SOME PROBLEMS IN THE THEORY OF THE SOLID STATE.

(A thesis submitted for the degree of Doctor of Philosophy at the University of Oxford, January 1965.)

by

V. M. HALPERN.

Mathematical Institute,
Oxford.

BODLEIAN LIBRARY OXFORD
22.VII.65

ACKNOWLEDGEMENTS.

The author wishes to express his thanks to his supervisor, Professor C. A. Coulson, for his constant help and encouragement, to Dr. J. Mayers for his advice with regard to the computer programmes and for providing the atomic π wave functions used, to Mrs. E. Chin and the staff of the Oxford University Computing Laboratory for running the programmes, to Dr. G. Doggett for many valuable discussions in connection with the first part of this thesis, to Dr. J. C. Phillips for providing the O.P.W. wave functions for silicon, to Professor P. T. Landsberg and Mr. C. J. Hearn of University College, Cardiff, for several valuable discussions in connection with the last chapter of this thesis, and to D.S.I.R. for a research studentship during the tenure of which the work reported herein was performed. Finally, he expresses his thanks to the Almighty for enabling him to carry out the research and produce this thesis.

CONTENTS.

	Page
ABSTRACT.	S 1.
References for abstract.	S 11.
CHAPTER 1. Molecular Orbitals In Crystals.	
1.1 The Hartree-Fock Equations.	1.
1.2 Crystal Symmetry and Group Theory.	11.
1.3 Methods of Calculating Molecular Orbitals in Crystals.	24.
1.4 The Relationship between L.C.A.O. and O.P.W. Methods.	39.
CHAPTER 2. Equivalent Orbitals In Tetrahedral Crystals.	
2.1 Equivalent Orbitals and Molecular Orbitals.	47.
2.2 A Variational Treatment of μ Equivalent Orbitals.	58.
2.3 The Formalism for Tetrahedral Semiconductors.	66.
CHAPTER 3. Our Method Of Calculating The Equivalent Orbitals.	
3.1 The Problem of Convergence.	71.
3.2 The Choice of Basis Functions.	75.
3.3 The Crystal Potential.	81.
3.4 Other Approximations Made in the Calculations.	86.
CHAPTER 4. The Results Of Our Calculations.	
4.1 The One-electron Energy Bands.	91.
4.2 Equivalent Orbitals in Diamond - type Crystals.	96.
4.3 Equivalent Orbitals in Zincblende - type Crystals.	98.
4.4 Further Discussion of our Results.	101.
CHAPTER 5. Overlap Functions In Planar Crystals.	
5.1 The Method Of Calculation.	108.
5.2 The Values of the Overlap Integrals.	114.
5.3 Theoretical Considerations and Conclusions.	117.
References for the thesis.	121.
APPENDICES.	
Appendix I. The Schmidt Orthonormalisation Procedure.	i.
Appendix II. Effective Masses.	iii.
Appendix III. The Computer Programmes.	v.

ABSTRACT

During the past thirty years, a large number of calculations on the electronic structure of crystals have been made. There are several reasons why such calculations are quite attractive. From the theoretical point of view, the advantage of crystals over large molecules is that their periodicity considerably simplifies the calculations; it is much easier to calculate the one-electron energy levels of a perfect crystal of sodium or germanium, even though it may contain some 10^{22} atoms, than to find the energy levels of a non-symmetric molecule containing twenty or thirty atoms. Thus, it is natural to proceed from calculations on isolated atoms and small molecules to calculations on simple crystals, rather than to work on large poly-atomic molecules. Since most materials used in electronics are crystalline, such calculations can also lead to important applications; the energy gap between the valence and conduction bands of a semiconductor is an extremely important parameter, and generally calculations of the band structure can enable experimental results to be explained and made use of. Typical applications include the selection of materials for use in transistors, lasers and masers.

The major part of this thesis is devoted to calculations on the valence bands of semiconducting crystals having the diamond or zincblende structures, which can also be called tetrahedral semi-conductors. While a lot of work had been done previously on diamond[1], silicon[2] and germanium[3], it is only in the past few years that an appreciable number of results have been published for compounds having the zincblende structure[4,5]. However, several standard methods, such as the orthogonalised plane wave[6] and augmented plane wave[7] methods, are now available and in use for full-scale calculations[8]. These methods generally use at least fifty plane waves in a variational procedure to find the one-electron eigenvalues and eigenfunctions for the valence and lowest conduction bands, and seem to be fairly accurate. We had neither the desire nor the computer facilities to perform such calculations, but were concerned with a different aspect

of these crystals. Classically, the ground state of these tetrahedral semiconductors can be represented by associating all the valence electrons with covalent bonds[9], and it therefore seemed of interest to investigate whether an accurate quantitative treatment could employ wave functions representing such bonds. Most work done previously using this approach has been based either on the LCAO method[10,11] or on a related interpolation scheme[12]. Since ours was, to the best of our knowledge, the first attempt to find bond functions without making the rather drastic a priori assumptions of the LCAO method, we decided to attempt approximate calculations for a wide range of substances rather than an extremely accurate calculation for just one or two compounds. This decision was also prompted by the fact that only a Mercury computer was available, and unlike Edsac, Atlas, and the other newer machines this was not fast enough and did not have enough storage space to make accurate calculations really feasible.

In the first part of chapter one of this thesis, we derive the Hartree-Fock equations in a form which is convenient for calculations on crystals. In section 1, we introduce a division of the Coulomb and exchange terms in these equations which differs from the conventional one. It is customary with a closed-shell system to write the electron-electron interaction terms as the sum of (i) the Coulomb potential due to all electrons in a different orbital from the one being considered, (ii) the Coulomb potential due to an electron in the same orbital as the one being considered but having opposite spin, and (iii) the exchange interaction due to all the electrons other than the one being considered. A disadvantage of this scheme is that none of these terms individually is invariant under a unitary transformation of the one-electron orbitals, although their sum of course is. We preferred, therefore, to divide the electron-electron interactions into just two parts, viz. the Coulomb potential due to all the electrons in the crystal, <u>including</u> the one being considered, and the exchange potential due to all the electrons, again including the one being

considered. The total potential is then the same as previously, but now each term separately is invariant under a unitary transformation of the wave functions.

The importance of this step first emerges in section 2, where we consider the effects of crystal symmetry. The section begins with a brief outline of some of the more useful aspects of group theory, and after this we introduce a one-electron Hamiltonian H. With our division of the electron-electron interactions, this can be so defined that the electronic wave functions enter into it only through a spin-independent first-order density matrix $q(\underline{r}' | \underline{r})$. This Hamiltonian H is equivalent to the Hartree-Fock operator in its effect on orbitals occupied in the ground state of the crystal. If we assume that the original many-electron ground state is non-degenerate, so that its wave function has the full crystal symmetry, then all the density matrices derived from it also have this symmetry. It therefore seems reasonable in the Hartree-Fock approximation to restrict our consideration to those Slater determinants for the ground state which lead to first-order density matrices having the full crystal symmetry. If this is done, our Hamiltonian H also has the full crystal symmetry, so that all the valuable results of group theory can be applied to its eigenfunctions: this is generally assumed in calculations on crystals, but is seldom explicitly stated or proved. For orbitals not occupied in the ground state, our one-electron Hamiltonian H is not equivalent to the Hartree-Fock operator; on the other hand, the approximation of a single Slater determinant is unlikely to be valid for such excited states. It is possible to give the name "generalised molecular orbital" to those eigenfunctions of H not corresponding to orbitals occupied in the ground state of the crystal, and the corresponding eigenvalues are in fact usually used to represent the conduction band.

The rest of chapter one, in a slight digression from the main theme of the thesis, examines various different methods of calculating

molecular orbitals in a crystal. Particular attention is paid to the orthogonalised plane wave method[6], and it is shown how the effect of orthogonalisation to the core states, which is often represented by a pseudopotential[1,13], can be very simply and effectively treated by a matrix routine. This approach can only be used if a finite set of basis functions is employed, but this restriction is not important in practice; moreover, our method copes with such problems as the possible over-completeness of a basis set once orthogonalisation to the core functions has taken place. Unlike the pseudo-potential method, our scheme allows exactly for the normalisation of the total wave function, and it can also be used with the modified plane wave method, L.C.A.O. methods, etc; we also used it in our bond orbital calculations.

After examining the LCAO and LCBO methods, we proceed in the last section of chapter one to establish the formal relationship between these methods and the OPW one[14]. To the extent that these methods are regarded as a means of calculating molecular orbitals (an alternative approach to the LCBO method is suggested in chapter two), we show that they differ from the OPW method only in attempting to expand the smooth parts of the molecular orbitals in terms of a few Bloch waves constructed from atomic orbitals instead of in terms of a much larger number of plane waves[15]. This procedure is not inherently unreasonable, but requires justification by comparing the resulting wave functions with those obtained by a more accurate method such as the OPW one, just as the use of Slater[16] orbitals for atoms can be justified by comparing them with Hartree-Fock orbitals. A formalism for carrying out such a comparison is given, and is applied to silicon; incidentally, it might be possible to develop this method into an interpolation scheme for wave functions, which could be considerably more accurate than the more conventional method of interpolating by use of energy parameters[12] in order to find the shape of the energy bands away from points of symmetry in reciprocal space.

We also show that the LCAO method, as usually employed, cannot be expected to give good values for the effective masses of the electrons, so that it is not justifiable to attack the method as a whole because of this failure.

Chapter 2 of the thesis is devoted to a derivation of the equations which equivalent orbitals[9] must satisfy. We first show how these equations can be derived from those for molecular orbitals, and how a self-consistent one-electron Hamiltonian can, in principle, be found much more easily for equivalent orbitals than for molecular orbitals. One problem with equivalent orbitals such as Wannier functions is that they are not very well localised, and since they do not in general correspond to any intuitive ideas of chemical bonds it ia difficult to know what sort of basis functions should be used for calculating them. Fortunately for tetrahedral semiconductors the position is rather simpler, since a set of equivalent orbitals exists having the same symmetry as the classical chemical bonds. These equivalent orbitals are henceforth called bond orbitals; the L.C.A.O.-type bond orbitals used by Stocker[10] and by Cohan[11] can be regarded as approximations to them, even though these approximations (especially Stocker's) do not satisfy the required orthonormality conditions. At the end of section 1, we introduce the model substance springium in an attempt to show how the relationship between equivalent and molecular orbitals resembles that between localised and normal coordinates in the classical theory of mechanical vibrations. In section 2 of this chapter we present an alternative derivation of the equivalent orbital equations, based on that given by Koster[17] but derived somewhat more rigorously for the special case of bond orbitals. This variational treatment of equivalent orbitals leads to a much clearer picture of what they should look like, and the molecular orbitals and their associated energies can easily be derived. Finally, we set up in section 3 an exact formalism for finding the bond orbitals and deriving the one-electron energies for the valence bands of tetrahedral semiconductors.

So far our treatment has been completely rigorous, and it is satisfying to know that such a sound basis exists for the calculation of bond orbitals. In chapter three, we describe the various approximations which we found it necessary or advisable to make so as to be able to carry out actual calculations on the Mercury computer. We first treat the problem of adapting the variational procedure so as to be able to obtain self-consistent solutions of the equivalent orbital equation: the problem of obtaining convergence was far more tricky than than been foreseen initially. We also introduce the localisation potential D, which is a pseudopotential representing the effect of requiring the bond orbitals to be orthogonal to each other. This localisation potential is the extra term in the equation for equivalent orbitals as compared with that for molecular orbitals, and while its expectation value for an exact bond orbital must be zero it has a profound effect on the shape of the self-consistent solutions of the equivalent orbital equation.

The next problem is the choice of basis functions in terms of which to expand the bond orbitals. We decided not to use Slater[16] orbitals, such as Stocker[10] and Cohan[11] used, because of the difficulty of evaluating exactly many of the integrals that would be needed. The idea of using Gaussian orbitals centred on the atoms at each end of the bond was also rejected, because it seemed likely to require too large a set of basis functions. Instead, we chose to use a set of Gaussian functions centred about the mid-point of the bond, and to orthogonalise them to the core orbitals by the matrix method derived in chapter one. We used a total of twelve basis functions, six of them symmetric about the mid-point of the bond and the remainder anti-symmetric about it. Because our basis set was so small, and could not conveniently be enlarged because of the limited storage facilities on the Mercury computer, we restricted ourselves to axially symmetric functions instead of only requiring them to have the C_{3v} symmetry required by group theory. If we denote distance along the bond by ζ, our

functions were $e^{-Cr^2} \times 1$, r^2, ζ^2, r^4, $r^2\zeta^2$, ζ^4; ζ, $r^2\zeta$, ζ^3, $r^4\zeta$, $r^2\zeta^3$, and ζ^5. The choice of the parameter C proved very difficult, as is not unusual in calculations involving Gaussian functions. It has the dimensions of $(\text{length})^{-2}$, and so we decided to make it inversely proportional to the square of the bond length. It was not obvious a priori how the constant of proportionality should vary with the atomic numbers and electronegativities of the atoms composing the crystal, and so we took it to be independent of these quantities; we chose it by examining the results of calculations on silicon, germanium, and tin using various different values of C. Our results suggest that this was not a terribly good choice for C, but this knowledge only became available after the calculations had been carried out.

In treating the crystal potential, our method of dividing up the electron-electron interaction was extremely valuable. Since our Coulomb potential term has the full crystal symmetry, we were able to represent it by its Fourier transform; a similar procedure was adopted with the exchange potential for which we used the Slater approximation[18]. It was thus extremely simple to find the matrix elements of the one-electron Hamiltonian between our basis functions. The Coulomb potential could easily be made self consistent since it is additive for the core and valence electrons, but since the Slater exchange potential is proportional to the cube root of the charge density we had to use an approximation of dubious validity in order to make it nominally self-consistent. In fact an exact exchange potential can be used fairly easily with Gaussian functions[19] - this is one of the chief advantages of such functions for the calculation of wave functions - but we felt that to do so would take more time on the computer than would be justified for our approximate calculations. Considerations of computing time also led us to make several other approximations, including a rather drastic one in connection with the localisation potential D. None of these approximations is

essential to the method, and our computer programmes can very easily be adapted to avoid all of them except that to the exchange potential, for which a new programme would have to be written. For our core orbitals, we used the Hartree-Fock wave functions and energies for the free atoms which Dr Mayers has calculated, and introduced only a slight shift in the energies to allow for the crystal environment.

In chapter four, we present the results of our calculations; while we did not expect to obtain definitive results, the energy bands that emerged from our work were in fact a lot poorer than we had hoped. It was found that the shape of the energy bands depended critically on the choice of the parameter C and on the contribution of the valence electrons to the crystal potential. The effect was not so noticeable for the homonuclear compounds silicon, germanium, and gray tin, for which our results agree qualitatively with those of other authors, but it was of crucial importance in the heteronuclear III-V and II-VI sompounds. We have drawn the energy bands for a number of compounds in which the results did not seem too unreasonable, but they must be regarded more as possible band shapes for heteronuclear compounds than as definitive energy bands; incidentally, we found that the anti-symmetric part of the bond orbital had a far greater effect on the energies than would have been predicted from $\underline{k}.\underline{p}$ perturbation theory[20]. Our bond orbitals may well be quite close to the correct shapes, but it would be dangerous to draw any specific conclusions from them without first performing some calculations on a much larger scale, perhaps using one of the larger and faster computers now coming into service.

However, there are one or two general conclusions that can be drawn from our results. Our calculations on Si, Ge, and Sn support the view of Mooser and Pearson[21] that the bonds tend to become less localised as the principal quantum number n of the valence electrons of the constituent atoms increases, but their suggestion about the effect of the ionicity of the bond on its degree of localisation cannot be properly judged from our results for the heteronuclear

compounds. It seems that it is certainly justifiable to talk of covalent chemical bonds in III-V compounds, and that such bonds can be represented by accurate analytical wave functions and are not just loose concepts. It is not really practicable to talk about bond ionicities or effective charges in our method, and the only meaningful concept of this nature in our scheme would seem to be the charge distribution in the bond orbital. This quantity, unlike the others, can be defined without the use of the LCAO approximation and without attempting to allocate among the atoms electrons which essentially belong to bonds and are shared by pairs of atoms. Our results for the II-VI compounds are less reliable, and more unexpected, than those for the III-V crystals. It may be that our choice of C and approximate exchange potential introduced a greater error for these compounds. It is also possible that the Gaussian functions should not be centred about the mid-point of the bond in view of the difference in the number of valence electrons which the neutral atoms would possess: we note that our whole theory applies with only a few minor alterations if the Gaussian functions are centred about some other point on the bond. However, it is quite possible that the bond orbital picture is not very suitable for II-VI compounds; there may be a tendency for the group II atom to surrender some of its electrons to the group VI atom, thereby producing ionic bonding, rather than for charge to be transferred in the opposite direction to produce covalent bonding. One could perhaps distinguish between the two types of bonding by examining the bond charge density: for covalent bonding one would expect it to show no minima and just one maximum, while for ionic binding one might find maxima in the valence charge density at the two ends of the bond and a minimum in between them. On this basis, our results suggest that in all the II-VI compounds we considered, with the exception of CdTe, there is ionic binding. It should be remembered that for certain ratios of the atomic radii the zincblende structure is the most favourable energetically for a classical ionic compound,

so that this structure does not necessarily imply the existence of covalent bonds.

The last chapter of our thesis is concerned with an entirely separate problem, namely the calculation of the overlap between the modulating part $u_{\underline{k}}(\underline{r})$ of the wave function $\psi_{\underline{k}}(\underline{r}) = e^{i\underline{k}.\underline{r}}u_{\underline{k}}(\underline{r})$ for electrons in the valence and conduction bands of planar crystals. These overlap functions are of some interest in connection with the Auger[22] effect in semiconductors, as well as in other contexts; they had previously been calculated for a one-dimensional Kronig-Penney[23] model, but not for any real crystal using wave functions obtained from a variational calculation of the energy bands. The crystals we considered were graphite and planar boron nitride, for which wave functions had been found by Coulson and Taylor[24]. The calculation of these integrals involved the use of bipolar coordinates and the summation of a double series, and this work was done on a Mercury computer. The results confirm that for bands which are not degenerate it is a reasonable approximation to set the overlap integral for electrons in the same band equal to one provided that the $\underline{k}$-vectors of the electrons do not differ by too much; however, where (as in graphite) the band is degenerate, this is not a valid approximation if $\underline{k}$ is anywhere near the point of degeneracy. An account of this work has been published in the Journal of the Physics and Chemistry of Solids[25].

