ ,  $\mathbf{x}_3(l)$  as the position vectors of three neighbors of the atom in unit cell ( $l$ ) relative to this latter atom and which become the lattice basis vectors when the distortion is removed. The strain energy of the bulk of the crystal may therefore be written

$$V = \frac{1}{2} \sum_{l,l'} \sum_{m,n,m'}^{\mathbf{s}} B_{mnm'n'}(l-l') \lambda_{mn}(l) \lambda_{m'n'}(l') + \dots, \quad (3)$$

where  $\lambda_{mn}(l) = (\mathbf{x}_m(l) \cdot \mathbf{x}_n(l) - \mathbf{X}_m \cdot \mathbf{X}_n) / 2a$  and is symmetric in  $(m,n)$ , where the sums over  $l, l'$  range over all the unit cells in the bulk, and where  $\{B_{mnm'n'}(l-l')\}$  is invariant under the operations of the translation subgroup. In actual fact, of course,  $\{B_{mnm'n'}(l-l')\}$  must be invariant under all the operations of the space group and must be positive definite. In addition, the  $B_{m'n'mn}(l-l')$  must fall off rapidly enough as  $l-l'$  increases so that  $\sum_l B_{m'n'mn}(l-l')$  is convergent. If the  $B$ 's have these properties and one uses the expression (3) for the energy, then we have a suitably invariant form for the strain energy and all the force constants are independent.

We now extend these considerations to the case of nonprimitive structures. One suitable set of scalar products for a diatomic structure consists of the set noted above, using the atoms on one sublattice, together with three extra scalars per unit cell of the diatomic structure necessary to locate the  $B$  atom relative to the  $A$  atom. However, a more convenient set is obtained by writing  $\mathbf{x}_1(l)$ ,  $\mathbf{x}_2(l)$ , and  $\mathbf{x}_3(l)$  as the position vectors of the  $B$  atoms in neighboring unit cells relative to the  $A$  atom of cell ( $l$ ) and writing  $\mathbf{x}_4(l)$  as the position vector of atom  $B$  in cell ( $l$ ) with respect to the  $A$  atom there. The nine scalar products per unit cell are the 10 scalar products  $\mathbf{x}_m(l) \cdot \mathbf{x}_n(l)$  ( $m, n = 1, 2, 3, 4$ ) less one of the off-diagonal products ( $\mathbf{x}_3 \cdot \mathbf{x}_4$ , for example, is determined for small displacements if the other 9 scalar products are known). Thus we can write the strain energy as

$$V = \frac{1}{2} \sum_{l,l'} \sum_{m,n,m'}^{\mathbf{s}} B_{m'n'mn}(l-l') \lambda_{m'n'}(l') \lambda_{mn}(l) + \dots, \quad (4)$$

where the prime denotes that terms involving  $\lambda_{34}$  are not included (unless they arise in the course of imposing invariance on the  $B$ 's under symmetry operations—see Sec. IV). The extension to structures with a greater number of atoms per unit cell is straightforward and will not be included here.

If we wish to remain within the harmonic approximation, then we can drop terms in  $u^2$  from the  $\lambda$ 's and Eq. (3) or (4) approximates to

$$V = \frac{1}{2} K_{ab}{}^{mn} u_a^m u_b^n, \quad (5)$$

where  $u_a^m$  is the  $a$ th component of the displacement of the  $m$ th nucleus and the  $K_{ab}{}^{mn}$  are linear combinations of the  $B_{mnm'n'}$ . This form is now suitable for calculational purposes and, when derived from Eqs. (3) or (4), has the correct invariance and other properties.