References

1. L. Kleinman and J.C. Phillips, Phys.Rev. 116, 880 (1959)
2. T.O. Woodruff, Phys.Rev. 103, 1159 (1956)
3. F. Herman, Physica 20, 801 (1954)
4. L. Kleinman and J.C. Phillips, Phys.Rev. 117, 460 (1960)
5. F. Bassani and Y. Yoshimine, Phys.Rev. 130, 20 (1963)
6. C. Herring, Phys.Rev. 57, 1169 (1940)
7. J.C. Slater, Phys.Rev. 92, 602 (1953)
   M.M. Saffren and J.C. Slater, Phys.Rev. 92, 1126 (1953)
8. F. Herman and S. Skillman, Proc.Internat.Conf. on Semiconductor Physics, Prague, 1960, p.20.
9. G.G. Hall, Phil.Mag. 43, 338 (1952)
10. D. Stocker, Proc.Roy.Soc.A 270, 397 (1962)
11. N.V. Cohan, D. Pugh, and R.H. Tredgold, Proc.Phys.Soc. 82, 65 (1963)
12. J.C. Slater and G.F. Koster, Phys.Rev. 94, 1498 (1954)
13. M.H. Cohen and V. Heine, Phys.Rev., 122, 1821 (1961)
14. R.H. Parmenter, Phys.Rev. 100, 573 (1955)
15. E. Brown, Phys.Rev. 126, 421 (1962)
16. J.C. Slater, Phys.Rev. 36, 57 (1930)
17. G.F. Koster, Phys.Rev. 89, 67 (1953)
18. J.C. Slater, Phys.Rev. 81, 385 (1951)
19. S.F. Boys, Proc.Roy.Soc. A 200, 542 (1950)
20. E.O. Kane, J.Phys.Chem.Sol. 1, 249 (1955)
21. E. Mooser and W.B. Pearson, Acta Cryst. 12, 1015 (1959)
22. A.R. Beattie and P.T. Landsberg, Proc.Roy.Soc. A 249, 16 (1959)
23. A.R. Beattie and P.T. Landsberg, Proc.Roy.Soc. A 258, 486 (1960)
24. C.A. Coulson and R. Taylor, Proc.Phys.Soc. A 65, 815 and 834 (1952)
25. V. Halpern, J.Phys.Chem.Sol. 24, 1495 (1963)

CHAPTER 1

MOLECULAR ORBITALS IN CRYSTALS

1.1 THE HARTREE-FOCK EQUATIONS.

The Schrodinger Equation.

The basic equation for all non-relativistic wave mechanics of systems in a state of static equilibrium is the Schrodinger equation

$$\mathcal{H}\Psi = E\Psi \qquad (1.1)$$

where $\mathcal{H}$ is the Hamiltonian of the system, E is a constant, and Ψ is a function of the position co-ordinates $\underset{\sim}{r}_j$ and spin co-ordinates σ_j of the particles comprising the system. For systems involving nuclei and electrons, it is convenient to use the Born-Oppenheimer approximation. In this, it is assumed that for any given arrangement of the nuclei the electrons are so distributed as to make stationary the total energy of the system with respect to changes in the electron distribution, and that this electron energy is part of the effective potential in which the ~~electrons~~ nuclei may be regarded as moving. The error in this approximation depends on the ratio of the mass of an electron to the mass of the nucleus and on the gradient of the nuclear wave function, and is usually very small, at least at low temperatures when the nuclei do not vibrate very much.

We assume this approximation, and write $\underset{\sim}{s}_j$ for the co-ordinates $\underset{\sim}{r}_j$, σ_j of the j'th electron. The Schrodinger equation for a system consisting of L' nuclei in fixed positions and N electrons can then be written

$$\mathcal{H}\,\Psi(\underline{s}_1, \underline{s}_2, \ldots \underline{s}_N) = E\,\Psi(\underline{s}_1, \underline{s}_2 \ldots \underline{s}_N) \qquad (1.2)$$

Let Z_λ be the charge on the λ^{th} nucleus, $\underline{R}_\lambda$ the position of this nucleus, and write

$$R_{\lambda\mu} = |\underline{R}_\lambda - \underline{R}_\mu|, \quad r_{\lambda i} = |\underline{R}_\lambda - \underline{r}_i|, \text{ and } r_{ij} = |\underline{r}_i - \underline{r}_j| \qquad (1.3)$$

Then, in the absence of any external fields and if the effect of electron-spin on the energy can be ignored, the N-electron Hamiltonian $\mathcal{H}$ involved in equation 1.2 can be written, in atomic units,

$$\mathcal{H} = \tfrac{1}{2} \sum_{\substack{\lambda,\mu=1 \\ \lambda\neq\mu}}^{L'} \frac{Z_\lambda Z_\mu}{R_{\lambda\mu}} - \sum_{i=1}^{N} \tfrac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{\lambda=1}^{L'} \frac{Z_\lambda}{r_{\lambda i}} + \sum_{\substack{i,j=1 \\ i\neq j}}^{N} \frac{1}{r_{ij}} \qquad (1.4)$$

We note that equation 1.2 will be satisfied by any function $\Psi(\underline{s}_1, \underline{s}_2 \ldots \underline{s}_N)$ which makes the expectation value of $\mathcal{H}$, viz. $\int \Psi^* \mathcal{H}\, \Psi\, (d\underline{s})^N$ stationary for arbitrary small variations in Ψ subject to the auxiliary condition $\int \Psi^* \Psi\, (d\underline{s})^N = N$. On this approach, the quantity E in equation 1.2 is the Lagrangian multiplier associated with the normalisation condition. In practice, it is frequently more convenient to use this variational approach than to attempt to solve equation 1.2 directly.

Density Matrices.

Consider the operator

$$\Lambda(1, 2, \ldots N) = \Lambda_0 + \sum_{i=1}^{N} \Lambda_i + \tfrac{1}{2!}\sum_{\substack{i,j=1 \\ i\neq j}}^{N} \Lambda_{ij} \qquad (1.5)$$

where Λ_0 is independent of the electrons, Λ_1 involves only electron i, and Λ_{ij} involves only electrons i and j. We note that $\mathcal{H}(\underline{s}_1, \underline{s}_2 \ldots \underline{s}_N)$ as defined by equation 1.4 is of this form, with

$$\Lambda_0 = \tfrac{1}{2} \sum_{\substack{\lambda,\mu=1 \\ \lambda\neq\mu}}^{L'} \frac{Z_\lambda Z_\mu}{R_{\lambda\mu}}, \quad \Lambda_i = -\tfrac{1}{2}\nabla_i^2 - \sum_{\lambda=1}^{L'} \frac{Z_\lambda}{r_{\lambda i}} \text{ and } \Lambda_{ij} = \frac{1}{r_{ij}} \qquad (1.6)$$

Assuming that $\Psi(\underline{s}_1, \underline{s}_2 \ldots \underline{s}_N)$ is a normalised N-electron wave function, we define the single-particle, or first order, density matrix

$$\gamma(\underline{s}_1'|\underline{s}_1) = N\int \Psi^*(\underline{s}_1', \underline{s}_2, \cdots \underline{s}_N)\,\Psi(\underline{s}_1, \underline{s}_2, \cdots \underline{s}_N)\,d\underline{s}_2\,d\underline{s}_3 \cdots d\underline{s}_N \qquad (1.7)$$

and the two-particle, or second order, density matrix

$$\Gamma(\underline{s}_1', \underline{s}_2'; \underline{s}_1, \underline{s}_2) = \frac{N(N-1)}{2!}\int \Psi^*(\underline{s}_1', \underline{s}_2', \underline{s}_3 \cdots \underline{s}_N)\,\Psi(\underline{s}_1, \underline{s}_2, \underline{s}_3 \cdots \underline{s}_N)\,d\underline{s}_3 d\underline{s}_4 \cdots d\underline{s}_N \qquad (1.8)$$

Then the expectation value $\langle \Lambda \rangle$ of the operator Λ is given by

$$\langle \Lambda \rangle \equiv \int \Psi^*(\underline{s}_1, \underline{s}_2, \cdots \underline{s}_N)\,\Lambda\,\Psi(\underline{s}_1, \underline{s}_2 \cdots \underline{s}_N)\,d\underline{s}_1 d\underline{s}_2 \cdots d\underline{s}_N$$

$$= \Lambda_0 + \int [\Lambda_1 \gamma(\underline{s}_1'|\underline{s}_1)]_{\underline{s}_1'=\underline{s}_1} d\underline{s}_1 + \iint \Lambda_{12}\,\Gamma(\underline{s}_1, \underline{s}_2; \underline{s}_1, \underline{s}_2)\,d\underline{s}_1 d\underline{s}_2 \qquad (1.9)$$

provided that, as in our case, Λ_{12} does not involve any differential operators.

In fact, all our knowledge about a system depends only on the density matrices, and the wave function tells us more than we need to, or can, know about the system, so that it is partly redundant. It is not physically possible to distinguish between two systems having different wave functions but identical density matrices of all orders. Unfortunately, it is not easy to calculate ab initio the density matrices for a many-electron system, since it is difficult to determine what types of matrices are compatible with the anti-symmetry of the electrons. For this reason, it is usual to concentrate on trying to calculate the many-electron wave function $\Psi(\underline{s}_1, \underline{s}_2 \cdots \underline{s}_N)$.

Slater Determinants.

For systems containing more than two or three electrons, it is usually necessary to assume that the wave function $\Psi(\underline{s}_1, \underline{s}_2 \cdots \underline{s}_N)$ is either separable or can be expressed as the sum of a number of separable wave functions. It is not strictly necessary to assume that one-electron wave functions can be defined, but it is most convenient to do so, and also

to assume that Ψ is adequately represented by the Slater determinant formed from N such wave functions in the manner described below. This assumption is used throughout the remainder of this thesis; one consequence of it is that any possible correlation between the positions of different electrons is ignored.

Let us define the complete anti-symmetric operator $\mathcal{A}$ for N particles by

$$\mathcal{A} = \left(\frac{1}{N!}\right)^{\frac{1}{2}} \sum_P \varepsilon_P P \qquad (1.10)$$

where $\varepsilon_P = \pm 1$ according as P is an even or odd permutation of the numbers 1, 2 ..... N and the sum is taken over all elements P of the permutation group on N numbers. Then we write

$$\Psi(\underline{s}_1, \underline{s}_2, \ldots \underline{s}_N) = \mathcal{A}\, \psi_1(\underline{s}_1)\psi_2(\underline{s}_2) \ldots \psi_N(\underline{s}_N)$$

$$= \left(\frac{1}{N!}\right)^{\frac{1}{2}} \sum_P \varepsilon_P \psi_1(\underline{s}_{P1})\psi_2(\underline{s}_{P2}) \ldots \psi_N(\underline{s}_{PN})$$

$$= \left(\frac{1}{N!}\right)^{\frac{1}{2}} \begin{vmatrix} \psi_1(\underline{s}_1) & \psi_2(\underline{s}_1) & \cdots & \psi_N(\underline{s}_1) \\ \psi_1(\underline{s}_2) & & & \\ \vdots & & & \\ \psi_1(\underline{s}_N) & \psi_2(\underline{s}_N) & \cdots & \psi_N(\underline{s}_N) \end{vmatrix} \qquad (1.11)$$

for the Slater determinant of the wave functions $\psi_1, \psi_2, \ldots \psi_N$. It is to be noted that P is understood to permute the electron co-ordinater $\underline{s}_j$ and not the wave functions ψ_j. If

$$\int \psi_i^*(\underline{s})\, \psi_j(\underline{s})\, d\underline{s} = \delta_{ij} \qquad (1.12)$$

i.e. if the ψ_j form an orthonormal set of functions, then $\Psi(\underline{s}_1, \underline{s}_2 \ldots \underline{s}_N)$ is normalised as required.

Consider now the 1 x N vector

$$\underline{\psi}(\underline{s}) = \left(\psi_1(\underline{s}), \psi_2(\underline{s}) \ldots \psi_N(\underline{s})\right) \qquad (1.13)$$

If we denote by A^+ the transpose of the matrix A and by $\tilde{A}$ the complex conjugate of the transpose of A, we can write equation 1.12 as

$$\int \underline{\Psi}^*(\underline{s})\, \underline{\Psi}^+(\underline{s})\, d\underline{s} = \underline{I}, \text{ the } N \times N \text{ identity matrix} \qquad (1.14)$$

Let us now choose a new set of functions $\psi_j'(\underline{s})$ which are linear combinations of the $\psi_i(\underline{s})$, so that

$$\psi_j'(\underline{s}) = \sum_i \psi_i(\underline{s})\, A_{ij} \qquad (1.15)$$

or

$$\underline{\Psi}'(\underline{s}) = \underline{\Psi}(\underline{s})\, A \qquad (1.16)$$

where, of course, the matrix A is independent of $\underline{s}$. Then since

$$\int \underline{\Psi}'^*(\underline{s})\, \underline{\Psi}'^+(\underline{s})\, d\underline{s} = \int \underline{\Psi}^*(\underline{s})\, A^* A^+\, \underline{\Psi}^+(\underline{s})\, d\underline{s} \qquad (1.17)$$

it follows from equation 1.14 that the functions ψ_j' are orthonormal if and only if

$$A^* A^+ = I \qquad (1.18)$$

i.e. A is a unitary matrix. In such a case, $|A| = 1$, and so

$$\underline{\Psi}'(\underline{s}_1, \underline{s}_2 \cdots \underline{s}_N) = \left(\frac{1}{N!}\right)^{\frac{1}{2}} \begin{vmatrix} \underline{\Psi}'(\underline{s}_1) \\ \vdots \\ \underline{\Psi}'(\underline{s}_N) \end{vmatrix} = \left(\frac{1}{N!}\right)^{\frac{1}{2}} \begin{vmatrix} \underline{\Psi}(\underline{s}_1) \\ \vdots \\ \underline{\Psi}(\underline{s}_N) \end{vmatrix} |A| = \underline{\Psi}(\underline{s}_1, \underline{s}_2 \cdots \underline{s}_N) \qquad (1.19)$$

Thus, it makes no difference to the many-electron wave function, or to the density matrices, whether the $\psi_i(\underline{s})$ or the $\psi_j'(\underline{s})$ are regarded as the occupied orbitals and used in equation 1.11. In terms of the $\psi_i(\underline{s})$, we find that if Ψ is described by the Slater determinant of equation 1.11, then

$$\gamma(\underline{s}_1' \mid \underline{s}_1) = \frac{N}{N!} \int \left[\sum_P \varepsilon_P P \psi_1^*(\underline{s}_1') \psi_2^*(\underline{s}_2) \cdots \psi_N^*(\underline{s}_N)\right] \left[\sum_Q \varepsilon_Q Q \psi_1(\underline{s}_1) \cdots \psi_N(\underline{s}_N)\right] d\underline{s}_2 \cdots d\underline{s}_N$$

$$= \frac{N}{N!} \times (N-1)! \sum_j \psi_j^*(\underline{s}_1') \psi_j(\underline{s}_1)$$

$$= \sum_{j=1}^{N} \psi_j^*(\underline{s}_1') \psi_j(\underline{s}_1) \qquad (1.20)$$

and

$$\Gamma(\underline{\xi}_1', \underline{\xi}_2'; \underline{\xi}_1, \underline{\xi}_2) = \frac{N(N-1)}{2!}\sum_{ij}\Big[\psi_i^*(\underline{\xi}_1')\,\psi_j^*(\underline{\xi}_2')\,\psi_i(\underline{\xi}_1)\psi_j(\underline{\xi}_2) - \psi_i^*(\underline{\xi}_1')\,\psi_j^*(\underline{\xi}_2')\,\psi_i(\underline{\xi}_2)\psi_j(\underline{\xi}_1)$$
$$- \psi_i^*(\underline{\xi}_2')\,\psi_j^*(\underline{\xi}_1')\,\psi_i(\underline{\xi}_1)\psi_j(\underline{\xi}_2) + \psi_i^*(\underline{\xi}_2')\psi_j^*(\underline{\xi}_1')\,\psi_i(\underline{\xi}_2)\psi_j(\underline{\xi}_1)\Big] \times (N-2)!$$

$$= \sum_{i,j=1}^{N}\left[\psi_i^*(\underline{\xi}_1')\psi_j^*(\underline{\xi}_2')\,\psi_i(\underline{\xi}_1)\psi_j(\underline{\xi}_2) - \psi_i^*(\underline{\xi}_1')\psi_j^*(\underline{\xi}_2')\,\psi_i(\underline{\xi}_2)\psi_j(\underline{\xi}_1)\right] \quad (1.21)$$

Hence, using equations 1.6, 1.9, 1.20, and 1.21, we find that

$$\langle \mathcal{H} \rangle = \tfrac{1}{2}\sum_{\lambda\neq\mu}\cdot\frac{Z_\lambda Z_\mu}{R_{\lambda\mu}} + \sum_{i=1}^{N}\int \psi_i^*(\underline{\xi})\,\mathcal{H}_1(\underline{\xi})\,\psi_i(\underline{\xi})\,d\underline{\xi}$$
$$+ \sum_{i,j=1}^{N}\iint \frac{1}{r_{12}}\left[\psi_i^*(\underline{\xi}_1)\psi_j^*(\underline{\xi}_2)\,\psi_j(\underline{\xi}_2)\psi_i(\underline{\xi}_1) - \psi_i^*(\underline{\xi}_1)\psi_j^*(\underline{\xi}_2)\psi_i(\underline{\xi}_2)\psi_j(\underline{\xi}_1)\right] d\underline{\xi}_1 d\underline{\xi}_2 \quad (1.2$$

where $\mathcal{H}_1(\underline{\xi}) = -\tfrac{1}{2}\nabla^2 - \sum_{\lambda=1}^{L}\frac{Z_\lambda}{|\underline{r}-\underline{R}_\lambda|}$ (1.23)

Furthermore, since equation 1.22 is derived from equation 1.9, and the latter involves only density matrices which are invariant under the unitary transformation $\underline{\psi}' = \underline{\psi}A$ if follows that equation 1.22 is still valid if we replace $\psi_i(\underline{\xi})$ by $\psi_i'(\underline{\xi})$ throughout it.

The Hartree-Fock Equations.

Let us write

$$\psi_i(\underline{\xi}) = \phi_i(\underline{r})\,\chi_i(\sigma) \quad (1.24)$$

where $\chi = \alpha$ or β is the spin-function; this form will generally be valid if the Hamiltonian $\mathcal{H}$ of the system does not involve the electron-spin. We restrict our attention to a closed-shell state, i.e. one in which all the space-orbitals $\phi_i(\underline{r})$ are

doubly occupied. In that case, on writing $d\underset{\sim}{s} = d\underset{\sim}{r}d\sigma$ in equation 1.22 and carrying out the integration over σ we find that

$$\langle \mathcal{H} \rangle = \tfrac{1}{2} \sum_{\lambda \neq \mu} \frac{Z_\lambda Z_\mu}{R_{\lambda\mu}} + 2 \sum_{i=1}^{\frac{1}{2}N} \int \phi_i^*(\underset{\sim}{r}_1)\, \mathcal{H}_1(\underset{\sim}{r}_1)\, \phi_i(\underset{\sim}{r}_1)\, d\underset{\sim}{r}_1 \qquad (1.25)$$

$$+ 2 \sum_{i,j=1}^{N} \iint \frac{\phi_i^*(\underset{\sim}{r}_1)\, \phi_j^*(\underset{\sim}{r}_2)\, \phi_j(\underset{\sim}{r}_2)\, \phi_i(\underset{\sim}{r}_1)}{r_{12}} d\underset{\sim}{r}_1 d\underset{\sim}{r}_2 - \sum_{i,j=1}^{\frac{1}{2}N} \iint \frac{\phi_i^*(\underset{\sim}{r}_1)\, \phi_j^*(\underset{\sim}{r}_2)\, \phi_i(\underset{\sim}{r}_2)\, \phi_j(\underset{\sim}{r}_1)}{r_{12}} d\underset{\sim}{r}_1 d\underset{\sim}{r}_2$$

since if ψ_i and ψ_j have opposite spin

$$\iint \chi_i^*(\sigma_1)\, \chi_j^*(\sigma_2)\, \chi_i(\sigma_2)\, \chi_j(\sigma_1)\, d\sigma_1 d\sigma_2 = 0 \qquad (1.26)$$

If we now write

$$\underline{\phi}' = \underline{\phi}\, A \qquad (1.27)$$

where A is henceforth an $\frac{1}{2}N \times \frac{1}{2}N$ unitary matrix, then since equation 1.22 applies with ϕ_j' everywhere replacing ϕ_j so will equation 1.25.

As mentioned previously, $\underline{\Psi}$ is a solution of the Schrodinger equation if it makes $\langle \mathcal{H} \rangle$ stationary subject to the normalisation condition. If we represent $\underline{\Psi}$ by a single Slater determinant, then for an arbitrary variation in it we can write

$$\phi_i \to \phi_i + \delta\phi_i \,, \quad \phi_i^* \to \phi_i^* + \delta\phi_i^* \,, \quad i = 1, 2, \ldots \tfrac{1}{2}N$$

while normalisation will be preserved if we still require that

$$\int \phi_i^*(\underset{\sim}{r})\, \phi_j(\underset{\sim}{r})\, d\underset{\sim}{r} = \delta_{ij} \qquad (1.12)$$

Thus, for an eigenfunction of $\mathcal{H}$ represented by a single Slater determinant and corresponding to a closed shell state, we require that

$$X = \langle \mathcal{H} \rangle - \sum_{i,j=1}^{\frac{1}{2}N} \varepsilon_{ij} \left[\int \phi_j^*(\underset{\sim}{r})\, \phi_i(\underset{\sim}{r})\, d\underset{\sim}{r} - \delta_{ij} \right] \qquad (1.28)$$

be stationary, where $\langle \mathcal{H} \rangle$ is given by equation 1.25 and the quantities ε_{ij} are undetermined Lagrangian multipliers.

Now

$$\delta X = 2\sum_i \left[\int \delta\phi_i^*(\underline{r}) H_i(\underline{r}) \phi_i(\underline{r}) d\underline{r} + \int \phi_i^*(\underline{r}) H_i(\underline{r}) \delta\phi_i(\underline{r}) d\underline{r} \right]$$

$$+ 4\sum_{i,j} \int \left[\delta\phi_i^*(\underline{r}) \phi_i(\underline{r}) + \delta\phi_i(\underline{r}) \phi_i^*(\underline{r}) \right] \left[\int \frac{\phi_j^*(\underline{r}') \phi_j(\underline{r}')}{|\underline{r}-\underline{r}'|} d\underline{r}' \right] d\underline{r}$$

$$- 2\sum_{i,j} \left[\iint \frac{\delta\phi_i^*(\underline{r}) \phi_j^*(\underline{r}') \phi_i(\underline{r}') \phi_j(\underline{r})}{|\underline{r}-\underline{r}'|} d\underline{r}' d\underline{r} + \iint \frac{\delta\phi_i(\underline{r}) \phi_j^*(\underline{r}) \phi_i^*(\underline{r}') \phi_j(\underline{r}')}{|\underline{r}-\underline{r}'|} d\underline{r} d\underline{r}' \right]$$

$$- 2\sum_{i,j} \varepsilon_{ji} \left[\int \delta\phi_i^*(\underline{r}) \phi_j(\underline{r}) d\underline{r} + \int \delta\phi_i(\underline{r}) \phi_j^*(\underline{r}) d\underline{r} \right]. \qquad (1.29)$$

Hence, equating the coefficients of $\delta\phi_i^*$ and $\delta\phi_i$ to zero and making use of the fact that $H_i(\underline{r})$ is Hermitian, we find that

$$H_i(\underline{r}) \phi_i(\underline{r}) + 2\sum_j \int \frac{\phi_j^*(\underline{r}') \phi_j(\underline{r}')}{|\underline{r}-\underline{r}'|} d\underline{r}' \phi_i(\underline{r}) - \sum_j \int \frac{\phi_j^*(\underline{r}') \phi_i(\underline{r})}{|\underline{r}-\underline{r}'|} d\underline{r}' \phi_j(\underline{r})$$

$$= \sum_{j=1}^{\frac{1}{2}N} \varepsilon_{ji} \phi_j(\underline{r}) \qquad (1.30)$$

and $H_i^*(\underline{r}) \phi_i^*(\underline{r}) + 2\sum_j \int \frac{\phi_j^*(\underline{r}') \phi_j(\underline{r}')}{|\underline{r}-\underline{r}'|} d\underline{r}' \phi_i^*(\underline{r}) - \sum_j \int \frac{\phi_j(\underline{r}') \phi_i^*(\underline{r}')}{|\underline{r}-\underline{r}'|} d\underline{r}' \phi_j^*(\underline{r})$

$$= \sum_{j=1}^{\frac{1}{2}N} \varepsilon_{ij}^* \phi_j^*(\underline{r}) \qquad (1.31)$$

Hence $\varepsilon_{ij} = \varepsilon_{ji}^*$ (1.32)

and $F\phi_i(\underline{r}) = \sum_{j=1}^{\frac{1}{2}N} \phi_j(\underline{r}) \varepsilon_{ji}, \quad i = 1, 2, \ldots\ldots\ldots \tfrac{1}{2}N$ (1.33)

where

$$F\phi_i(\underline{r}) = \left[-\tfrac{1}{2}\nabla^2 - \sum_{\lambda=1}^{L} \frac{Z_\lambda}{|\underline{r}-\underline{R}_\lambda|} + 2\sum_{j=1}^{\frac{1}{2}N} \int \frac{\phi_j^*(\underline{r}') \phi_j(\underline{r}')}{|\underline{r}-\underline{r}'|} d\underline{r}' \right] \phi_i(\underline{r})$$

$$- \sum_{j=1}^{\frac{1}{2}N} \int \frac{\phi_j^*(\underline{r}') \phi_i(\underline{r}')}{|\underline{r}-\underline{r}'|} d\underline{r}' \phi_j(\underline{r}) \qquad (1.34)$$

These are the Hartree-Fock equations.

Since the derivation of these equations depends only on equation 1.25, which was shown to apply if ϕ_j' was everywhere substituted for ϕ_j these equations also hold under such conditions, i.e.

$$F' \phi_i'(\underline{r}) = \sum_{j=1}^{\frac{1}{2}N} \phi_j'(\underline{r})\, \varepsilon_{ji}' \qquad (1.35)$$

The relationships between F' and F and between ε_{ji}' and ε_{ji} can most easily be derived using a matrix notation. Equation 1.34 can be written

$$F \underline{\Phi}(\underline{r}) = \left[-\tfrac{1}{2}\nabla^2 - \sum_{\lambda=1}^{L'} \frac{Z_\lambda}{|\underline{r}-\underline{R}_\lambda|} + 2\int \frac{1}{|\underline{r}-\underline{r}'|}\, \underline{\Phi}^*(\underline{r}')\, \underline{\Phi}^+(\underline{r}')\, d\underline{r}'\right] \underline{\Phi}(\underline{r}) \qquad (1.36)$$

$$- \int \frac{1}{|\underline{r}-\underline{r}'|}\, \underline{\Phi}^*(\underline{r}')\, \underline{\Phi}^+(\underline{r})\, \underline{\Phi}(\underline{r}')\, d\underline{r}'$$

so that

$$F' \underline{\Phi}'(\underline{r}) = \left[-\tfrac{1}{2}\nabla^2 - \sum_{\lambda=1}^{L'} \frac{Z_\lambda}{|\underline{r}-\underline{R}_\lambda|} + 2\int \frac{1}{|\underline{r}-\underline{r}'|}\, \underline{\Phi}^*(\underline{r}')\, A^*\, A^+\, \underline{\Phi}^+(\underline{r}')\, d\underline{r}'\right] \underline{\Phi}(\underline{r})\, A$$

$$- \int \frac{1}{|\underline{r}-\underline{r}'|}\, \underline{\Phi}^*(\underline{r}')\, A^*\, A^+\, \underline{\Phi}^+(\underline{r})\, \underline{\Phi}(\underline{r}')\, A\, d\underline{r}' \qquad (1.37)$$

$$= F \underline{\Phi}(\underline{r})\, A$$

Thus, multiplying on the right by $A^{-1} = \tilde{A}$ we find that

$$F' \underline{\Phi}(\underline{r}) = F \underline{\Phi}(\underline{r}) \qquad (1.38)$$

Also, since equation 1.33 can be written

$$F \underline{\Phi} = \underline{\Phi}\, \underline{\underline{\varepsilon}} \qquad (1.39)$$

we find that

$$F' \underline{\Phi}' = F \underline{\Phi}' = F \underline{\Phi} A = \underline{\Phi}\, \underline{\underline{\varepsilon}} A = \underline{\Phi} A \tilde{A}\, \underline{\underline{\varepsilon}} A = \underline{\Phi}'(\tilde{A}\, \underline{\underline{\varepsilon}}\, A) \qquad (1.40)$$

Molecular Orbitals.

Since in equation 1.39 the operator F is Hermitian, all its eigenvalues are real and eigenvectors associated with different eigenvalues are mutually orthogonal. In addition, mutually orthogonal linear combinations of any degenerate

eigenvectors can be chosen. Hence if $\underline{\phi}$ satisfies equation 1.39, it is possible to find a unitary matrix A such that $\tilde{A} \underline{\underline{\varepsilon}} A$ is a diagonal matrix and the functions ϕ_j' defined by $\underline{\phi}' = \underline{\phi} A$ are mutually orthogonal eigenfunctions of F. These functions ϕ_j' are called molecular orbitals.

One advantage of molecular orbitals is that if any given orbital ϕ_i is being considered, the other occupied orbitals ϕ_j only enter the equation which ϕ_i must satisfy through the operator F, since $\underline{\underline{\varepsilon}}$ is diagonal. It may be possible to approximate quite well to F without first knowing exactly all the functions ϕ_j, while if $\underline{\underline{\varepsilon}}$ were not diagonal it would not in general be possible even to estimate the right hand side of equation 1.33 without knowing all the other orbitals. This is why it is generally considered simplest to calculate the molecular orbitals for a system. The remainder of this chapter will be concerned with such calculations for crystals. However, it will be shown in the next chapter that under certain rather restrictive conditions, such as are believed to hold for the ground states of tetrahedral semiconductors, it may be both practicable and more convenient to choose a different matrix, A and calculate the equivalent orbitals.

Another advantage of molecular orbitals is that, according to Koopman's theorem, if one electron is removed from a molecular orbital while all the other orbitals remain unchanged, the ionisation potential, i.e. the energy required, is the corresponding ε_{ii} . It follows that if one electron goes from a molecular orbital ϕ_j to the previously unoccupied one ϕ_i the energy required will be $\varepsilon_{ii} - \varepsilon_{jj}$. This is often a fairly good approximation, so that it is possible to compare the energies calculated using molecular orbitals with those observed experimentally. The eigenvalues of F are known as the one-electron energies.

Ideally, one would like to be able to find self-consistent solutions of the Hartree-Fock equations, i.e. solutions in which the same functions appear in equation 1.34 (defining F) as in equation 1.33. However, for a general system containing N electrons numerical solutions for 3-dimensional wave functions are not practicable, and so unless spherical symmetry, as in atoms, or at least axial symmetry, as in diatomic molecules, permits parts of the wave functions to be expressed exactly in an analytical form, it is necessary to use approximate methods. In the commonest of these, Roothaan's self-consistent field method, [2] an attempt is made to expand all the molecular orbitals in terms of a finite number of members of a complete basis set of analytic functions. For a general system, a vast number of such functions will be required if N is large, and even if F is known approximately the order of the secular equations which will have to be solved to find the molecular orbitals is prohibitively large. If, however, the system has a certain symmetry, e.g. if it is invariant under the group G of symmetry operators, then the molecular orbitals can be classified according to their behaviour under the various elements of G, and the order of the secular equations for a given type of orbital is greatly reduced. It is for this reason that calculations can most easily be carried out on atoms, very symmetric molecules, and crystals. In the next section, we examine briefly how use can be made of symmetry in calculations on crystals.

1.2 CRYSTAL SYMMETRY AND GROUP THEORY.

Space Groups and the Crystal Lattice.

The feature which distinguishes a crystal from other forms of matter is its translational symmetry. In any

three-dimensional crystal, it is possible to find three non-coplanar vectors $\underset{\sim}{a}_1$, $\underset{\sim}{a}_2$, and $\underset{\sim}{a}_3$ such that the placing of a certain block, known as a unit cell, in corresponding positions at each of the lattice points

$$\underset{\sim}{n} = n_1\underset{\sim}{a}_1 + n_2\underset{\sim}{a}_2 + n_3\underset{\sim}{a}_3, \qquad n_j = 0, \pm 1, \pm 2, \ldots . \tag{1.41}$$

will produce a crystal indistinguishable from that being considered. The vectors $\underset{\sim}{n}$ will be called, for short, lattice vectors.

That the choice of unit cell is not unique can be seen from considering the zincblende structure. Taking a right-handed orthogonal set of co-ordinate axes, so oriented that the join of one pair of adjacent atoms is in the (1,1,1) direction, we find that

$$\underset{\sim}{a}_1 = (2L,0,0)\ , \quad \underset{\sim}{a}_2 = (0,2L,0) \quad \text{and} \quad \underset{\sim}{a}_3 = (0,0,2L) \tag{1.42}$$

provides a translation lattice for the crystal, and so does the set

$$\underset{\sim}{a}_1 = (0,L,L)\ , \quad \underset{\sim}{a}_2 = (L,0,L) \quad \text{and} \quad \underset{\sim}{a}_3 = (L,L,0). \tag{1.43}$$

Since the volume Ω of a unit cell is given by the formula

$$\Omega = |\underset{\sim}{a}_1 . \underset{\sim}{a}_2 \times \underset{\sim}{a}_3| = \mathrm{mod} \begin{vmatrix} a_{1x} & a_{1y} & a_{1z} \\ a_{2x} & a_{2y} & a_{2z} \\ a_{3x} & a_{3y} & a_{3z} \end{vmatrix} \tag{1.44}$$

the first of these lattices has a unit cell of volume four times greater than that of the second lattice. Although for some purposes, such as crystallography, the first lattice may be more convenient since it is cubic, the second one, with basic translation vectors given by equation 1.43, is more convenient for our calculations and will be used throughout this thesis.

Even when the lattice has been determined, the unit cell can still be chosen in a variety of ways. For instance, it can be taken to be the parallelepiped with sides $\underset{\sim}{a}_1$, $\underset{\sim}{a}_2$, and $\underset{\sim}{a}_3$, and such a choice is very convenient for a simple cubic lattice. However, it is usually most convenient to choose the symmetric unit cell, also known as the Wigner-Seitz cell, which is the cell bounded by the planes bisecting the joins of a given lattice point to its nearest neighbours. Such a cell contains all points nearer to the given lattice point than to any other lattice point, and has the full symmetry of the crystal. For the zincblende structure, which has a face-centred cubic lattice, the symmetric unit cell is a rhombododecahedron. Fortunately, all calculations can be performed by considering only the cube contained inside it, which is shown in figure I; this cube is sometimes erroneously stated to be the unit cell of the structure, but we prefer to call it the basic cell or unit cube.

In addition to possessing translational symmetry, a crystal may also be invariant under symmetry operations such as rotations, inversions and reflections. The group of all symmetry operations which leave a crystal invariant is known as the space group G of the crystal.[3] It is convenient to denote the symmetry operation consisting of a rotation α followed by a translation $\underset{\sim}{t}$ by the symbol $(\alpha|\underset{\sim}{t})$. $(\varepsilon|\underset{\sim}{t})$ is a pure translation, and $(\alpha|0)$ a pure rotation. If $(\alpha|\underset{\sim}{t})$ and $(\beta|\underset{\sim}{s})$ are two symmetry operators, it follows from our definitions that their product

$$(\alpha|\underset{\sim}{t})(\beta|\underset{\sim}{s}) = (\alpha\beta|\underset{\sim}{t}+\alpha\underset{\sim}{s}) \qquad (1.45)$$

and so that

$$(\alpha|\underset{\sim}{t})^{-1} = (\alpha^{-1}|-\alpha^{-1}\underset{\sim}{t}). \qquad (1.46)$$

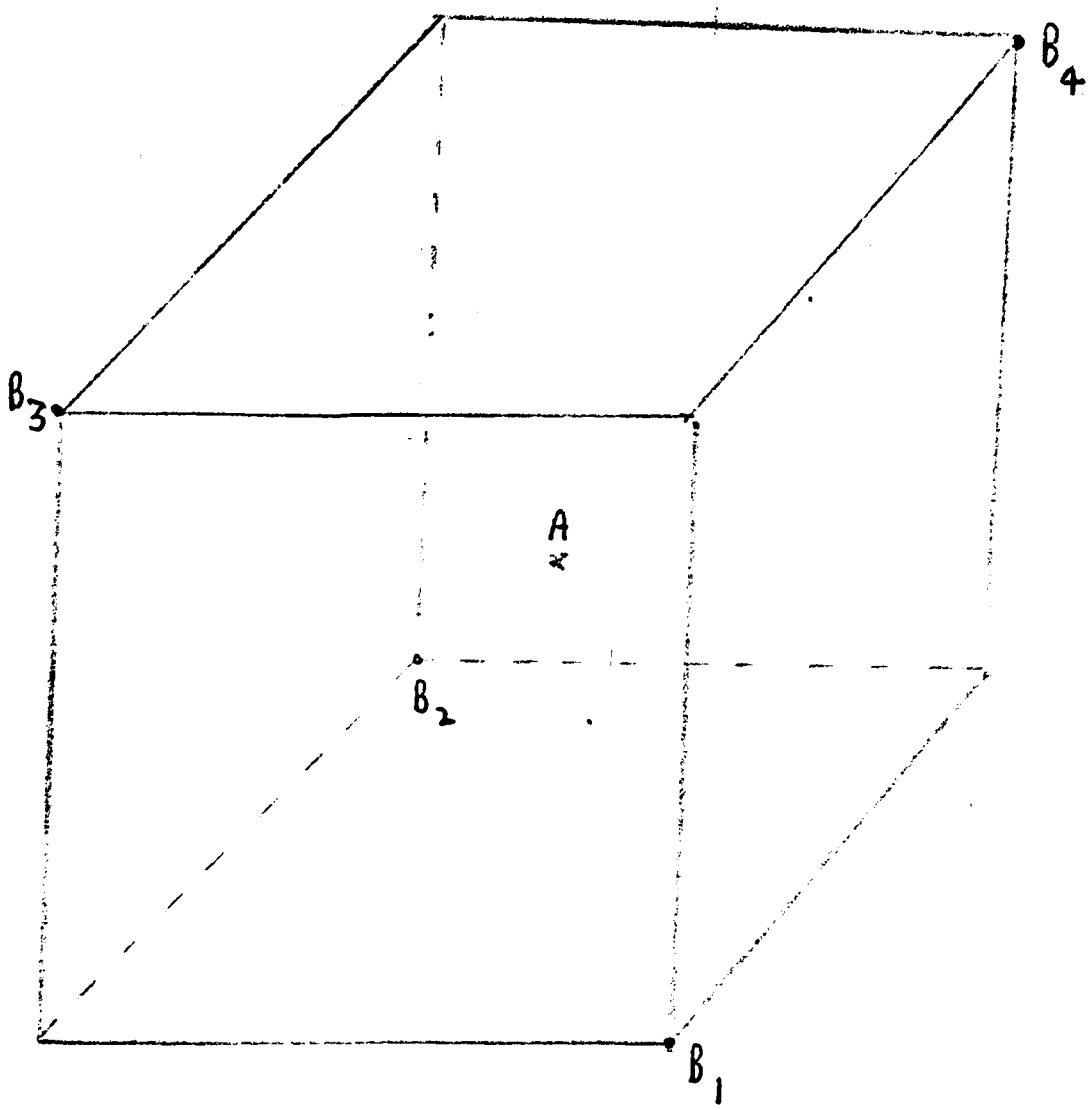


Figure 1. The unit cube of the Zinchlende structure.