The expansion (5) is essentially the starting point of Born and Huang<sup>3</sup> in their method of imposing rotational and displacement invariance on  $V$ . In this latter method, one tries to find a set of coefficients  $K_{ab}{}^{mn}$  which satisfy the rotational-invariance condition<sup>12</sup>

$$K_{ab}{}^{mn} X_c^m = K_{cb}{}^{mn} X_a^m \quad (6)$$

<sup>12</sup> See, for example, Eq. (23.23) of Ref. 3 or (5.11) of Ref. 5.

and displacement-invariance condition<sup>13</sup>

$$\sum_m K_{ab}{}^{mn} = 0, \quad (7)$$

and the crystal-symmetry conditions. It is completely equivalent to obtaining the correct form (5) from (3) or (4), but the latter method offers one or two operational advantages. It is, for example, somewhat more difficult to apply (6) and (7) than to write down an expansion in the form (3) or (4) and reduce it to the form (5). Furthermore, the physical significance of the force constants is much more readily apparent in the form (3) or (4) than in the form (5). As an example of this, we see plainly from Eqs. (2), (3), or (4) that the only nearest-neighbor harmonic interaction which arises between  $k$  and its nearest neighbor  $k'$  is the term in  $(\lambda_{kk'kk'})^2$ . *This term is clearly a central interaction and thus there are no noncentral purely first-neighbor interactions present in any nonmetallic crystal.* This result, which generalizes Lax's result<sup>5</sup> for a one-dimensional lattice to all cases, can, in principle, be derived from conditions (4) and (5), but this is more difficult and has not yet been carried out.

We note that we have also established that an arbitrary potential energy which is suitably invariant is decomposable into parts which are separately invariant and have therefore confirmed Harrison's suggestion<sup>14</sup>; the validity of this decomposition depends only on the validity of the series expansion of  $V$ . In actual fact, the result regarding noncentral first-neighbor interactions follows as soon as the decomposition is made; Lax, for example, has pointed out<sup>5</sup> that pair forces must be central. Finally, we see from Eqs. (3) and (4) that the invariance requirements give rise to appreciable contributions to the anharmonic part of the strain energy (i.e., terms in  $u^3$ , etc.) from the  $B_{mnm'n'}(l-l')$ , which lends some support to the crude procedure (see Ref. 8, for example) of estimating the magnitude of the cubic anharmonic coefficients as  $K^{(3)} \sim K^{(2)}/a$ . This result, which is apparently new, is not too surprising if we remember that, if the first-order potential coefficients are not made to vanish by imposing the extremum condition, Eq. (6) is a relation between the first- and second-order coefficients.<sup>12</sup> Thus we might expect a similar relation between the second- and third-order coefficients if the Born-Huang invariance conditions<sup>8</sup> were derived for an anharmonic strain energy. If one wishes to include cubic terms in Eq. (5), one must, of course, also include terms cubic in the  $\lambda$ 's in Eqs. (3) or (4) before the approximation to the form (5). It is clear that the procedure proposed here for imposing the necessary invariance requirements is even more convenient in comparison with the Born-Huang method if anharmonic terms are to be included.

To summarize, the proposed prescription for finding

a realistic set of force constants for monatomic or diatomic crystals is to write out expression (3) or (4), eliminate those terms which are to be ignored for the sake of simplicity by putting the corresponding  $B$ 's equal to zero, and ensure  $\{B_{mnm'n'}(l-l')\}$  is positive definite and has the required invariance properties. In Sec. IV, we shall apply this method to the calculation of the elastic constants of the diamond type of crystal in order to illustrate the mechanics and results of the method and also to show how the inconsistencies noted in Sec. II vanish if a physically realistic set of force constants is used.

#### IV. CALCULATION OF THE ELASTICITY OF CRYSTALS OF THE DIAMOND TYPE

The elasticity and lattice dynamics of the diamond structure have been the subject of a large number of calculations since the original work of Born.<sup>2</sup> This latter calculation gives unrealistic results, as we noted in Sec. II, because it is invalid by virtue of the use of a noncentral first-neighbor force constant, prohibited by the result obtained in Sec. III. The two-constant model has been extended to include more constants,<sup>15-17</sup> notably by Smith<sup>15</sup> and Herman.<sup>16</sup> Herman<sup>16</sup> included interactions out to fifth neighbors in an attempt to fit both elastic and phonon-dispersion data and showed that Smith was in error due to the neglect of some antisymmetric second-neighbor terms. Harrison<sup>14</sup> has used a model with three force constants, two describing central first- and second-neighbor interactions and a second-neighbor constant representing the effect of angle changes. Huntington<sup>18</sup> has used the first-neighbor central and second-neighbor noncentral part of this model to obtain a relation analogous to the Born relation.<sup>2,4</sup>