We choose to regard our operators as acting on the co-ordinate axes, so that for a function $\phi(\underset{\sim}{r})$ of the electron co-ordinates

$$(\alpha|\underset{\sim}{t})^{-1}\ \phi(\underset{\sim}{r}) = \phi(\alpha\underset{\sim}{r} + \underset{\sim}{t})\ ; \qquad (1.47)$$

the nuclei are assumed to be fixed in the axes.

The set of translations $(\varepsilon|\underset{\sim}{n})$ where $\underset{\sim}{n}$ is a lattice vector, form a subgroup of the space group known as the translational group T. It is an invariant subgroup, since

$$(\alpha^{-1}|-\alpha^{-1}\underset{\sim}{t})(\varepsilon|\underset{\sim}{n})(\alpha|\underset{\sim}{t}) = (\varepsilon|\alpha^{-1}\underset{\sim}{n}) = (\varepsilon|\underset{\sim}{m}) \qquad (1.48)$$

where $\underset{\sim}{m}$ is another lattice vector. The set of rotations $(\alpha|0)$ belonging to a given space group G form a subgroup which we call the point group P of G; P is isomorphic to a rotation group of finite order. It is often convenient to write $\alpha\phi$ for $(\alpha|0)\,\phi$ just as we write $\alpha\underset{\sim}{r}$ for $(\alpha|0)\,\underset{\sim}{r}$,

A space group is said to be symmorphic if every element in it can be expressed uniquely in the form $(\varepsilon|\underset{\sim}{n})(\alpha|0)$, where $(\alpha|0)$ is in P and $(\varepsilon|\underset{\sim}{n})$ is in T. A non-symmorphic space group contains elements of the form $(\beta|\underset{\sim}{v})$ where $\underset{\sim}{v}$ is not a lattice vector. The zincblende structure has a symmorphic space group, the point group of which is the tetrahedral group T_d containing 24 rotations. The diamond structure has a non-symmorphic space group, since it has inversion symmetry about the mid-point of a bond. However, a sub-group of its space group is the symmorphic space group of the zincblende structure, and we found it convenient for most purposes to use only this subgroup. As a consquence, we are able to restrict our attention to symmorphic space groups for the remainder of this thesis.

Irreducible Representations of Space Groups.

In order to restrict our attention to finite space groups,

we impose cyclic boundary conditions on the crystal, i.e. we assume that after M_i translations through the basic lattice vector $\underset{\sim}{a}_i$ we are back at the place from which we started. This condition can be written in the form

$$(\varepsilon|\underset{\sim}{a}_1)^{M_1} = (\varepsilon|\underset{\sim}{a}_2)^{M_2} = (\varepsilon|\underset{\sim}{a}_3)^{M_3} = (\varepsilon|0) \tag{1.49}$$

Let us define three new vectors $\underset{\sim}{T}_i$ such that

$$\underset{\sim}{T}_i \cdot \underset{\sim}{a}_j = 2\pi\delta_{ij}, \quad i,j = 1,2,3. \tag{1.50}$$

It follows from this definition that

$$\underset{\sim}{T}_i = \frac{2\pi}{\Lambda}\, \underset{\sim}{a}_j \times \underset{\sim}{a}_k \tag{1.51}$$

where (ijk) is a cyclic permutation of the numbers (1,2,3)
For the zincblende structure,

$$\underset{\sim}{a}_1 = \frac{2R}{\sqrt{3}}(0,1,1), \quad \underset{\sim}{a}_2 = \frac{2R}{\sqrt{3}}(1,0,1), \quad \underset{\sim}{a}_3 = \frac{2R}{\sqrt{3}}(1,1,0) \tag{1.52}$$

where R is the bond strength, i.e. the distance between a pair of adjacent atoms, and so

$$\Lambda = \frac{16R^3}{3\sqrt{3}} \tag{1.53}$$

and

$$\underset{\sim}{T}_1 = \frac{\pi\sqrt{3}}{2R}(-1,1,1), \quad \underset{\sim}{T}_2 = \frac{\pi\sqrt{3}}{2R}(1,-1,1), \quad \underset{\sim}{T}_3 = \frac{\pi\sqrt{3}}{2R}(1,1,-1) \tag{1.54}$$

The vectors $\underset{\sim}{T}_j$ form the basis of the reciprocal lattice of the crystal; the reciprocal lattice vectors are of the form

$$\underset{\sim}{K} = K_1\underset{\sim}{T}_1 + K_2\underset{\sim}{T}_2 + K_3\underset{\sim}{T}_3, \quad K_j = 0, \pm 1, \pm 2, \ldots \tag{1.55}$$

The main advantage of finding the irreducible representations of a group is the fact that functions belonging to different representations are orthogonal, and the matrix element between them of any operator invariant under the group is zero. Thus, if H is invariant under G and ϕ_1 and ϕ_2 belong to different irreducible representations of G, then

$$\langle\phi_1|H|\phi_2\rangle = 0, \quad \langle\phi_1|\phi_2\rangle = 0. \tag{1.56}$$

If ϕ_1 and ϕ_2 belong to the same irreducible representation of G but to different irreducible representations of some other group G' under which H is not invariant, then $\langle \phi_1 | \phi_2 \rangle = 0$ but $\langle \phi_1 | H | \phi_2 \rangle$ is not necessarily zero.

The property expressed in equation 1.56 enables us to reduce the order of the secular equations which have to be solved in order to find a molecular orbital using Roothaan's self-consistent field method. Thus in a crystal containing M unit cells, where M may be of order 10^{22}, molecular orbital calculations would be quite impracticable were it not for the existence of M different one-dimensional representations of the translation group T. These are given by the correspondence

$$(\varepsilon | \underset{\sim}{n}) \rightarrow e^{-i\underset{\sim}{k}\cdot\underset{\sim}{n}} \qquad (1.57)$$

where

$$\underset{\sim}{k} = k_1 \underset{\sim}{T}_1 + k_2 \underset{\sim}{T}_2 + k_3 \underset{\sim}{T}_3, \quad k_j = S_j / M_j, \quad S_j = 0, \pm 1, \pm 2 \ldots\ldots \pm (M_j - 1), \pm M_j. \qquad (1.58)$$

We have chosen the range of the S_j so that $\underset{\sim}{k}$ lies within a single unit cell of the reciprocal lattice.

If $\underset{\sim}{k}$ and $\underset{\sim}{k}'$ differ by a reciprocal lattice vector $\underset{\sim}{K}$, then for any lattice vector $\underset{\sim}{n}$ $e^{-i\underset{\sim}{k}\cdot\underset{\sim}{n}} = e^{-i\underset{\sim}{k}'\cdot\underset{\sim}{n}}$ so that $(\varepsilon | \underset{\sim}{n}) \rightarrow e^{-i\underset{\sim}{k}\cdot\underset{\sim}{n}}$ and $(\varepsilon | \underset{\sim}{n}) \rightarrow e^{-i\underset{\sim}{k}'\cdot\underset{\sim}{n}}$ are then the same representation of the translation group; in such a case $\underset{\sim}{k}$ and $\underset{\sim}{k}'$ are said to be equivalent. Hence, if we define a sub-lattice of the reciprocal lattice with basis vectors $\underset{\sim}{T}_1/M_1$, $\underset{\sim}{T}_2/M_2$, and $\underset{\sim}{T}_3/M_3$, those values of $\underset{\sim}{k}$ corresponding to all the points of the sub-lattice lying within any given unit cell of the reciprocal lattice will provide all the distinct irreducible represenations of the translation group T.

Equation 1.58 indicates one such set of sub-lattice points. Frequently it is convenient to restrict $\underset{\sim}{k}$ to the symmetric unit cell of the reciprocal lattice, which corresponds to the Wigner-Seitz cell in configuration space. This cell is known as the first Brillouin zone, and has the full symmetry of the point group P of the crystal. Unless otherwise stated, it will be assumed henceforth that $\underset{\sim}{k}$ is restricted to points lying within and on the surface of the first Brillouin zone. In practice, we let the M_j tend to infinity, and regard $\underset{\sim}{k}$ as varying continuously within this zone, but for theoretical purposes it is more convenient to keep the M_j finite so as to have a space group of finite order and be able to use sums instead of integrals.

From equation 1.57, it follows that a function belonging to a representation $\underset{\sim}{k}$ of the translation group satisfies the equation

$$(\varepsilon|\underset{\sim}{n})\,\phi(\underset{\sim}{r}) = e^{-i\underset{\sim}{k}\cdot\underset{\sim}{n}}\,\phi(\underset{\sim}{r}) \qquad (1.58)$$

i.e.

$$\phi(\underset{\sim}{r}-\underset{\sim}{n}) = e^{-i\underset{\sim}{k}\cdot\underset{\sim}{n}}\,\phi(\underset{\sim}{r}) \qquad (1.59)$$

Let

$$u_{\underset{\sim}{k}}(\underset{\sim}{r}) = e^{-i\underset{\sim}{k}\cdot\underset{\sim}{r}}\,\phi(\underset{\sim}{r}) \qquad (1.60)$$

Then

$$u_{\underset{\sim}{k}}(\underset{\sim}{r}-\underset{\sim}{n}) = e^{-i\underset{\sim}{k}\cdot(\underset{\sim}{r}-\underset{\sim}{n})}\,\phi(\underset{\sim}{r}-\underset{\sim}{n}) = e^{-i\underset{\sim}{k}\cdot\underset{\sim}{r}}\,\phi(\underset{\sim}{r}) = u_{\underset{\sim}{k}}(\underset{\sim}{r}) \qquad (1.61)$$

Hence

$$\phi(\underset{\sim}{r}) = e^{i\underset{\sim}{k}\cdot\underset{\sim}{r}}\,u_{\underset{\sim}{k}}(\underset{\sim}{r}) \qquad (1.62)$$

where $u_{\underset{\sim}{k}}(\underset{\sim}{r})$ is invariant under the translation group T of the lattice. Since it is possible so to choose the molecular orbitals

that each of them belongs to some irreducible representation of T, it follows that the molecular orbitals can be written in the form 1.62, which is just Bloch's theorem. We note that for any function $\phi(\underset{\sim}{r})$ the Bloch sum $\sum_{\underset{\sim}{\alpha}} e^{i\underset{\sim}{k}\cdot\underset{\sim}{\alpha}}(\alpha|\underset{\sim}{\alpha})\,\phi(\underset{\sim}{r})$ either vanishes identically or belongs to representation $\underset{\sim}{k}$ of T; Bloch sums are also known as Bloch waves.

Consider now the set $P_{\underset{\sim}{k}}$ of all elements of P which leave $\underset{\sim}{k}$ invariant i.e. the set $(\alpha;\ \alpha \in P,\ \alpha\underset{\sim}{k} = \underset{\sim}{k})$. $P_{\underset{\sim}{k}}$ is a subgroup of P, and we call it the little group of $\underset{\sim}{k}$. Let $S_{\underset{\sim}{k}}$ be the set of all vectors of the form $(\beta\underset{\sim}{k},\ \beta \in P)$; we call $S_{\underset{\sim}{k}}$ the star of $\underset{\sim}{k}$. Then, as shown by Redei[4] if one considers a basis ϕ_λ for an irreducible representation of $P_{\underset{\sim}{k}}$ and forms from it all the possible Bloch sums with $\underset{\sim}{k}'$ in $S_{\underset{\sim}{k}}$, the vector space spanned by these sums is invariant and irreducible under the space group G; this result is very useful for L.C.A.O. calculations. We note that the matrix element of any operator H which is invariant under G between two Bloch waves associated with non-equivalent $\underset{\sim}{k}'$ in $S_{\underset{\sim}{k}}$ must, of course, be zero.

Points on the Surface of the Brillouin zone.

For certain points of high symmetry on the surface of the first Brillouin zone, there exist elements $(\alpha|0)$ of P such that

$$\alpha\underset{\sim}{k} - \underset{\sim}{k} = \underset{\sim}{K}$$

a non-zero reciprocal lattice vector. Since in such a case $\underset{\sim}{k}$ and $\alpha\underset{\sim}{k}$ are equivalent, it is convenient to introduce for these points an augmented little group $P'_{\underset{\sim}{k}}$ consisting of all the operators α in P such that $\underset{\sim}{k}$ and $\alpha\underset{\sim}{k}$ are equivalent; the corresponding star of $\underset{\sim}{k}$ will be denoted by $S'_{\underset{\sim}{k}}$, and consists of all the vectors of form $(\beta\underset{\sim}{k},\ \beta \in P)$ which are not equivalent to $\underset{\sim}{k}$. It is tempting to apply Redei's result

to the augmented little group $P'_{\underset{\sim}{k}}$, but we now proceed to show that great care is needed in doing so.(5)

For a given $\underset{\sim}{k}$, consider a set of functions $\phi^{lm}(\underset{\sim}{r})$ forming a basis for the irreducible representations of $P'_{\underset{\sim}{k}}$ and such that $\phi^{ji}(\underset{\sim}{r})$ is the i^{th} element of the j^{th} irreducible representation of $P'_{\underset{\sim}{k}}$; then we can write

$$(\alpha|0)\,\phi^{ji}(\underset{\sim}{r}) = \sum_l \alpha^j_{li}\,\phi^{jl}(\underset{\sim}{r}) \qquad (1.63)$$

Putting $\underset{\sim}{m} = \alpha\underset{\sim}{n}$, we find that

$$(\alpha|0)\,(\varepsilon|\underset{\sim}{n})\,\phi^{ji}(\underset{\sim}{r}) = \sum_l \alpha^j_{li}(\varepsilon|\underset{\sim}{m})\,\phi^{jl}(\underset{\sim}{r}) \qquad (1.64)$$

Let

$$\psi^{ji}_{\underset{\sim}{k}}(\underset{\sim}{r}) = \sqrt{M}\sum_{\underset{\sim}{n}} e^{i\underset{\sim}{k}\cdot\underset{\sim}{n}}(\varepsilon|\underset{\sim}{n})\,\phi^{ji}(\underset{\sim}{r}) \qquad (1.65)$$

Then

$$\begin{aligned}(\alpha|0)\,\psi^{ji}_{\underset{\sim}{k}}(\underset{\sim}{r}) &= \frac{1}{\sqrt{M}}\sum_{\underset{\sim}{n}} e^{i\underset{\sim}{k}\cdot\underset{\sim}{n}}(\varepsilon|\underset{\sim}{m})\sum_l \alpha^j_{li}\,\phi^{jl}(\underset{\sim}{r}) \\ &= \frac{1}{\sqrt{M}}\sum_l \alpha^j_{li}\sum_{\underset{\sim}{m}} e^{i\underset{\sim}{k}\cdot(\alpha^{-1}\underset{\sim}{m})}(\varepsilon|\underset{\sim}{m})\,\phi^{jl}(\underset{\sim}{r}) \\ &= \sum_l \alpha^j_{li}\,\psi^{jl}_{\alpha\underset{\sim}{k}}(\underset{\sim}{r})\end{aligned}$$

Hence, if $\alpha \in P'_{\underset{\sim}{k}}$,

$$\text{then}\quad (\alpha|0)\,\psi^{ji}_{\underset{\sim}{k}}(\underset{\sim}{r}) = \sum_l \alpha^j_{li}\,\psi^{jl}_{\underset{\sim}{k}}(\underset{\sim}{r}) \qquad (1.66)$$

and so the functions $\psi^{ji}_{\underset{\sim}{k}}(\underset{\sim}{r})$ also form a basis for the irreducible representations of $P'_{\underline{k}}$, the augmented little group of $\underline{k}$.

However, a similar analysis shows that if

$$\chi^{ji}_{\underset{\sim}{k}}(\underset{\sim}{r}) = \frac{1}{\sqrt{M}}\sum_{\underset{\sim}{n}} e^{i\underset{\sim}{k}\cdot\underset{\sim}{n}}(\varepsilon|\underset{\sim}{n}+\underset{\sim}{a})\,\psi^{ji}(\underset{\sim}{r}) \qquad (1.67)$$

where $\underset{\sim}{a}$ is not a lattice vector so that $e^{i\underset{\sim}{K}.\underset{\sim}{a}} \neq 1$ for every receiprocal lattice vector $\underset{\sim}{K}$, then the above result holds for $P_{\underset{\sim}{k}}$ but not for the augmented little group $P'_{\underset{\sim}{k}}$. These groups will only differ for points on the surface of the first Brillouin zone, but at such points considerable care is needed if a localised orbital method is used and the crystal contains more than one atom per unit cell. The importance of these points lies in the fact that extra information can be obtained by using $P'_{\underset{\sim}{k}}$ instead of $P_{\underset{\sim}{k}}$, and this makes it easier to calculate the molecular orbitals at these points than elsewhere in reciprocal space.[40]

The One-Election Hamiltonian.