Herman<sup>16</sup> could fit the experimental data, but this is hardly surprising in view of the large number of parameters used. In actual fact, it appears that this fit is not physically significant since, by dropping the rigid-ion model in favor of a shell-model approach, Cochran has shown<sup>19</sup> that a near-neighbor model can then fit the data. The shell model is more realistic physically than the Herman model, and thus it appears that distant-neighbor interactions are not too important in diamond-like crystals. The shell model<sup>19</sup> introduces a relatively large number of parameters but, for many lattice dynamical purposes, is equivalent to a rigid-ion model with a smaller number of force "constants" which are wavelength-dependent (see Appendix); hence the fit with the lattice dynamical data. However, for the nondispersive long-wavelength acoustic modes, the

<sup>15</sup> H. M. J. Smith, Phil. Trans. Roy. Soc. **A241**, 105 (1948).

<sup>16</sup> F. Herman, J. Phys. Chem. Solids **8**, 405 (1959).

<sup>17</sup> See, for example, N. S. Nagandra Nath, Proc. Indian Acad. Sci. **A1**, 333 (1934); K. G. Ramanathan, *ibid.* **A26**, 481 (1947); D. Krishnamurti, *ibid.* **A33**, 325 (1951).

<sup>18</sup> H. B. Huntington, Solid State Phys. **7**, 213 (1958).

<sup>19</sup> W. Cochran, Proc. Roy. Soc. (London) **A258**, 260 (1959).

<sup>13</sup> See, for example, Eq. (23.16) of Ref. 3 or (A.2) of Ref. 5.

<sup>14</sup> W. A. Harrison, thesis, University of Illinois, 1956 (unpublished).

wavelength dependence of the rigid-ion force constants is negligible (see Appendix), and thus, as long as we wish to fit only the long-wavelength acoustic data (i.e., the elastic data), a small number of near-neighbor constants should give an adequate description. We shall therefore consider a model with only two types of interaction, a nearest-neighbor central term and a noncentral second-neighbor term. The long-range quadrupolar interaction introduced by Lax<sup>6</sup> falls off as  $r^{-5}$  and will be ignored except for close neighbors, where it will be absorbed into the two force constants. The shell-shell interaction introduces more important medium-range effects,<sup>6,8</sup> but these will generally be small<sup>8</sup> for third and more distant neighbors. The two-constant model we use is somewhat similar to the Huntington model,<sup>18</sup> but the noncentral interaction is different and gives rise to rather different results, as we shall see.

The basic unit cell of the diamond structure is a rhombohedron with two atoms (1 and 0 of Fig. 3) on its major axis, which is directed along the [111] direction. The three neighboring unit cells of interest contain atoms 2 and 5, 3 and 6, 4 and 7, respectively. We use Eq. (4) for the strain energy and will obtain our two-

constant model by including only diagonal products of the  $\lambda$ 's. Thus we write

$$V = \frac{1}{2} \sum_l \sum'_{m,n=1}^4 B_{mnmn}(O) \lambda_{mn}^2(l) \\ = \frac{1}{2} \sum_l \left[ \frac{\alpha}{4a^2} \sum_{i=1}^4 (x_{0i}^2(l) - 3a^2)^2 \right. \\ \left. + \frac{\beta}{2a^2} \sum_{i,j>i}^4 (\mathbf{x}_{0i}(l) \cdot \mathbf{x}_{0j}(l) + a^2)^2 \right], \quad (8)$$

where the atomic labeling is as in Fig. 3 and the required symmetry has been imposed on the  $B_{mnmn}(O)$ . The crystal symmetry requires that  $B_{mmmm}(O) = \alpha$  (all  $m$ ) and  $B_{mnmn}(O) = 6\beta/5$  (all  $m, n, m \neq n$ ) and the term in  $\lambda_{34}^2$  becomes included when we impose this symmetry (i.e.,  $B_{3434} = B_{1212}$ , etc.). If both  $\alpha$ , the central first-neighbor constant, and  $\beta$ , the noncentral second-neighbor constant, are positive, the condition of positive definiteness is satisfied. Equation (8) can be rewritten:

$$V = \frac{1}{2} \alpha \sum_l \{ [(u_{01} + v_{01} + w_{01})^2 + (u_{02} - v_{02} - w_{02})^2 + (u_{03} + v_{03} - w_{03})^2 + (u_{04} - v_{04} + w_{04})^2] \\ + \frac{1}{4} \beta \sum_l \{ [(u_{01} + u_{02} + v_{12} + w_{12})^2 + (u_{32} - v_{32} + w_{02} + w_{03})^2 + (u_{03} + u_{04} + v_{43} - w_{43})^2 + (u_{42} + v_{02} + v_{04} - w_{42})^2 \\ + (u_{31} - v_{01} - v_{03} + w_{31})^2 + (u_{41} + v_{41} - w_{01} - w_{04})^2] \} + O(u^3). \quad (9)$$

At this stage it becomes necessary to choose a method of calculation of the elastic constants. The Born-Huang long-waves method<sup>3,5</sup> consists of a calculation of the sound-wave combinations of elastic constants in terms of the  $K_{ab}^{mn}$  [using, for example, (26.32) and (26.33) of Ref. 3], followed by the use of Eq. (27.26) of Ref. 3 to determine the constants themselves. There is always an appreciable amount of work involved and this is particularly true in the case of nonprimitive lattices, such as diamond. We shall use an extension to nonprimitive lattices of a method which seems to have been first used by Kittel<sup>7</sup> for the simple cubic case. We assume that the strain, while inhomogeneous in general, varies sufficiently slowly so that it may be considered constant over a few unit cells; it is only in this case, of course, that macroscopic elasticity is a useful concept. The energy density  $U$  is obtained from  $V$  by dropping the sum over all unit cells and dividing by the unit cell volume. From an equation such as (9) we express the energy density  $U$  in terms of the displacement-gradient components  $u^i_{,j}$  and the  $\xi_{pq}^i$ , the components of the differences in internal strain between each pair  $pq$  of sublattices. The conditions  $\partial U / \partial \xi_{pq}^i = 0$  establish the relationship  $\xi_{pq}^k = \xi_{pq}^k(u^i_{,j})$  and one can then write the energy density as  $U(u^i_{,j})$ . This expression is then compared with the well-known macroscopic expression to yield relations between the elastic coefficients and the force constants. The method is straightforward and involves a relatively small amount of computational work, as the present example shows.

We write the strains as  $e_{xx} = \partial u / \partial x$ ,  $e_{yz} = (\partial v / \partial z) + (\partial w / \partial v)$ , etc., and write the components of the internal strain as  $u', v', w'$ . Equation (9) becomes

$$U = \frac{\alpha}{32a} \left\{ \left[ \left( e_d + e_{yz} + e_{zx} + e_{xy} + \frac{u' + v' + w'}{a} \right)^2 + \left( e_d + e_{yz} - e_{zx} - e_{xy} + \frac{u' - v' - w'}{a} \right)^2 \right. \right. \\ \left. \left. + \left( e_d - e_{yz} - e_{zx} + e_{xy} - \frac{u' + v' - w'}{a} \right)^2 + \left( e_d - e_{yz} + e_{zx} - e_{xy} - \frac{u' - v' + w'}{a} \right)^2 \right] \right\} \\ + \frac{\beta}{16a} \left\{ \left[ \left( e_{xx} - e_{yy} - e_{zz} - e_{yz} + \frac{u'}{a} \right)^2 + \left( e_{xx} + e_{yy} - e_{zz} - e_{xy} + \frac{w'}{a} \right)^2 + \left( e_{xx} - e_{yy} - e_{zz} + e_{yz} - \frac{u'}{a} \right)^2 \right. \right. \\ \left. \left. + \left( e_{xx} - e_{yy} + e_{zz} - e_{xz} + \frac{v'}{a} \right)^2 + \left( e_{xx} - e_{yy} + e_{zz} + e_{zx} - \frac{v'}{a} \right)^2 + \left( e_{xx} + e_{yy} - e_{zz} + e_{xy} - \frac{w'}{a} \right)^2 \right] \right\},$$

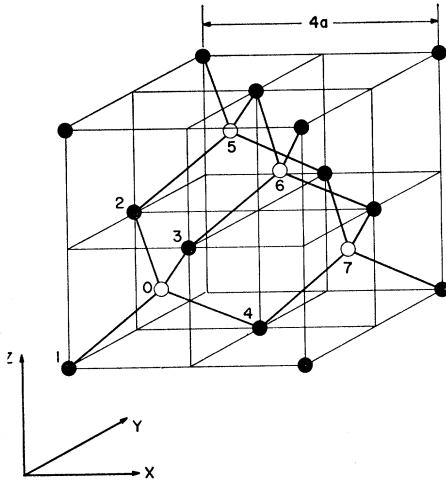


FIG. 3. The crystal model. The open and filled circles represent the atoms on the two different sublattices.

where  $e_a = e_{xx} + e_{yy} + e_{zz}$ . By imposing the condition

$$\frac{\partial U}{\partial u'} = \frac{\partial U}{\partial v'} = \frac{\partial U}{\partial w'} = 0,$$

one obtains  $u' = -a\zeta e_{yz}$ ,  $v' = -a\zeta e_{zx}$ ,  $w' = -a\zeta e_{xy}$ , where  $\zeta = (\alpha - \beta)/(\alpha + \beta)$ . After substitution for  $u'$ , etc., we compare the resulting expression for the energy density with the well-known macroscopic expression<sup>7</sup> for cubic crystals:

$$U = \frac{1}{2}c_{11}(e_{xx}^2 + e_{yy}^2 + e_{zz}^2) + c_{12}(e_{yy}e_{zz} + e_{zz}e_{xx} + e_{xx}e_{yy}) + \frac{1}{2}c_{44}(e_{yz}^2 + e_{zx}^2 + e_{xy}^2)$$

and obtain

$$\begin{aligned} c_{11} &= (\alpha + 3\beta)/4a, \\ c_{12} &= (\alpha - \beta)/4a, \\ c_{44} &= \alpha\beta/a(\alpha + \beta), \end{aligned} \quad (10)$$

which are somewhat different from the invalid expressions obtained by Born.<sup>2,4</sup> Because there are three elastic coefficients and two force constants, expressions (10) predict a relation between the elastic coefficients analogous to the well-known but invalid Born relation.<sup>2,4</sup> In the present case,

$$2c_{44}(c_{11} + c_{12}) = (c_{11} - c_{12})(c_{11} + 3c_{12}) \quad (11)$$

is the predicted relation between the elastic constants.

## V. DISCUSSION OF THE RESULTS OF THE CALCULATION

The validity of the model chosen for this calculation depends in part on Eq. (11) being well satisfied by the relevant experimental data for crystals of the diamond type. In Table I, we present theoretical fits of the elasticity data, and values of

$$\frac{2c_{44}(c_{11} + c_{12})}{(c_{11} - c_{12})(c_{11} + 3c_{12})}$$

for diamond, silicon, and germanium, using the data listed by Kittel.<sup>7</sup> We note from Table I that expression (11) is very well satisfied by all these crystals. The agreement is much better than that obtained from the invalid Born relation<sup>2,4,18</sup> and very much better than that obtained from the Huntington relation.<sup>18</sup> The Huntington second-neighbor interaction is somewhat different from that utilized in the present work and is equivalent to a strain energy in which there are the additional contributions

$$\frac{1}{2} \sum_l [(3/5)2\beta \sum_{m,n>m} \lambda_{mm}(l)\lambda_{nn}(l) - (4/5)\beta \sum_{m \neq n \neq n'} \lambda_{mn}(l)\lambda_{nn'}(l)],$$

over and above Eq. (8). The Huntington model is less adequate because it treats the second-neighbor interaction in an unsatisfactory way. For example, in this latter model, a variational displacement of atom 1 along the [111] direction has no effect on atoms 2, 3, 4, whereas the shell-shell and shell-core interactions will ensure that this type of second-neighbor interaction should be comparable with the other second-neighbor terms—as it is in the model used in the present work. The success of the present model in fitting the elasticity data suggests that long-range interactions are, indeed, unimportant in these crystals and thus confirms that the shell-model fit<sup>19</sup> of the phonon-dispersion data is physically more significant than the fit obtained by Herman.<sup>16</sup>