Consider a system, invariant under the group G of symmetry operators, which has a non-degenerate ground state. The many-electron wave function of this state, $\Psi_0(\underset{\sim}{r}_1, \underset{\sim}{r}_2 \ldots \underset{\sim}{r}_N)$ satisfies equation 1.2, i.e.

$$\mathcal{H}\Psi_0 = E\Psi_0 \qquad (1.2)$$

The effect of operating with an element $(\alpha^{-1}|-\alpha^{-1}\underset{\sim}{a})$ of G on Ψ_0 is given by the formula

$$(\alpha|\underset{\sim}{a})^{-1}\Psi(\underset{\sim}{r}_1, \underset{\sim}{r}_2, \ldots \underset{\sim}{r}_N) = \Psi(\alpha\underset{\sim}{r}_1+\underset{\sim}{a}, \alpha\underset{\sim}{r}_2+\underset{\sim}{a}, \ldots \alpha\underset{\sim}{r}_N+\underset{\sim}{a}) \qquad (1.67)$$

Since $\mathcal{H}$ is invariant under G, any element of G commutes with $\mathcal{H}$, and so operating on both sides of equation 1.2 with $(\alpha|\underset{\sim}{a})^{-1}$ leads to the equation

$$\mathcal{H}\Psi_0(\alpha\underset{\sim}{r}_1+\underset{\sim}{a}, \alpha\underset{\sim}{r}_2+\underset{\sim}{a}, \ldots \alpha\underset{\sim}{r}_N+\underset{\sim}{a}) = E\Psi_0(\alpha\underset{\sim}{r}_1+\underset{\sim}{a}, \alpha\underset{\sim}{r}_2+\underset{\sim}{a}, \ldots \alpha\underset{\sim}{r}_N+\underset{\sim}{a}) \qquad (1.68)$$

However, by hypothesis, the ground state is non-degenerate, and so

$$\Psi_0(\alpha\underset{\sim}{r}_1+\underset{\sim}{a}, \alpha\underset{\sim}{r}_2+\underset{\sim}{a}, \ldots \alpha\underset{\sim}{r}_N+\underset{\sim}{a}) = A\Psi_0(\underset{\sim}{r}_1, \underset{\sim}{r}_2 \ldots \underset{\sim}{r}_N) \qquad (1.69)$$

where $AA^* = 1$ since a symmetry operation cannot affect the normalisation.

A consequence ofthis is that for the ground state of the system all the density matrices have the full symmetry of the system, and in particular for any operator $(\alpha|\underset{\sim}{a})$ in G

$$\gamma(\alpha\underset{\sim}{s}_1' + \underset{\sim}{a} \mid \alpha\underset{\sim}{s}_1 + \underset{\sim}{a}) = \gamma(\underset{\sim}{s}_1' \mid \underset{\sim}{s}_1) \qquad (1.70)$$

It therefore seems reasonable that when we consider which single-electron wave functions to use in order to produce a single Slater determinant as an approximate many-electron wave function for the non-degenerate ground state of the system, we should only select a set of solutions of equation 1.33 such that the first-order density matrix derived from them is invariant under G. This is not a constraint on the possible solutions of equation 1.33 but only a device to help us decide which of the solutions we choose to regard as occupied molecular orbitals, and so it should not introduce any extra terms into equation 1.33. It will be shown in the next chapter that, under certain conditions, if it is assumed that the solutions will have a certain symmetry <u>before</u> the variational procedure used to obtain equation 1.29 is carried out, one still obtains equation 1.33 in spite of the extra restriction on the variation.

Let us now define, for a closed-shell system, a spinless first-order density matrix

$$q(\underset{\sim}{r}_1' \mid \underset{\sim}{r}_1) = \sum_{i=1}^{\frac{1}{2}N} \phi_i^*(\underset{\sim}{r}_1')\,\phi_i(\underset{\sim}{r}_1) \qquad (1.71)$$

Then, in view of equation 1.20

$$q(\underset{\sim}{r}_1' \mid \underset{\sim}{r}_1) = \tfrac{1}{2}\int \gamma(\underset{\sim}{r}_1', \sigma_1 \mid \underset{\sim}{r}_1, \sigma_1)\, d\sigma_1 \qquad (1.72)$$

Hence, since $\gamma(\underset{\sim}{s}'|\underset{\sim}{s})$ is invariant under G, so is $q(\underset{\sim}{r}'|\underset{\sim}{r})$, and in

particular if $\underline{r}' = \alpha \underline{s}' + \underline{n}$ then

$$q(\underline{r}'|\underline{r}') = q(\underline{s}'|\underline{s}') \tag{1.73}$$

and

$$q(\underline{r}'|\alpha\underline{r}+\underline{n}) = q(\underline{s}'|\underline{r}) \tag{1.74}$$

Equation 1.34, which defined the Hartree-Fock operator F, can be re-written

$$F\phi_i(\underline{r}) = \Lambda_1 \phi_i(\underline{r}) + 2\int \frac{q(\underline{r}'|\underline{r}')}{|\underline{r}-\underline{r}'|} d\underline{r}' \phi_i(\underline{r}) - \int \frac{q(\underline{r}'|\underline{r})}{|\underline{r}-\underline{r}'|} \phi_i(\underline{r}') d\underline{r}' \tag{1.75}$$

where, as in equation 1.6,

$$\Lambda_1 = -\tfrac{1}{2}\nabla^2 - \sum_{\lambda=1}^{L} \frac{Z_\lambda}{|\underline{R}_\lambda - \underline{r}|}$$

Hence, if $(\alpha|\underline{n})^{-1}$ is an element of G, so that

$$(\alpha|\underline{n})^{-1} \Lambda_1 \phi(\underline{r}) = \Lambda_1 (\alpha|\underline{n})^{-1} \phi(\underline{r}) = \Lambda_1 \phi(\alpha\underline{r}+\underline{n}) \tag{1.76}$$

then, using equations 1.73 and 1.74, we find that

$$(\alpha|\underline{n})^{-1}[F\phi_i(\underline{r})] = \Lambda_1 \phi_i(\alpha\underline{r}+\underline{n}) + 2\int \frac{q(\underline{r}'|\underline{r}')}{|\alpha\underline{r}+\underline{n}-\underline{r}'|} d\underline{r}' \phi_i(\alpha\underline{r}+\underline{n}) - \int \frac{q(\underline{r}'|\alpha\underline{r}+\underline{n})}{|\alpha\underline{r}+\underline{n}-\underline{r}'|} \phi_i(\underline{r}')\, d\underline{r}'$$

$$= \Lambda_1 \phi_i(\alpha\underline{r}+\underline{n}) + 2\int \frac{q(\underline{s}'|\underline{s}')}{|\alpha\underline{r}-\alpha\underline{s}'|} d\underline{s}' \phi_i(\alpha\underline{r}+\underline{n}) - \int \frac{q(\underline{s}'|\underline{r})}{|\alpha\underline{r}-\alpha\underline{s}'|} \phi_i(\alpha\underline{s}'+\underline{n})\, d\underline{s}' \tag{1.77}$$

Since α is a rotation,

$$|\alpha\underline{r} - \alpha\underline{s}'| = |\underline{r}-\underline{s}'| \tag{1.78}$$

and so

$$(\alpha|\underline{n})^{-1} F\phi_i(\underline{r}) = F\phi_i(\alpha\underline{r}+\underline{n}) = F[(\alpha|\underline{n})^{-1}\phi_i(\underline{r})] \tag{1.79}$$

i.e. F is invariant under G, and all the elements of G commute with F.

This proof only applies to the ground state of the system, for which $q(\underset{\sim}{r}'|\underset{\sim}{r})$ is invariant under G. However, we can define an operator $H(\underset{\sim}{r})$, which we call the one-electron Hamiltonian, such that for an arbitrary well-behaved function $\phi(\underset{\sim}{r})$

$$H(\underset{\sim}{r})\,\phi(\underset{\sim}{r}) = -\tfrac{1}{2}\nabla^2\phi(\underset{\sim}{r}) - \sum_{\lambda=1}^{L'} \frac{Z_\lambda}{|\underset{\sim}{r}-\underset{\sim}{R}_\lambda|}\,\phi(\underset{\sim}{r}) + 2\int \frac{q(\underset{\sim}{r}'|\underset{\sim}{r}')}{|\underset{\sim}{r}-\underset{\sim}{r}'|}\,d\underset{\sim}{r}'\,\phi(\underset{\sim}{r}) - \int \frac{q(\underset{\sim}{r}'|\underset{\sim}{r})}{|\underset{\sim}{r}-\underset{\sim}{r}'|}\,\phi(\underset{\sim}{r}')\,d\underset{\sim}{r}' \qquad (1.80)$$

where $q(\underset{\sim}{r}'|\underset{\sim}{r})$ is derived (in accordance with equation 1.71 or 1.72) from the first-order density matrix for the ground state of the system. With this definition, H is invariant under the space group G of the system, and so all the results of group theory can be applied to its matrix elements. In particular, every eigenfunction of H can be required to belong to an irreducible representation of G. Some of the eigenfunctions of H with smallest eigenvalues (or one-election energies) will be molecular orbitals occupied in the ground state of the system, while the other eigenfunctions can be called unoccupied molecular orbitals. We must emphasise that these unoccupied orbitals do not satisfy the Hartree-Fock equations for an excited state of the system, since for such a state the Hartree-Fock operator F differs from its value for the ground state, while our H is only equivalent to the ground state F in its effect on orbitals occupied in the ground state. Even in the unlikely event of a single Slater determinant being a reasonable approximation to the wave function for an excited state, the energies of our "unoccupied molecular orbitals" will only be approximations to the one-electron energies which should be used to calculate the excitation energy in accordance with Koopmans' theorem. However,

it is the eigenfunctions of our one-electron Hamiltonian H which are normally used, without comment on the validity of the approximation, to represent the conduction bands of semiconducting crystals.

1.3 METHODS OF CALCULATING MOLECULAR ORBITALS IN CRYSTALS.

General Survey.

The most commonly-used methods for calculating electronic wave functions in crystals are the cellular method, plane wave method, linear combination of atomic orbital (L.C.A.O.) methods, and equivalent orbital methods. The first three of these groups of methods involve calculating the molecular orbitals directly, while in the last group the basic function calculated is an equivalent orbital, and the molecular orbitals are then derived from it; this method is considered at length in the next chapter.

Among all these methods, the cellular method[6] is unique in that it attempts to find exact solutions of the Hartree-Fock equations which only approximately satisfy the boundary conditions arising from the translational symmetry of the crystal; all the other methods use a set of basis functions which satisfy the boundary condisions exactly but are only approximate solutions of the Hartree-Fock equations.

In the cellular method, the function $u_{\underset{\sim}{k}}(\underset{\sim}{r})$ defined by equation 1.62 is calculated inside and on the surface of the Wigner-Seitz cell for different values of $\underset{\sim}{k}$, and $u_{\underset{\sim}{k}}(\underset{\sim}{r})$ and its derivatives are required to be continuous across the surface of the cell. At first sight, it might seem that, since this method is concerned with localised functions, it should be

closely connected with chemical ideas in much the same way as are equivalent orbitals. That this is not really so is shown by the dependence of $u_{\underset{\sim}{k}}(\underset{\sim}{r})$ on $\underset{\sim}{k}$, a vector which is only meaningful in connection with long-range order and translational symmetry. For instance, if $\phi_{\underset{\sim}{k}}(\underset{\sim}{r}) = \sum_{\underset{\sim}{n}} e^{i\underset{\sim}{k}.\underset{\sim}{n}} a(\underset{\sim}{r}-\underset{\sim}{n})$ where $a(\underset{\sim}{r})$ is a bond-orbital i.e. a wave function appropriate to an electron localised in a bond, then in each unit cell $u_{\underset{\sim}{k}}(\underset{\sim}{r})$ has at best the form $e^{-i\underset{\sim}{k}.\underset{\sim}{r}}\, a(\underset{\sim}{r})$, which cannot readily be interpreted chemically. Since the cellular method is not closely related to the other methods mentioned above, it will not be considered further in this thesis.

In the plane wave and L.C.A.O. methods, the molecular orbitals belonging to representation $\underset{\sim}{k}$ of the translational group are found as the sums of plane waves or Block waves having the required translational symmetry. These orbitals will be denoted by $\phi_t^{\underset{\sim}{k}}(\underset{\sim}{r})$ where the superscript $\underset{\sim}{k}$ indicates the translational symmetry of the orbital while the subscript t is needed as a band label since there are in general several molecular orbitals associated with each $\underset{\sim}{k}$. The eigenvalue of H associated with $\phi_t^{\underset{\sim}{k}}(\underset{\sim}{r})$ will be denoted by $E_t^{\underset{\sim}{k}}$.

The Orthogonalised Plane Wave Method.

One of the simplest methods of approximating to the one-electron energies in a crystal consists of starting with a set of plane waves, which are the eigenfunctions for electrons in an empty lattice, and treating the periodic potential in the crystal as a perturbation. This seems a rather drastic assumption, but it works quite well for the conduction bands of crystals of monovalent metals, for instance, provided that the basis set used contains only a few sets of planes waves of

the desired symmetry. However, if the ordinary one-electron crystal Hamiltonian is used, as more and more waves are included in the basis set the lowest eigenvalue approaches not the energy of a valence or conduction electron but that of an inner core electron. On the other hand, a vast number of plane waves is needed in order to obtain reasonable energies for core states, since the wave functions for such states are highly localised, so that their Fourier expansion involves a large number of significant terms.

In order to overcome this difficulty, Herring[7] proposed that a basis set consisting of plane waves orthogonalised to the atomic core functions be used. The lowest eigenfunction obtained using a complete basis set of this form will be the lowest valence band (or, for metals, conduction band) molecular orbital. It should be possible to obtain a reasonable approximation to its associated eigenvalue with the use of only a few such orthogonalised plane waves, provided that orthogonalisation to the core functions is an adequate way of representing the rapid oscillations of the valence and conduction band wave functions near the atomic nuclei. Orthogonalisation to core may not be adequate if, for instance, one is considering valence wave functions of a symmetry different from that of any core function, such as a p-band in diamond or a d-band in iron; in such cases it may be necessary (or at least convenient) to introduce one or more sets of auxiliary functions.[7][8] However, under other circumstances the requirement of orthogonality to a whole series of functions, plus a minimum number of nodes so as to keep the kinetic energy as low as possible, may well suffice to fix the form of the wave functions close to the nucleus. It is, in principle, possible to test whether

the basis set contains enough plane waves by seeing whether an enlargement of it has any appreciable effect on the eigenvalues and eigenfunctions in which one is interested; such a procedure is only possible where, as here, the basis set is chosen systematically from a complete set of functions. We note that in general if the original plane waves comprise an orthonormal set then the orthogonalised plane waves are neither mutually orthogonal nor normalised.

A development of the O.P.W. method was introduced by Kleinman and Phillips,[9] who suggested that instead of using a basis set of orthogonalised plane waves one could expand the wave functions in terms of ordinary plane waves and then orthogonalise the total wave function to the core functions. Such a procedure is equivalent to introducing a non-local pseudo-potential V_r into the equation which the wave function must satisfy, and so to finding the eigenfunctions of the operator $H' = H + V_r$ instead of finding the eigenfunctions of H. This potential V_r is not uniquely defined, and Cohen and Heine[10] have shown how to find the pseudo-potential which minimises the kinetic energy of the lowest eigenfunctions of H' and so should minimise the number of plane waves needed to expand these eigenfunctions. This pseudo-potential effectively cancels out H in the regions near atomic nuclei; physically this means that because of the Pauli exclusion principle the genuine valence and conduction band wave functions have an increase in their kinetic energy in these regions which more or less balances the gain in potential energy from being near the positively-charged nucleus. As a result, H' contains a net potential V' which is only slowly-varying; this explains why the method of treating the crystal potential as a perturbation on a system of free electrons can give such good results.

The nature of the pseudo-potential in this approach can be seen from the following analysis. Let $\Phi_t^{\underline{k}}$ be the core functions with translational symmetry $\underline{k}$, and let $\underline{Q}$ be the projection operator

$$\underline{Q} = \sum_{t,\underline{k}} |\Phi_t^{\underline{k}}\rangle\langle\Phi_t^{\underline{k}}| \qquad (1.81)$$

Then $\underline{Q}$ projects an arbitrary function into the Hilbert space spanned by the core states; thus if $\underline{P} = 1 - \underline{Q}$, then $\underline{P}$ projects an arbitrary function into the space orthogonal to the core states. Hence the equation

$$H\underline{P}|\psi\rangle = E\underline{P}|\psi\rangle \qquad (1.82)$$

has the same eigenvalues as does $H|\phi\rangle = E|\phi\rangle$ in the Hilbert space orthogonal to the core states. However, equation 1.82 does not uniquely determine the function $|\psi\rangle$ corresponding to the eigenfunction $\underline{P}|\psi\rangle$ of H, since if $|\chi\rangle$ is an arbitrary function we can add $\underline{Q}|\chi\rangle$ to $|\psi\rangle$ without affecting $\underline{P}|\psi\rangle$. From the definition of $\underline{P}$ and equation 1.82, we find that

$$H|\psi\rangle + (E-H)\underline{Q}|\psi\rangle = E|\psi\rangle \qquad (1.83)$$

where

$$(E-H)\underline{Q} = \sum_{t,\underline{k}} (E - \varepsilon_t^{\underline{k}})|\Phi_t^{\underline{k}}\rangle\langle\Phi_t^{\underline{k}}| \qquad (1.84)$$

is the pseudo-potential operator V_r. Just like the exchange potential, it is a non-local operator which for convenience is usually approximated to by a local potential; the form of this local potential will depend on the sort of function $|\psi\rangle$ that is being used, but it will generally introduce some errors. Another possible source of error is that while $\underline{P}|\psi\rangle$ is the function which should be normalised and used in the Slater

determinant, and so in equation 1.80 which defines H, in general only $|\psi\rangle$ is found from equations 1.83 and 1.84.

An Alternative Approach to Orthogonalisation to the Core States.

Although the pseudo-potential approach to orthogonalisation to the core states has certain advantages in helping one to understand what is happening in the crystal and so to predict the nature of the eigenfunctions and eigenvalues, it is possible to adopt an alternative procedure which may well be simpler for calculations. This procedure, which is described below, avoids the problems of normalisation and non-local potentials mentioned above, and also avoids the problems of orthonormality associated with orthogonal plane waves. In addition, it can conveniently be used even if the original set of basis functions is not an orthonormal set.

Let us attempt to expand the molecular orbital $\phi_s^{\underset{\sim}{k}}(\underset{\sim}{r})$ in terms of m orthonormal core functions $\Phi_t^{\underset{\sim}{k}}(\underset{\sim}{r})$ and n other linearly independent functions $v_t^{\underset{\sim}{k}}(\underset{\sim}{r})$ which need not be plane waves but which have translational symmetry $\underset{\sim}{k}$. We write

$$w_l(\underset{\sim}{r}) = \begin{cases} \Phi_l^{\underset{\sim}{k}}(\underset{\sim}{r}) & \text{if } l \le m \\ v_{l-m}^{\underset{\sim}{k}}(\underset{\sim}{r}) & \text{if } l > m \end{cases}, \qquad (1.85)$$

$$\phi_s^{\underset{\sim}{k}}(\underset{\sim}{r}) = \sum_{l=1}^{m} x_l \Phi_l^{\underset{\sim}{k}}(\underset{\sim}{r}) + \sum_{l=m+1}^{m+n} x_l v_{(l-m)}^{\underset{\sim}{k}}(\underset{\sim}{r}), \qquad (1.86)$$

$$H_{pq} = \langle w_p | H | w_q \rangle \quad \text{and} \quad S_{pq} = \langle w_p | w_q \rangle \qquad (1.87)$$

In general the functions $w_1(\underset{\sim}{r})$ will not constitute an orthonormal set, but by means of the Schmidt orthonormalisation procedure (see Appendix I) it is possible to find an orthonormal set of functions $w_1'(\underset{\sim}{r})$ which are linear combinations of the functions $w_j(\underset{\sim}{r})$. If the original functions were not linearly independent there will be less than m + n different functions $w_1'(\underset{\sim}{r})$, but in order to simplify the analysis we ignore this case, which is unlikely to arise in practise.