We note that the bulk modulus given by Eqs. (10) is again dependent on the noncentral constant  $\beta$ , but this is now satisfactory since the noncentral interaction of the present model contributes to the strain energy when the lattice constant is changed but the angles are not. This is because  $\mathbf{x}_{01} \cdot \mathbf{x}_{02}$  does not vanish when the bond angle is at its equilibrium value (since this value is not  $\pi/2$  in diamond) and is one reason why the present model gives results which are more realistic than those obtained from the Huntington model, as discussed in the previous paragraph. The compliance constant  $s_{11}$  does not now exhibit the unacceptable

TABLE I. The validity of relation (11) for diamond-like crystals. (Elastic constants in  $10^{12}$  dyn/cm<sup>2</sup>, force constants in  $10^6$  dyn/cm).

	Diamond	Silicon	Germanium
$\frac{2c_{44}(c_{11} + c_{12})}{(c_{11} - c_{12})(c_{11} + 3c_{12})}$	0.99	0.99	1.07
$\alpha$	1.29	0.485	0.38
$\beta$	0.85	0.138	0.12
$\beta/\alpha$	0.66	0.29	0.32
$c_{11}$ (expt)	10.76	1.66	1.29
$c_{11}$ (theoret)	10.76	1.66	1.31
$c_{12}$ (expt)	1.25	0.64	0.48
$c_{12}$ (theoret)	1.25	0.64	0.46
$c_{44}$ (expt)	5.76	0.79	0.67
$c_{44}$ (theoret)	5.75	0.79	0.65

behavior noted in Sec. II for the Born results. Thus we see that the inconsistencies which arise from the invalid Born model are no longer present if a realistic set of force constants is chosen.

Our analysis readily yields theoretical values of the Kleinman internal strain parameter,<sup>20</sup> which is the ratio of the internal base strains with and without non-central forces. Thus, since  $\zeta=1$  if  $\beta=0$ , the Kleinman parameter is identical with the  $\zeta$  parameter of this article. We see that

$$\zeta = \frac{\alpha - \beta}{\alpha + \beta} = \frac{2c_{12}}{(c_{11} + c_{12})}, \text{ using Eq. (10).}$$

Hence, we expect  $\zeta$  to be 0.55 for silicon, 0.53 for germanium, and 0.21 for diamond. These values predicted for germanium and silicon are roughly 10% lower than the experimental values of Segmüller and Neyer,<sup>21</sup> while no experimental value is yet available for diamond. The displacements of the two types of  $\{111\}$  plane under  $\langle 111 \rangle$  uniaxial stress are clearly comparable for germanium and silicon. Thus the force constants  $k, k'$  between neighboring Wannier states used in a previous one-dimensional calculation<sup>8</sup> of the second-order electric moment in these crystals are comparable, as was assumed. The aside<sup>8</sup> regarding the reason why the Born relation is less good for diamond is no longer very relevant, of course. The fact that the  $a_{13}$  of that work is not small in diamond is no reason for failure of the new relation (11) since second-neighbor terms are now included. Instead, the larger value of  $a_{13}$  suggests that the second-neighbor constant is larger relative to the first-neighbor constant in diamond than in either silicon or germanium; this is confirmed in the present work, as the values of  $\beta/\alpha$  given in Table I show.

It is also of interest to consider the elasticity of the zincblende III-V compounds from the viewpoint of the present treatment. An extension of the present theory to these solids would be possible (by replacing  $\beta$  by the mean value of  $\beta$  for the two components of the compound) were it not for the increased importance of long-range forces due to the presence of coulomb effects. In fact, the deviations from relation (11) correlate well with the values of  $e_B^*$ , the Born effective charge,<sup>22</sup> as shown in Table II. Thus the anomaly discussed by Westbrook,<sup>23</sup> for example, has now been resolved.