Let us write

$$w_i'(\underset{\sim}{r}) = \sum_{j=1}^{m+n} t_{ij}\, w_j(\underset{\sim}{r}) \qquad (1.88)$$

and

$$H'_{ij} = \langle w_i' | H | w_j' \rangle \qquad (1.89)$$

Then if $T = (t_{ij})$ and $\tilde{T}$ is the complex conjugate of its transpose, we find that $H' = T H \tilde{T}$ (1.90)

while $T S \tilde{T} = I$ the identity matrix. (1.91)

We now partition our matrices into blocks

$$\left(\begin{array}{c|c} m\times m & m\times n \\ \hline n\times m & n\times n \end{array}\right)$$

and, making use of the fact that H is a Hermitian operator, write

$$H = \left(\begin{array}{c|c} P & Q \\ \hline \tilde{Q} & R \end{array}\right), \quad S = \left(\begin{array}{c|c} I & C \\ \hline \tilde{C} & D \end{array}\right); \qquad (1.92)$$

if the expansion functions $v_t^{\underset{\sim}{k}}(\underset{\sim}{r})$ are normalised plane waves, then D = I, but this is not essential.

With this partitioning, we can write

$$T = \left(\begin{array}{c|c} I & 0 \\ \hline A & B \end{array}\right) \qquad (1.93)$$

since the core functions $\bar{\Phi}_t^k(r)$ were orthonormal; then from equation 1.91 we deduce that

$$A + B\tilde{C} = 0 \tag{1.94}$$

and

$$BD\tilde{B} + B\tilde{C}\tilde{A} = B(D - C\tilde{C})\tilde{B} = I \tag{1.95}$$

Also

$$H' = THT = \left(\begin{array}{c|c} P & P\tilde{A} + Q\tilde{B} \\ \hline AP + BQ & AP\tilde{A} + BQ\tilde{A} + AQ\tilde{B} + BR\tilde{B} \end{array}\right) \tag{1.96}$$

If, now, $Q = -PC$, i.e. $\langle u_i | H | v_j \rangle = E_i \langle u_i | v_j \rangle$ for each i, which will be true if the functions u_j are exact eigenfunctions of our one-electron Hamiltonian H, then

$$P\tilde{A} + Q\tilde{B} = P(\tilde{A} + C\tilde{B}) = 0 \tag{1.97}$$

and

$$H' = \left(\begin{array}{c|c} P & 0 \\ \hline 0 & B(R - \tilde{C}PC)\tilde{B} \end{array}\right) \tag{1.98}$$

Thus, the eigenfunctions of H' are just those of P, i.e. the core functions, and those of $B(R - \tilde{C}PC)\tilde{B}$.

If we now write $\tilde{T}\underline{Y} = \underline{x}$, and $\underline{x}$ is an eigenvector of H, then since $H\underline{x} = ES\underline{x}$ (1.99)

we have that

$$T(H\tilde{T}\underline{Y}) = ET(S\tilde{T}Y) \quad \text{i.e.} \quad H'\underline{Y} = E\underline{Y} \tag{1.100}$$

Also, it is apparent from equation 1.98 that the eigenvectors of H' either correspond to core functions or are of the form $\begin{pmatrix} 0 \\ \underline{y} \end{pmatrix}$ where $\underline{y}$ is a $\sim \times 1$ column vector, while since T is non-singular every eigenfunction of H' corresponds to an eigenfunction of H. Hence, finally, if $\underline{y}$ is an eigenfunction of $B(R - \tilde{C}PC)\tilde{B}$ and

$$\underline{x} = \tilde{T}\begin{pmatrix} 0 \\ \underline{y} \end{pmatrix} = \begin{pmatrix} \tilde{A}\underline{y} \\ \tilde{B}\underline{y} \end{pmatrix} \tag{1.101}$$

then $\sum_j x_j w_j(r)$ is an eigenfunction of H orthogonal to the core functions.

Thus, if a set of basis functions not orthogonal to the core functions is used for the expansion of the valence and conduction band wave functions, then our method derives from the matrix of the ordinary one-electron Hamiltonian H a matrix of order n whose eigenvalues are just those of the non-core eigenfunctions of H in this basis. In order to relate this to the pseudo-potential method for the case where our expansion functions $v_t^{\underset{\sim}{k}}(\underset{\sim}{r})$ are normalised plane waves, let us write $\tilde{B}\underset{\sim}{y} = \underset{\sim}{z}$; we note that the matrix B, like T, is non-singular.

Then if, as assumed previously,

$$B(R - \tilde{C} P C)\tilde{B}\,\underset{\sim}{y} = E\,\underset{\sim}{y} \qquad (1.102)$$

we find that

$$(R - \tilde{C} P C)\,\underset{\sim}{z} = E\,(\tilde{B}B)^{-1}\,\underset{\sim}{z} \qquad (1.103)$$

Also, since $P = I$ in this case,

$$(\tilde{B}B)^{-1} = I - C\tilde{C} \qquad (1.104)$$

in view of equation 1.95, and so

$$R\underset{\sim}{z} + \tilde{C}(EI - P)C\underset{\sim}{z} = E\underset{\sim}{z}\ ; \qquad (1.105)$$

this corresponds exactly to equation 1.83 applied to our set of basis functions. In addition, the normalisation condition $\tilde{\underset{\sim}{y}}\,\underset{\sim}{y} = 1$ leads to

$$\tilde{\underset{\sim}{z}}\,(I - \tilde{C}C)\,\underset{\sim}{z} = 1 \qquad (1.106)$$

Thus in this case we obtain in matrix form the same equations as those of the usual O.P.W. method.

An advantage of the above analysis is that, provided we omit the requirement $Q = PC$, it can be applied to arbitrary sets of expansion functions. For instance, the core functions we use may be eigenfunctions of the isolated-atom Hamiltonian rather than of the crystal Hamiltonian H, in which case a rigorous application of our analysis corresponds to the modified

plane wave (M.P.W.) method of Brown and Krumhansl.[12] Also, auxiliary localised functions such as were suggested by Herring[7] and Callaway[8] can easily be introduced. The scheme can also be used in L.C.A.O. and equivalent orbital calculations, inter alia, and was used by us in our calculations of the bond orbitals in tetrahedral semiconductors.

The Augmented Plane Wave Method.

As remarked above, there is no certainty that the orthogonalisation of a smooth wave function to the core orbitals leads to a wave function with the correct behaviour near the atomic nuclei. An alternative approach to this problem is to divide the crystal into two parts, and assume that the crystal potential is spherically symmetric inside spheres surrounding the nuclei and is constant in the regions between these spheres. With such a potential, it is possible to find solutions of Schrodinger's equation inside the atomic spheres which join smoothly on to plane waves (the solutions of this equation in regions of constant potential) at the boundaries of the spheres to form augmented plane waves.[13] The difference between the actual and the hypothetical crystal potentials can then be treated as a perturbation. This method bears certain resemblances to the approximation to the cellular method in which the Wigner-Seitz cell is replaced by a sphere, but is a lot more accurate.

The augmented plane wave method will be expected to be most successful for crystals such as sodium in which there is a high degree of rotational symmetry about each atom, since in that case the difference between the hypothetical and actual potentials should be fairly small. It is more doubtful whether

this difference is small enough to be treated as a perturbation for crystals of lower symmetry, such as those having the zinc-blende structure where the symmetry at each atomic site is only tetrahedral. A disadvantage of this method is that the molecular orbital wave functions obtained do not have a simple analytical form, and so although very good results have been obtained by the method [14] we will not consider it further in this thesis.

L.C.A.O. Methods.

All the methods considered so far have involved large-scale computations, and these are probably necessary if accurate energies and wave functions are to be obtained, especially for conduction bands, i.e. orbitals not occupied in the ground state. However, for many purposes it is convenient to have wave functions which are easily calculated at all points of the Brillouin zone, even if they are not so accurate. In that case, the method of linear combinations of atomic orbitals (L.C.A.O.) is very useful for crystals, just as it is in the case of molecules.[15] In particular, the L.C.A.O. method is one of the most helpful if one wishes to compare and explain the properties of crystals and molecules in terms of bonds, ionicity parameters, charge transfer, and the like.

The basis of the L.C.A.O. method for crystals is that, starting with a few wave functions $\psi_\ell(\underline{r})$ appropriate to the outer electrons in the ground and/or excited states of the free atoms or ions, one forms the linear combinations (or Bloch waves)

$$\psi_\ell^{\underline{k}}(\underline{r}) = \frac{1}{\sqrt{M}} \sum_{\underline{n}} e^{i\underline{k}\cdot\underline{n}} (\varepsilon|\underline{n}) \psi_\ell(\underline{r}) = \frac{1}{\sqrt{M}} \sum_{\underline{n}} e^{i\underline{k}\cdot\underline{n}} \psi_\ell(\underline{r}-\underline{n}) \qquad (1.107)$$

which belong to representation $\underline{k}$ of the translation group. The t'th molecular orbital with this translational symmetry,

$\phi_t^{\underline{k}}(\underline{r})$ is then expanded as a linear combination of these,

$$\phi_t^{\underline{k}}(\underline{r}) = \sum_l c_{tl}^{\underline{k}} \psi_l^{\underline{k}}(\underline{r}) \qquad (1.108)$$

The method is not exact, since the bound states of an atom do not form a complete set of functions, so that in general the functions $\psi_l^{\underline{k}}(\underline{r})$ do not form a complete set either. However, if only a few functions $\psi_l(\underline{r})$ corresponding to the highest occupied and lowest unoccupied states of the atom are used, the calculations are usually of manageable proportions and may be expected to give results which are qualitatively correct. Incidentally, a combination of a few such functions $\psi_l^{\underline{k}}(\underline{r})$ and a number of plane waves might well provide an extremely good basis set for an exact calculation, since the behaviour of valence and conduction band wave functions near the atomic nuclei may well closely resemble that of the ground and excited state wave functions of the isolated atom. With a little care, it would not be diffucult to ensure that such a basis set was not over-complete and did not contain linearly dependent functions. Unfortunately, such a calculation involves the evaluation of some very difficult integrals which cannot be estimated from experimental data or similar sources.

The simplest application of the L.C.A.O. method occurs with π -electron systems such as graphite or planar boron nitride. In these there are just two atoms per unit cell, and as in the L.C.A.O. treatment of benzene only one orbital per atom is used. Hence the approximate molecular orbitals are of the form

$$\phi_t^{\underline{k}}(\underline{r}) = \frac{1}{\sqrt{MN_{\underline{k}}}} \sum_{\underline{n}} e^{i\underline{k}\cdot\underline{n}} (\varepsilon|\underline{n}) \left[\psi_1(\underline{r}-\underline{b}) + \lambda_{\underline{k}} \psi_2(\underline{r}+\underline{b}) \right], \qquad (1.109)$$

where $N_{\underline{k}}$ is a normalisation factor, $\lambda_{\underline{k}}$ is an ionicity parameter, and the secular determinant is only of order two.

If in addition a tight-binding approximation is used and all overlaps and matrix elements are neglected except those between orbitals on the same or nearest-neighbour atoms, one obtains the wave functions reported by Coulson and Taylor,[16] which we use in the last chapter of this thesis.

It has been objected that the L.C.A.O. method, at least in the tight-binding approximation, must be very inaccurate since it completely fails to reproduce the effective masses which are found experimentally.[17] This criticism is not really fair, since by its very nature a tight-binding approximation cannot be expected to deal adequately with effective masses. For if we write, dropping the band label for convenience,

$$E_{\underset{\sim}{k}} = \sum_{\underset{\sim}{n}} e^{-i\underset{\sim}{k}\cdot\underset{\sim}{n}} \varepsilon(\underset{\sim}{n}) , \qquad (1.110)$$

which is always possible since $E_{\underset{\sim}{k}}$ is periodic in $\underset{\sim}{k}$, then the Fourier coefficients $\varepsilon(\underset{\sim}{n})$ are related to the matrix elements of H between orbitals about atoms in cells distant $\underset{\sim}{n}$ from one another. Differentiation of equation 1.110 shows that, for instance,

$$\frac{\partial^2 E}{\partial k_x^2} = \sum_{\underset{\sim}{n}} n_x^2 \, \varepsilon(\underset{\sim}{n}) \, e^{-i\underset{\sim}{k}\cdot\underset{\sim}{n}} \qquad (1.111)$$

so that even if for large $|\underset{\sim}{n}|$ we can ignore $\varepsilon(\underset{\sim}{n})$ in finding $E_{\underset{\sim}{k}}$ without introducing too great an error, it may not be valid to ignore it when calculating the xx component of the effective mass tensor because of the weight factor n_x^2. Hence, even if the values obtained for effective masses by the L.C.A.O. method are erroneous, the method may still give quite accurate results for such quantities as energy levels, the density of states curve, and the charge distribution, especially for the

ground state. A more detailed treatment of effective masses for the special case of bond orbitals in tetrahedral semiconductors is presented in Appendix II; we show there that, because of equation 1.111, any localised orbital method would require extreme accuracy in the calculation of some very small parameters in order to give reliable effective masses.

The L.C.B.O. Method.

For crystals having the zincblende structure, in which we are principally interested, it is usual to use four orbitals per atom in an L.C.A.O. approach. Since there are two atoms per unit cell, this leads to eight bands of states, the lowest four of which should correspond to the four valance bands which are filled in the ground state of the crystal; whether the upper four bands give an adequate representation of the four lowest conduction bands is more dubious. In any case, the valence band energies and wave functions can be found at a general point in reciprocal space by solving a set of eight secular equations, while at points or on lines of special symmetry the secular determinant factorises into two or more determinants of smaller order. The basic atomic orbitals used will normally be either the highest s and p orbitals occupied in the ground state of the free atom, e.g. $4s$, $4p_x$, $4p_y$, and $4p_z$ for Cu, Zn, Ga, Ge, As, and Se, or the four hybridised sp^3 orbitals which can be constructed from these. It is apparent that the two sets of orbitals are equivalent, and if matrix elements are calculated rather than estimated semi-empirically the use of sp^3 orbitals would seem to involve an extra amount of work without conferring any advantages. This is especially true for certain symmetry-directions in

reciprocal space, such as (ν , 0, 0,) where simple combinations of the s and p orbitals belong to the irreducible representations of $P_{\underset{\sim}{k}}$.

A modification of the L.C.A.O. method for tetrahedral semiconductors is the method of linear combinations of bond orbitals (L.C.B.O.) [16],[19] in which the basis functions are not atomic orbitals $\psi_i(r)$ but rather bond orbitals of the form

$$\chi_i(\underset{\sim}{r}) = \chi_i^A(\underset{\sim}{r}) + \lambda \chi_i^B(\underset{\sim}{r}) \quad , \quad i=1,2,3,4, \tag{1.112}$$

where χ_i^A and χ_i^B are sp^3 orbitals centred on atoms A and B_i respectively and pointing along the AB_i bond. The value of λ in equation 1.112 appropriate to the valence bands can be chosen either so as to minimise the total energy of the crystal[20] or so as to give self-consistent effective charges on the A and B atoms.[18] In either case a real value of λ is obtained, while for the homonuclear diamond-type crystals $\lambda = 1$ automatically. The valence band molecular orbitals are then expressed as linear combinations of the Bloch waves formed from these bond orbitals, i.e. of the functions

$$\chi_i^{\underset{\sim}{k}}(\underset{\sim}{r}) = \frac{1}{\sqrt{MN}} \sum_{\underset{\sim}{n}} e^{i\underset{\sim}{k}\cdot\underset{\sim}{n}} (\varepsilon|\underset{\sim}{n}) \left[\chi_i^A(\underset{\sim}{r}) + \lambda \chi_i^B(\underset{\sim}{r})\right], \quad i=1,2,3,4 \tag{1.113}$$

As a result, it is only necessary to solve a set of four secular equations to find the eigenvalues and eigenfunctions at a general point in reciprocal space. This halving of the order of the secular determinant as compared with the L.C.A.O. method is obtained, in effect, by assuming that in equation 1.112 λ is independent of $\underset{\sim}{k}$ and i, and also does not depend on which valence band is being considered. Although λ can be interpreted as an ionicity parameter of a type familiar from the study of diatomic molecules, these assumptions are not easily justified

and it seems probable that lower energies can be obtained if they are not made. Although the analogy is not very good, we find for instance that in graphite[16] while $|\lambda| = 1$ the best value of λ, i.e. that producing the lowest energies for any given value of $\underset{\sim}{k}$, has a phase factor depending on $\underset{\sim}{k}$.

So far, we have discussed the L.C.B.O. method as an approximation to the L.C.A.O. Method. It is, however, also possible to regard the bond orbitals as approximations to equivalent orbitals, in which case λ will in fact be a constant independent of $\underset{\sim}{k}$, l, or the valence band being considered. The theory for such an approach is developed in the next chapter.

1.4 THE RELATIONSHIP BETWEEN L.C.A.O. AND O.P.W. METHODS.

Alternative Forms of the L.C.A.O. Method.

Just as in the O.P.W. method one can derive the eigenvalues and eigenfunctions for the valence and conduction bands of a crystal by starting with a set of plane waves and then orthogonalising either these plane waves or the final eigenfunctions to the core functions, so in the L.C.A.O. method one can start with a set of smooth functions corresponding to the outer parts of atomic orbitals and then orthogonalise these to core functions in various ways. We take as our initial basis set the localised smooth atomic functions $p_j(\underset{\sim}{r})$, which could for instance be Slater orbitals,[21] and consider the various ways in which they can be used to form a suitable basis set for calculating the molecular orbitals $\phi_s^{\underset{\sim}{k}}(\underset{\sim}{r})$ with translational symmetry $\underset{\sim}{k}$; for convenience, we drop the band label s. As previously, we assume that the crystalline core

functions are of the form

$$\Phi_t^{\underset{\sim}{k}}(\underset{\sim}{r}) = \frac{1}{\sqrt{M}} \sum_{\underset{\sim}{n}} e^{i\underset{\sim}{k}\cdot\underset{\sim}{n}} \phi_t(\underset{\sim}{r}-\underset{\sim}{n}) \qquad (1.114)$$

where $\phi_t(\underset{\sim}{r})$ is an atomic core function and

$$\langle \phi_s(\underset{\sim}{r}) | \phi_t(\underset{\sim}{r}-\underset{\sim}{n}) \rangle = \delta_{st}\,\delta_{\underset{\sim}{n},\underset{\sim}{0}} \qquad (1.115)$$

A simple linear transformation connects the set of functions $\Phi_t^{\underset{\sim}{k}}(\underset{\sim}{r})$, where $\underset{\sim}{k}$ lies in the first Brillouin zone, with the set of functions $\phi_t(\underset{\sim}{r}-\underset{\sim}{n})$, which are the ones more commonly thought of as core functions. The advantage of the former set is that arbitrary functions associated with non-equivalent vectors $\underset{\sim}{k}$ are automatically orthogonal if the volume of integration extends over all the M unit cells in the crystal.