<sup>20</sup> L. Kleinman, in *Proceedings of the International Conference on the Physics of Semiconductors, Exeter, 1962* (The Institute of Physics and the Physical Society, London, 1962).

<sup>21</sup> A. Segmüller and H. R. Neyer, *Physik Kondensierten Materie* 4, 63 (1965).

<sup>22</sup> Because of the delocalized nature of the charge in these crystals, the Lorentz treatment of the internal field is not valid and the Born charge is more useful than the Szigeti charge [see, for example, E. Burstein, in *Phonons and Phonon Interactions*, edited by T. Bak (W. J. Benjamin, Inc., New York, 1964), where the effective-charge data were also obtained].

<sup>23</sup> J. H. Westbrook, in *Mechanical Properties of Intermetallic Compounds*, edited by J. H. Westbrook (John Wiley & Sons, New York, 1960).

TABLE II. Deviations from relation (11) in zinc blende III-V crystals.<sup>a</sup>

	$2c_{44}(c_{11}+c_{12})$	$e_B^*$
	$(c_{11}-c_{12})(c_{11}+3c_{12})$	$e$
InAs	1.22	2.7
InSb	1.16	2.5
GaAs	1.13	2.2
AlSb	1.11	2.2
GaSb	1.11	2.0

<sup>a</sup> Elasticity data from O. Madelung, *Physics of III-V Compounds* (John Wiley & Sons, Inc., New York, 1964).

## ACKNOWLEDGMENTS

The author wishes to acknowledge a number of extremely valuable discussions with Dr. Melvin Lax (particularly his pointing out the importance of the rotational-invariance requirement) and the advice and encouragement of Dr. G. Rupprecht.

## APPENDIX

For the purposes of lattice dynamics, a shell-model description of a harmonic crystal is equivalent to a rigid-ion description with wavelength-dependent force "constants" and this wavelength dependence is negligible for nondispersive acoustic waves. We shall illustrate this by considering a one-dimensional infinite monatomic lattice supporting longitudinal waves. We take the strain energy as

$$V = \frac{1}{2}k \sum_m (u_m - u_{m-1})^2 + \frac{1}{2}k_1 \sum_m (u_m - v_m)^2 + \frac{1}{2}k_2 \sum_m (v_m - v_{m-1})^2,$$

where  $u_m$  and  $v_m$  are the longitudinal displacements of the core and shell, respectively, of the  $m$ th atom (see Ref. 8). This form is not rotationally invariant because only one dimension is considered. The equations of motion for core and shell are

$$M\ddot{u}_m = k(u_{m+1} + u_{m-1} - 2u_m) - k_1(u_m - v_m), \quad (\text{A1})$$

$$0 = k_1(u_m - v_m) + k_2(v_{m+1} + v_{m-1} - 2v_m). \quad (\text{A2})$$

We can write the particular solutions of these equations as  $u_m = Ae^{i(\omega t - kxm)}$  and  $v_m = Be^{i(\omega t - kxm)}$ , where  $B/A$  is real because there is no phase difference—the shells follow the nuclear motion in full. From (A2),

$$k_1(A - B) = 2k_2B(1 - \cos ka)$$

or

$$B = k_1A[2k_2(1 - \cos ka) + k_1]^{-1}.$$

Thus, Eq. (A1) becomes

$$M\omega^2 = 2k(1 - \cos ka) + \frac{2k_1k_2(1 - \cos ka)}{2k_2(1 - \cos ka) + k_1}. \quad (\text{A3})$$

We see that the lattice behaves as a rigid-ion model would if the interatomic force "constant" were

$$k + \frac{k_1k_2}{k_1 + 2k_2(1 - \cos ka)},$$



i.e., if it were wavelength-dependent. Thus the three shell-model force constants are equivalent to one wavelength-dependent rigid-ion model force "constant"; this wavelength dependence is negligible for nondispersive modes. Equation (A3) can be written, for small  $\kappa$ ,

$$M\omega^2 = (k+k_2)\kappa^2 a^2 - \left[ \frac{2(k+k_2)}{4!} - k_2^2/k_1 \right] \kappa^4 a^4 + O(\kappa^6).$$

The concept of elasticity is useful only for slowly varying deformations, i.e., when terms in  $\kappa^3 a^3$  and higher are small compared with the term in  $\kappa^2 a^2$ . Thus the three shell-model force constants are equivalent to one wavelength-independent rigid-ion constant  $k' \equiv (k+k_2)$  as long as we are concerned only with elasticity data. The modifications due to the use of the shell model appear only at shorter wavelengths.