Let $q\phi_j(\underset{\sim}{r})$ be the function formed by orthogonalising $p\phi_j(\underset{\sim}{r})$ to the core functions on its own atom, i.e.

$$q_j(\underset{\sim}{r}) = \lambda_j(\underset{\sim}{r}) - \sum_t^{j} \langle \phi_t(\underset{\sim}{r}) | \lambda_j(\underset{\sim}{r}) \rangle \phi_t(\underset{\sim}{r}) \qquad (1.116)$$

where the symbol $\sum^{j}$ denotes summation over those core functions associated with the same atom as $p\phi_j(\underset{\sim}{r})$. We note that these functions $q\phi_j(\underset{\sim}{r})$ are approximations to the exact atomic orbitals normally used in the L.C.A.O. method. Also, let $c\phi_j(\underset{\sim}{r})$ be the function formed by orthogonalising $p\phi_j(\underset{\sim}{r})$ to all the core orbitals in the crystal, i.e.

$$c_j(\underset{\sim}{r}) = \lambda_j(\underset{\sim}{r}) - \sum_t \sum_{\underset{\sim}{n}} \langle \phi_t(\underset{\sim}{r}-\underset{\sim}{n}) | \lambda_j(\underset{\sim}{r}) \rangle \phi_t(\underset{\sim}{r}-\underset{\sim}{n}) \qquad (1.117)$$

From the localised functions $p\phi_j(\underset{\sim}{r})$, $q\phi_j(\underset{\sim}{r})$, and $c\phi_j(\underset{\sim}{r})$ we form the Bloch waves

$$P_j^{\underset{\sim}{k}}(\underset{\sim}{r}) = \frac{1}{\sqrt{M}} \sum_{\underset{\sim}{n}} e^{i\underset{\sim}{k}\cdot\underset{\sim}{n}} \lambda_j(\underset{\sim}{r}-\underset{\sim}{n}) \qquad (1.118)$$

$$Q_j^{\underset{\sim}{k}}(\underset{\sim}{r}) = \frac{1}{\sqrt{M}} \sum_{\underset{\sim}{n}} e^{i\underset{\sim}{k}\cdot\underset{\sim}{n}} q_j(\underset{\sim}{r}-\underset{\sim}{n}) \qquad (1.119)$$

and $$C_j^{\underset{\sim}{k}}(\underset{\sim}{r}) = \frac{1}{\sqrt{M}} \sum_{\underset{\sim}{n}} e^{i\underset{\sim}{k}\cdot\underset{\sim}{n}} c_j(\underset{\sim}{r}-\underset{\sim}{n}). \qquad (1.120)$$

We note that

$$\langle \frac{1}{\sqrt{M}} \sum_{\underset{\sim}{n}} e^{i\underset{\sim}{k}\cdot\underset{\sim}{n}} a(\underset{\sim}{r}-\underset{\sim}{n}) \Big| \frac{1}{\sqrt{M}} \sum_{\underset{\sim}{m}} e^{i\underset{\sim}{k}\cdot\underset{\sim}{m}} b(\underset{\sim}{r}-\underset{\sim}{m}) \rangle \qquad (1.121)$$

$$= \frac{1}{M} \sum_{\underset{\sim}{m}} \sum_{\underset{\sim}{n}} e^{i\underset{\sim}{k}\cdot(\underset{\sim}{m}-\underset{\sim}{n})} \langle a(\underset{\sim}{r}-\underset{\sim}{n}) | b(\underset{\sim}{r}-\underset{\sim}{m}) \rangle = \sum_{\underset{\sim}{l}} e^{i\underset{\sim}{k}\cdot\underset{\sim}{l}} \langle a(\underset{\sim}{r}) | b(\underset{\sim}{r}-\underset{\sim}{l}) \rangle$$

From the functions $P_j^{\underset{\sim}{k}}(\underset{\sim}{r})$ we can form the set of suitably orthogonalised functions

$$\Pi_j^{\underset{\sim}{k}}(\underset{\sim}{r}) = \rho_j^{\underset{\sim}{k}}(\underset{\sim}{r}) - \sum_t \langle \Phi_t^{\underset{\sim}{k}}(\underset{\sim}{r}) | \rho_j^{\underset{\sim}{k}}(\underset{\sim}{r}) \rangle \Phi_t^{\underset{\sim}{k}}(\underset{\sim}{r}) \qquad (1.122)$$

and similar sets could be formed from the functions $\varrho_j^{\underset{\sim}{k}}(\underset{\sim}{r})$ and $C_j^{\underset{\sim}{k}}(\underset{\sim}{r})$. However, while on applying equations 1.121, 1.117, and 1.115 we find that

$$\langle \Phi_t^{\underset{\sim}{k}}(\underset{\sim}{r}) | C_j^{\underset{\sim}{k}}(\underset{\sim}{r}) \rangle = \sum_{\underset{\sim}{l}} e^{-i\underset{\sim}{k}\cdot\underset{\sim}{l}} \langle \phi_t(\underset{\sim}{r}-\underset{\sim}{l}) | \rho_j(\underset{\sim}{r}) - \sum_s \sum_{\underset{\sim}{n}} \langle \phi_s(\underset{\sim}{r}-\underset{\sim}{n}) | \rho_j(\underset{\sim}{r}) \rangle \phi_s(\underset{\sim}{r}-\underset{\sim}{n}) \rangle$$

$$= \sum_{\underset{\sim}{l}} e^{-i\underset{\sim}{k}\cdot\underset{\sim}{l}} \left[\langle \phi_t(\underset{\sim}{r}-\underset{\sim}{l}) | \rho_j(\underset{\sim}{r}) \rangle - \sum_s \sum_{\underset{\sim}{n}} \langle \phi_s(\underset{\sim}{r}-\underset{\sim}{n}) | \rho_j(\underset{\sim}{r}) \rangle \delta_{st}\delta_{\underset{\sim}{n},\underset{\sim}{l}} \right]$$

$$= 0 \qquad (1.123)$$

so that the functions $C_j^{\underset{\sim}{k}}(\underset{\sim}{r})$ are themselves a suitably orthogonalised basis set, this is <u>not</u> true in general of the $\varrho_j^{\underset{\sim}{k}}(\underset{\sim}{r})$, which are the functions normally used in applying the L.C.A.O. method to crystals. Thus unless a sort of tight-binding approximation is used, and the overlaps between valence functions on one atom and core orbitals on any other atom are assumed to be negligible, the usual L.C.A.O. method does not allow fully for the requirement that the valence and conduction band wave functions must be orthogonal to all the crystal core functions.

Incidentally, it is easy to show that

$$C_j^{\underset{\sim}{k}}(\underset{\sim}{r}) = \Pi_j^{\underset{\sim}{k}}(\underset{\sim}{r}) \qquad (1.124)$$

For, from equations 1.120 and 1.117,

$$C_j^{\underline{k}}(\underline{r}) = \rho_j^{\underline{k}}(\underline{r}) - \frac{1}{\sqrt{M}} \sum_t \sum_{\underline{n}} \sum_{\underline{m}} e^{i\underline{k}\cdot(\underline{n}+\underline{m})} \phi_t(\underline{r}-\underline{n}-\underline{m}) \langle e^{i\underline{k}\cdot\underline{m}} \phi_t(\underline{r}-\underline{m}) | \rho_j(\underline{r}) \rangle$$

$$= \rho_j^{\underline{k}}(\underline{r}) - \sum_t \frac{1}{\sqrt{M}} \sum_{\underline{l}} e^{i\underline{k}\cdot\underline{l}} \phi_t(\underline{r}-\underline{l}) \left[\frac{1}{M} \sum_{\underline{m}} \sum_{\underline{m}'} \langle e^{i\underline{k}\cdot\underline{m}} \phi_t(\underline{r}-\underline{m}) | e^{i\underline{k}\cdot\underline{m}'} \rho_j(\underline{r}-\underline{m}' \right.$$

$$= \rho_j^{\underline{k}}(\underline{r}) - \sum_t \Phi_t^{\underline{k}}(\underline{r}) \langle \Phi_t^{\underline{k}}(\underline{r}) | \rho_j^{\underline{k}}(\underline{r}) \rangle = \Pi_j^{\underline{k}}(\underline{r}) \qquad \underline{Q.E.D.}$$

Hence it makes no difference in the L.C.A.O. method whether the orbitals used are orthogonalised to all the core orbitals in the crystal before or after the Bloch waves are formed. In practice, it may well be simpler to form the Bloch waves first, especially as in such a case pseudopotentials can be introduced to allow for orthogonalisation to the core functions just as in the O.P.W. method. Thus, in the L.C.A.O. approximation we can write for a molecular orbital with translational symmetry $\underline{k}$

$$\phi^{\underline{k}}(\underline{r}) = \sum_j b_j \Pi_j^{\underline{k}}(\underline{r}). \qquad (1.125)$$

Formal Similarity of O.P.W. and L.C.A.O. Methods.

If we write

$$x_t = \sum_j b_j \langle \Phi_t^{\underline{k}}(\underline{r}) | \rho_j^{\underline{k}}(\underline{r}) \rangle, \quad 1 \le t \le m \; ; \; x_t = b_{(t-m)}, \quad m \le t \le m+n \qquad (1.126)$$

then, from equations 1.125 and 1.122 we find that

$$\phi^{\underline{k}}(\underline{r}) = \sum_{l=1}^{m} x_l \Phi_l^{\underline{k}}(\underline{r}) + \sum_{l=m+1}^{m+n} x_l \rho_{(l-m)}^{\underline{k}}(\underline{r}) \qquad (1.127)$$

which is essentially the same as equation 1.86. Thus, there is a formal similarity between the L.C.A.O. and O.P.W. methods,[11] and the main difference between the methods lies in whether Bloch waves or plane waves are used for the expansion functions. In practice, it is usually a lot simpler to calculate matrix elements of the Hamiltonian between plane waves than between

the localised functions used to form the Bloch waves (although for the latter these elements can sometimes be estimated semi-empirically). As a result, a much greater number of expansion functions $U_l^k(r)$ can conveniently be used in the O.P.W. method than in the L.C.A.O. method. Since the accuracy of a variational calculation usually increases with the number of independent parameters being varied, other things being equal, it would seem that the O.P.W. method is more accurate, and so more useful, than the L.C.A.O. method. It also has the advantage of using a fairly systematic set of expansion functions, in which the number of terms can readily be increased so as to test convergence.

On the other hand, calculations by Herman[22] and others suggest that at least fifty sets of plane waves may be needed in order to obtain convergence at points of symmetry in reciprocal space, and far more at a general point. Also, the O.P.W. method does not show clearly the relationship between a whole range of substances, such as the inter-metallic semi-conducting compounds, or even between two forms of the same compound, such as the zincblende and wurtzite forms for compounds which can exist in both structures. This latter case is of particular interest, since the structures differ only in the location of third-neighbour and more distant atoms; thus they ought to be very similar chemically, while the O.P.W. results for the cubic zincblende form would look very different from those for the hexagonal wurtzite form. Another disadvantage of the L.C.A.O. method, if it were reliable, would be that the energies at a general point in reciprocal space could readily be calculated and so a density of states curve, which is of vital importance in the theory of many observable effects, could be obtained. For these reasons, it would be very useful

if a few localised atomic orbitals could be found, the Bloch waves formed from which constituted an adequate basis for expanding the valence and conduction band molecular orbitals. It would not be too surprising if such functions could be found, just as Slater orbitals[21] fulfil a similar purpose in atomic calculations, but there seem to be no a priori reasons why the same orbitals should be appropriate for crystals as for free atoms. We note that, for a crystal containing more than one atom per unit cell, the functions we are looking for need not resemble the ordinary (and rather diffuse) Wannier functions.[23]

A possible method of finding suitable orbitals for a particular substance is to attempt to match the Bloch waves obtained from them with the eigenfunctions obtained from a large-scale O.P.W. calculation. It is not unreasonable to hope that smooth crystal functions $p_j(\underline{r})$ could be found which correspond to the plane wave parts of the O.P.W. eigenfunctions, although as the latter are not uniquely defined it might be necessary to add on a certain amount of core functions in order to obtain convergence. This problem, coupled with the fact that O.P.W. eigenfunctions are usually not known nearly as accurately as the eigenvalues, constitutes the major practical difficulty in the way of performing such a calculation. This method of matching wave functions is to be contrasted with that of interpolating energy parameters proposed by Slater and Koster.[24] In order to carry it out, it should be convenient to make use of the formalism developed by Parmenter.[25] His conclusions concerning the uselessness of the L.C.A.O. method are probably not valid, however, since he represented his atomic orbitals by Gaussian

functions, and it is well known that the Fourier components of these decrease far more rapidly with $|\underset{\sim}{k}|$ than do those of the Slater orbitals more commonly used in L.C.A.O. calculations. This matter is discussed further in chapter four.

Fourier Analysis of Bloch Waves.

Let $$\rho^{\underset{\sim}{k}}(\underset{\sim}{r}) = \frac{1}{\sqrt{M}} \sum_{\underset{\sim}{n}} e^{i\underset{\sim}{k}\cdot\underset{\sim}{n}} \rho(\underset{\sim}{r}-\underset{\sim}{n}) \qquad (1.128)$$

where $\rho(\underset{\sim}{r})$ is a fairly localised function, and let us write

$$\rho^{\underset{\sim}{k}}(\underset{\sim}{r}) = \frac{1}{\sqrt{M}} \sum_{\underset{\sim}{K}} A^{\underset{\sim}{k}}_{\underset{\sim}{K}} e^{i(\underset{\sim}{k}+\underset{\sim}{K})\cdot\underset{\sim}{r}} \qquad (1.129)$$

the sum being taken over all reciprocal lattice vectors $\underset{\sim}{K}$. Then, on multiplying both sides of the equation by $e^{-i(\underset{\sim}{k}+\underset{\sim}{K})\underset{\sim}{r}}/\sqrt{\Omega}\sqrt{M}$ and integrating over all the M unit cells in the crystal, we find that

$$A^{\underset{\sim}{k}}_{\underset{\sim}{K}} = \frac{1}{M\Omega} \sum_{\underset{\sim}{n}} e^{i\underset{\sim}{k}\cdot\underset{\sim}{n}} \int e^{-i(\underset{\sim}{k}+\underset{\sim}{K})\cdot\underset{\sim}{r}'} \rho(\underset{\sim}{r}'-\underset{\sim}{n})\, d\underset{\sim}{r}'$$

$$= \frac{1}{\sqrt{\Omega}} \int \rho(\underset{\sim}{r})\, e^{-i(\underset{\sim}{k}+\underset{\sim}{K})\cdot\underset{\sim}{r}}\, d\underset{\sim}{r} \qquad (1.130)$$

A corollary of this is that if $P^{\underset{\sim}{k}}_{\underset{\sim}{a}}(\underset{\sim}{r})$ is a Bloch wave formed from functions centred about a point distant $\underset{\sim}{a}$ from the centre of the unit cell, so that

$$\rho^{\underset{\sim}{k}}_{\underset{\sim}{a}}(\underset{\sim}{r}) = \frac{1}{\sqrt{M}} \sum_{\underset{\sim}{n}} e^{i\underset{\sim}{k}\cdot\underset{\sim}{n}} \rho(\underset{\sim}{r}-\underset{\sim}{n}-\underset{\sim}{a}) \qquad (1.131)$$

then $$\rho^{\underset{\sim}{k}}_{\underset{\sim}{a}}(\underset{\sim}{r}) = \frac{1}{\sqrt{M}} \sum_{\underset{\sim}{K}} A^{\underset{\sim}{k}}_{\underset{\sim}{K}} e^{-i(\underset{\sim}{k}+\underset{\sim}{K})\cdot\underset{\sim}{a}}\, e^{i(\underset{\sim}{k}+\underset{\sim}{K})\cdot\underset{\sim}{r}} \qquad (1.132)$$

while if $$Q^{\underset{\sim}{k}}_{a}(\underset{\sim}{r}) = \frac{1}{\sqrt{M}} \sum_{\underset{\sim}{n}} e^{i\underset{\sim}{k}\cdot(\underset{\sim}{n}+\underset{\sim}{a})} \rho(\underset{\sim}{r}-\underset{\sim}{n}-\underset{\sim}{a}) \qquad (1.133)$$

then $$Q^{\underset{\sim}{k}}_{a}(\underset{\sim}{r}) = \frac{1}{\sqrt{M}} \sum_{\underset{\sim}{K}} A^{\underset{\sim}{k}}_{\underset{\sim}{K}} e^{-i\underset{\sim}{K}\cdot a}\, e^{i(\underset{\sim}{k}+\underset{\sim}{K})\cdot\underset{\sim}{r}} \qquad (1.134)$$

It is comparatively simple to calculate the coefficients of $A^{\underset{\sim}{k}}_{\underset{\sim}{K}}$ for Slater functions of s and p symmetry. Thus, if $\underset{\sim}{K} = \underset{\sim}{K} + \underset{\sim}{k}$,

then for an s - funetion with $p(\underline{r}) = r^n e^{-cr}$ we find that, provided $K \neq 0$,

$$A_{\underline{K}}^{\underline{k}} = \frac{4\pi}{\sqrt{\Omega}} \int_0^\infty r^{n+2} e^{-cr} \frac{\sin Kr}{Kr} dr = \frac{4\pi}{\sqrt{\Omega K}} \frac{\Gamma(n+2)}{\beta^{n+2}} \sin(n+2)\theta \qquad (1.135)$$

where $\theta = \tan^{-1} K/c$ and $\beta = (K^2+c^2)^{\frac{1}{2}}$ (1.136)

Similarly for a p - funetion, if we write

$$\cos\alpha = -K_z/K \quad \text{and} \quad \phi = \tan^{-1} K_y/K_x \qquad (1.137)$$

then if $[h_x(\underline{r}),\ h_y(\underline{r}),\ h_z(\underline{r})] = (x, y, z)\, r^{n-1} e^{-cr}$ (1.139)

we find that

$$A_{\underline{K}}^{\underline{k}} (h_x, h_y, h_z) = (\sin\alpha\cos\phi,\ \sin\alpha\sin\phi,\ \cos\alpha)\, B_{\underline{K}}^{\underline{k}} \qquad (1.139)$$

where

$$B_{\underline{K}}^{\underline{k}} = \frac{4\pi i}{\sqrt{\Omega}} \left[\frac{\Gamma(n+1)\sin(n+1)\theta}{K\beta^{n+1}} - \frac{\Gamma(n+2)\cos(n+2)\theta}{K\beta^{n+2}} \right] \qquad (1.140)$$

if $K \neq 0$.