PHYSICAL REVIEW

VOLUME 145, NUMBER 2

13 MAY 1966

### Adiabatic Theory of Nearly Small Polarons\*

D. M. EAGLES

*Post Office Research Station, London, N. W. 2, England*†

(Received 1 June 1965; revised manuscript received 20 December 1965)

An adiabatic theory of polarons is developed under conditions when the polarons are "nearly small," that is, when their binding energy is greater than or of the order of half the rigid-lattice bandwidth, but is not so great that small-polaron theory is applicable. In such an adiabatic theory there are a set of minimum-energy nuclear configurations, each of which has associated electronic wave-functions concentrated mainly on one positive ion, but spreading slightly to its neighbors. At and near these configurations trial electronic wave functions are taken in the form of linear combinations of single-ion functions for a particular ion and its nearest neighbors, with coefficients determined by minimizing the energy. Conditions for localized normal modes to be associated with any minimum are examined. Properties of wave functions describing the nuclear motion are studied within the framework of a generalized tight-binding approach. It is shown that, if localized modes are not formed, then bandwidths will decrease with increasing temperature as in small-polaron theory, but that when localized modes are present, then at nonzero temperatures, a discrete distribution of bandwidths will occur, and the thermal average of these widths may increase with increasing temperature. Optical absorption due to transitions between a wide valence band for which electron-phonon coupling is neglected and an adiabatic nearly-small-polaron conduction band with localized modes is considered. Results for absorption at absolute zero are similar to those obtained previously for a small-polaron conduction band without localized modes, but the temperature dependence of the absorption obtained here shows some new features. Parameters occurring in the theory are estimated using a continuum-polarization model for electron-phonon interactions, and numerical values are found for a simplified model of a possible conduction band in SrTiO<sub>3</sub>. The calculations indicate that electronic states may exist from which both adiabatic nearly small polarons and weak-coupling large polarons can be formed, and that the lowest energy polaron state may suddenly change from one type to the other as electronic overlap integrals or electron-phonon interactions are altered in magnitude.

#### I. INTRODUCTION

**A**N excess electron in an ionic crystal produces a polarization of the lattice around it by Coulomb interaction with the surrounding ions. The complex of electron plus its surrounding lattice polarization is usually called a polaron, and the study of polarons is thus equivalent to the study of electrons interacting with a phonon field. Different methods of approach to the polaron problem are suitable according to the strength of the electron-phonon coupling<sup>1-2</sup>—for weak coupling it is a fair approximation to assume that the

ionic displacements follow the motion of the electron, while for strong coupling the opposite approximation that the electron adjusts to the ionic motion adiabatically is generally more suitable.

If one looks at the polarization as a function of distance from the center of the polaron, which may be the instantaneous position of the electron (for weak coupling), or its averaged position over motion in its self-induced potential well (for strong coupling), then at large distances the polarization is just that which would be induced by a fixed point charge at the polaron center, while at short distances from the center the polarization potential flattens off.<sup>1</sup> The distance below which this potential flattens off may be said to define a polaron radius.

If the radius is appreciably greater than a lattice constant, a fair description of the system of one electron plus optical phonons is given by a Hamiltonian

\* The work described here, together with some previous work on optical properties of small polarons (Ref. 1) has been used in a thesis accepted for the examination for the Ph.D. degree in the Faculty of Science at London University.

† Acknowledgment is due to the Engineer-in-Chief, Post Office, for permission to publish this work.

<sup>1</sup> H. Fröhlich, *Advan. Phys.* **3**, 325 (1954).

<sup>2</sup> G. R. Allcock, *Advan. Phys.* **5**, 412 (1956).