An attempt to apply the above formalism to L.C.A.O. wave functions for the valence bands of silicon did not prove very successful; an analysis of the reasons why is presented in Chapter four, where we consider the relationship between L.C.A.O. wave functions, our equivalent orbitals, and O.P.W. eigenfunctions for the valence bands of tetrahedral semiconductors.

CHAPTER 2

EQUIVALENT ORBITALS IN TETRAHEDRAL CRYSTALS.

2.1 EQUIVALENT ORBITALS AND MOLECULAR ORBITALS.

Introduction.

In the last chapter, we were concerned chiefly with molecular orbitals, which we defined to be those solutions of the Hartree-Fock equations 1.39 which made the matrix $\underline{\underline{\xi}}$ diagonal. More generally, if we introduce the one-electron Hamiltonian H defined by equation 1.80, the molecular orbitals are just the eigenfunctions of H. The lowest 1/2 N of these, subject to the symmetry conditions that we assumed, are each doubly-occupied in the ground state of the system, and so enter into equation 1.71 defining the spinless first-order density matrix $q(\underline{r}' | \underline{r})$ on which H depends; the other eigenfunctions are unoccupied or virtual molecular orbitals. Since H is invariant under the total symmetry group G of the system, the set of eigenfunctions of H having a given eigenvalue will span one or more of the irreducible representations of G. Hence, all the eigenfunctions of H (or molecular orbitals) can be so chosen that each of them behaves according to some row of some irreducible representation of G.

In this chapter, we consider equivalent orbitals. A set of equivalent orbitals[26] in a system can be defined to be "a set of ortho-normal one-electron functions, any one of which may be transformed into any other by an operation belonging to a certain subgroup +B of G and each of which is transformed into some other by every operation of B". In other words, a set of equivalent orbitals is closed under all the operations of some subgroup B of the symmetry group G of the system, and the whole set can be

generated from a single given member, $a_1(\underset{\sim}{r})$ say, by operating on it with all the elements of B. For instance, the set of Wannier functions appropriate to a given band in a crystal is a set of equivalent orbitals associated with the translation group T.

There are two distinct problems associated with equivalent orbitals. The first, which is by far the simpler, is to show their relationship to the eigenfunctions of H, and also to derive the equations that they must satisfy. The second concerns whether a given set of equivalent orbitals corresponds to a set of molecular orbitals which are all doubly occupied in the ground state of the crystal; this cannot in general be solved without calculating the energies of the molecular orbitals.

Derivation of Equivalent Orbitals from Molecular Orbitals.

The procedure suggested by Lennard-Jones [1] for finding the equivalent orbitals in a system, and for using them to find the one-electron energies, depends on the fact that the molecular orbitals can be classified according to the irreducible representations of B to which they belong. Suppose that the functions $a_i(\underset{\sim}{r})$ form a set of equivalent orbitals, and that Γ_j are the irreducible representations of B; let χ denote the character of $a_i(\underset{\sim}{r})$ and γ_j that of Γ_j with respect to the classes of the group. Then the existence of the set $a_i(\underset{\sim}{r})$ implies, according to group theory, that there exist c_j such that

$$\chi = \sum_j c_j \gamma_j \qquad (2.1)$$

where the coefficients c_j are integers or zero.

Let L be the number of elements in B, and let the matrices of $\sum c_j \Gamma_j$ in their full L-dimensional form be (E_L), (A_L), (B_L) ...

.... (P_L) with respect to the various classes of B. Let the corresponding matrices for the $a_i(\underset{\sim}{r})$ be (ε_L), (α_L), (β_L) .... (π_L); each of the latter contains L-1 elements equal to zero and one element equal to unity in each row and each column. According to group theory, there exists a unitary transformation operator $\underline{T}$, represented by the L-dimensional unitary matrix t say, such that

$$A_L = t\alpha_L \tilde{t}, \quad B_L = t\beta_L \tilde{t}, \text{ etc.} \tag{2.2}$$

If (ψ_i) denotes a set of molecular orbitals ψ_i with the representation $\Sigma c_j \Gamma_j$ then

$$(a_i) = \underline{T}\,(\psi_j). \tag{2.3}$$

Alternatively, if the orbitals $a_i(\underset{\sim}{r})$ are known, the ψ_i can be found by means of the transformation

$$(\psi_i) = \underline{T}^{-1}\,(a_j) \tag{2.4}$$

and the energies from a similar transformation (c.f. equation 1.40) which diagonalises the matrix $\underline{\underline{\varepsilon}}$.

In practice, it is not always easy to find the matrix t for a crystal, since L will normally be a multiple of the number M of unit cells in the crystal, and this may be of order 10^{22}. For this reason, it is easier to use the essentially similar formalism of section 2.2, which leads to an explicit formula for determining the molecular orbitals from the equivalent orbitals in a crystal. However, the above analysis is a convenient way of showing that the inverse transformation may exist; the requirement that the coefficients c_j in equation 2.1. must be integers or zero enables one to discover whether molecular orbitals of a certain symmetry can be derived from (or give rise to) a given type of equivalent orbital (27). As this problem does not arise for tetrahedral

semiconductors, we will not consider it further for the moment.

Since any given set of equivalent orbitals is related to a set of molecular orbitals by a unitary transformation, there exist molecular orbitals ψ_j such that, if we write

$$\underline{a} = (a_1, a_2, \ldots\ldots a_L), \quad \underline{\psi} = (\psi_1, \psi_2, \ldots\ldots \psi_L) \tag{2.5}$$

then $$\underline{a} = \underline{\psi} A \tag{2.6}$$

for some unitary matrix A. Hence, in accordance with equation 1.40

$$H_{\underline{a}} = \underline{a} (\hat{A} \underline{\underline{\beta}}), \tag{2.7}$$

where $\underline{\underline{\beta}}$ is a diagonal matrix of order L, or

$$H_{\underline{a}} = \underline{a} \underline{\underline{\varepsilon}}. \tag{2.8}$$

Hence $$H a_j = \sum_{i=1}^{L} a_i \varepsilon_{ij}. \tag{2.9}$$

Since the equivalent orbitals form an orthonormal set,

$$\langle a_j | H | a_i \rangle = \sum_{i=1}^{L} \varepsilon_{ik} \delta_{jk} = \varepsilon_{ji} \tag{2.10}$$

Also, from the definition of equivalent orbitals, if b_1, b_2 $\ldots . b_L$ are the elements of B and b_1 is the identity element, we can write

$$a_j(\underline{r}) = b_j a_1(\underline{r}) \tag{2.11}$$

so that $$H a_1(\underline{r}) = \sum_{j=1}^{L} \langle b_j a_1 | H | a_1 \rangle b_j a_1(\underline{r}) \tag{2.12}$$

For a given H, this equation depends only on $a_1(\underline{r})$, instead of depending on L different functions ϕ_j as it would for a general set of orbitals. It is for this reason that, apart perhaps from molecular orbitals, equivalent orbitals are in prineiple the simplest solutions of the Hartree-Fock equations (1.39) to find.

A Self-consistent Hamiltonian H.

In section 1.2, we showed that the one-electron Hamiltonian

operator $H(\underline{r})$, which is equivalent to the Hartree-Fock operator for orbitals occupied in the ground state of the system, depends on the occupied orbitals only through the spinless first-order density matrix

$$q(\underline{r}'|\underline{r}) = \sum_{i=1}^{\frac{1}{2}N} \phi_i^*(\underline{r}')\,\phi_i(\underline{r}). \qquad (2.13)$$

Suppose now that the equivalent orbitals a_1 to a_L are related by a unitary transformation to the first L functions $\phi_j(\underline{r})$. Then if we write

$$\underline{X} = (a_1, a_2, \ldots, a_L, \phi_{L+1}, \phi_{L+2}, \ldots \phi_{\frac{1}{2}N}), \quad \underline{\Phi} = (\phi_1, \phi_2, \ldots \phi_{\frac{1}{2}N}) \qquad (2.14)$$

we have that

$$\underline{X} = \underline{\Phi}\, A \qquad (2.15)$$

where A is a unitary matrix, and so

$$q(\underline{r}'|\underline{r}) = \underline{\Phi}^*(\underline{r}')\,\underline{\Phi}^\dagger(\underline{r}) = \underline{X}^*(\underline{r}')\left(A^{-1}\right)^*\left(A^{-1}\right)^\dagger \underline{X}^\dagger(\underline{r}) = \underline{X}^*(\underline{r}')\,\underline{X}^\dagger(\underline{r}). \qquad (2.16)$$

Hence

$$q(\underline{r}'|\underline{r}) = \sum_{i=1}^{L} a_i^*(\underline{r}')\,a_i(\underline{r}) + \sum_{i=L+1}^{\frac{1}{2}N} \phi_i^*(\underline{r}')\,\phi_i(\underline{r}). \qquad (2.17)$$

a result which could also be derived directly from equation 1.20.

Thus, provided it is known that a given set of equivalent orbitals is occupied in the ground state of the system, it is much easier to calculate H self-consistently, and so to find self-consistent solutions of the Hartree-Fock equations, by using them instead of using L independent molecular orbitals; this is especially true if L is large, as it normally is for a crystal. This leads us to the following conclusions. If H is assumed known, it is usually much easier to find its eigenfunctions, i.e. the molecular orbitals, than to solve equation 2.12 for the equivalent orbitals. If, however it is desired to calculate H self-consistently by an iterative procedure and it is known that one or more sets of equivalent orbitals are occupied in the ground state, then in

principle it is better to use equation 2.12 than to find the corresponding molecular orbitals since a much smaller number of equations will have to be solved at each iteration. In addition, it is sometimes possible to use a second-order convergence procedure with equivalent orbitals which allows for the change in H as $a_l(\underset{\sim}{r})$ changes; such a procedure is much more diffieult to apply to molecular orbitals.

Examples of Equivalent Orbitals in Crystals.

A simple example of sets of equivalent orbitals used in most accurate calculations of wave functions in crystals are the sets of atomic core functions; these are, according to our definition, just sets of orbitals equivalent under G or under the translation group T which is an invariant subgroup of G. Any set of molecular orbitals $\psi^{\underset{\sim}{k}}$ containing one funetion for each possible value of $\underset{\sim}{k}$ in the first Brillouin zone can also be transformed by a unitary transformation into a set of orbitals equivalent under T. For if we write

$$a^{\underset{\sim}{n}}(\underset{\sim}{r}) = \frac{1}{\sqrt{M}} \sum_{\underset{\sim}{k}} e^{-i\underset{\sim}{k}\cdot\underset{\sim}{a}} \psi^{\underset{\sim}{k}}(\underset{\sim}{r}), \qquad (2.18)$$

then the set of functions $a^{\underset{\sim}{n}}(\underset{\sim}{r})$ are orthonormal since

$$\sum_{\underset{\sim}{k}} e^{i\underset{\sim}{k}\cdot(\underset{\sim}{m}-\underset{\sim}{a})} = M\,\delta_{\underset{\sim}{m},\underset{\sim}{a}}. \qquad (2.19)$$

According to equation 1.59

$$\psi^{\underset{\sim}{k}}(\underset{\sim}{r}-\underset{\sim}{a}) = e^{-i\underset{\sim}{k}\cdot\underset{\sim}{a}} \psi^{\underset{\sim}{k}}(\underset{\sim}{r}), \qquad (2.20)$$

and so we find that

$$a^{\underset{\sim}{a}}(\underset{\sim}{r}) = \frac{1}{\sqrt{M}} \sum_{\underset{\sim}{k}} \psi^{\underset{\sim}{k}}(\underset{\sim}{r}-\underset{\sim}{m}) = a^{\underset{\sim}{a}}(\underset{\sim}{r}-\underset{\sim}{a}). \qquad (2.21)$$

Thus, the functions $a^{\underset{\sim}{n}}(\underset{\sim}{r})$ form a set of orbitals equivalent underT. They are in faet just the Wannier functions corresponding to the molecular orbitals $\psi^{\underset{\sim}{k}}$ and our analysis shows that any set of orbitals corresponding to a filled band in a crystal, i.e. to a

band in which for each $\underset{\sim}{k}$ there is a doubly-occupied orbital with translational symmetry $\underset{\sim}{k}$, can be replaced by an equivalent set of Wannier functions. If however in the ground state of the crystal there is one band which is only partly filled, it will not in general be possible to find a set of Wannier functions, i.e. functions equivalent under the full translation group T, corresponding to the occupied orbitals in this band. This is the position in crystals of monovalent metals such as potassium and sodium, for instance. If H is known, it is still possible to solve equation 2.12 and obtain a set of Wannier functions from which the highest occupied and lowest unoccupied molecular orbitals can be derived, but it is not clear whether either this or the inverse process represented by equation 2.18 is very useful in such a case. If it did not involve too much work, it is possible that such a process might be useful for finding the relevant effective masses, since as we show in section 2.2

$$E^{\underset{\sim}{k}} = \sum_{\underset{\sim}{n}} e^{i\underset{\sim}{k}.\underset{\sim}{n}} \langle a^{\underset{\sim}{n}}(\underset{\sim}{r}) \,|H|\, a^{\underset{\sim}{0}}(\underset{\sim}{r}) \rangle \qquad (2.22)$$

and this expression is very easy to differentiate with respect to $\underset{\sim}{k}$. These Wannier functions are related to Pauling's theory of resonating bonds in a metal.[28]

For tetrahedral semiconductors, with which we are principally concerned in this thesis, the position is fortunately very much simpler. All the experimental evidence suggests that in diamond, silicon, germanium, grey tin, the inter-metallic III - V compounds, and possibly also in II - VI and I - VII compounds having the zincblende structure, there are four full valence bands separated by a finite energy gap from the conduction bands, and that the symmetries of these bands are compatible with the existence of equivalent orbitals having the same symmetry as the bonds. These conclusions are supported by those molecular orbital calculations

that have been carried out. It should be noted that not all symmetry types, i.e. irreducible represenatations of G to which molecular orbitals can belong, are compatible with equivalent orbitals having this symmetry; in the case of the zincblende structure, representations derived from the 2-dimensional irreducible representation of T_d are not compatible with such bond orbitals. However, Hall (27) has suggested that whereever chemical bonds are expected in classical valence theory there will exist a set of equivalent orbitals, having the same symmetry as the bonds, which is completely occupied in the ground state of the system, and we assume that this is true for the crystals that we are considering.

Springium.

The relationship between molecular and equivalent orbitals in a crystal is very similar to that between normal co-ordinates and local co-ordinates for a symmetric vibrating system in classical mechanics. Since the latter system is conceptually simpler, as it does not involve the problems associated with anti-symmetry and the exclusion principle, we will use it to obtain a picture of the relationship between these two types of orbitals. Rather than consider the very complicated case of the dynamics of an actual crystal lattice, we take for our model the simple one-dimensional substance "springium", which we now describe.

A single bit of springium consists of a perfect light spring with sticky masses $1/2\ m_1$ and $1/2\ m_2$ attached to its ends. Two bits can combine to form a linear molecule, in which either there is a mass $1/2(m_1 + m_2)$ at the centre with masses $1/2\ m_1$ and $1/2\ m_2$ at the ends, or there is a mass m_j at the centre and $1/2\ m_k$ at

each end. Similarly, a large number of bits can combine to form chains, planar networks, etc. The two simplest linear chains are the homogeneous one, in which there is a mass of $1/2(m_1 + m_2)$ at the join of each pair of bits, and the heterogeneous one of period two in which masses m_1 and m_2 occur alternately. For convenience, we consider the homogeneous case, which could be said to correspond to a homonuclear crystal. If the chain is long enough, the conditions satisfied by the end bits are not very important for most purposes, and it is convenient to consider the chain as forming a circle containing M bits, where M is very large, with the back of the M'th bit of springium adhering to the front of the first bit. We then speak of cyclic boundary conditions, and a homogeneous chain of springium satisfying these conditions will be called springcom.

A single bit of springium possesses various excited states, corresponding to different amplitudes of free oscillation excited in the spring. Similarly, there are a large number of possible excited states of springcom. These can most simply be described either by using local co-ordinates such as the extension of each bit or the position of each mass or by using normal co-ordinates. We can require that the possible state of springcom be quantised in such a way that as t tends to infinity the mean value of $\frac{1}{t}\int_0^t$(potential energy of system) dt' shall be an integral multiple of a certain fundamental quantity. It follows that if each normal mode is separately quantised then the superposition of any number of normal modes will also be; in addition, a definite energy, related to its frequency, can be associated with each normal mode, at least for small oscillations in which the different normal modes are essentially independent of each other. With local co-ordinates

the problem is more complicated, since they are not independent of each other and quantised states of the system as a whole do not necessarily involve quantised excitations in each bit separately.

If normal co-ordinates are used, any state of the system can to a first approximation be regarded as the superposition of a number of independent modes, each of which behaves as if it were due to a single bit having a different natural frequency, and the detailed structure of the chain is only relevant for determining the frequency spectrum. In addition, many of the possible quantised states of the system will be able to be described as the superposition of quantised excitations of a set of normal modes. This type of description is very suitable for dealing with small changes in the system such as might be produced by an external field. The ease of performing calculations with normal co-ordinates might seem at first sight to be so great that there is never any advantage in using local co-ordinates. However, one disadvantage of normal co-ordinates is that they do not make it easy to compare single bits or small molecules of springium with springeom, or to compare types of springeom with different spring constants or masses; for such comparisons, local co-ordinates are far better. There is one type of state for which the difficulties associated with quantisation do not arise for local co-ordinates, namely a state in which all the normal modes, and all the bits, have the same quantised excitation. In such a state it is possible to find an explicit equation in just one set of local co-ordinates for the behaviour of a typical bit, which compares quite favourably with having to solve equations for all the M normal co-ordinates. In addition, in such a state it is no more meaningful to state that all the normal modes are equally excited than to state that all the bits of springium are. The two descriptions are entirely equivalent and equally valid. Which

description to use will depend on whether one is more interested in the properties of springcom while it is in this state or in transitions to and from such a state; local co-ordinates are more convenient for the former, and ~~local~~ normal co-ordinates for the latter.

The electronic states in molecules and crystals correspond in many ways to the excitations of springcom, since in each case the mathematical problem concerns quantised combinations of the eighnfunctions of a Hamiltonian which is invariant under a given symmetry group G. It is often convenient to describe the state of a system by the molecular orbitals which are occupied, although this is strictly justified only for states which can be represented by a single Slater determinant and is not a necessary consequence of quantisation or anti-symmetry requirements; similarly in springcom it is sufficient, but not necessary, that each normal mode be separately quantised. In fact, the molecular orbitals in an atomic system correspond fairly closely to the normal co-ordinates in springcom, and have most of their advantages and disadvantages. A periodic potential in a crystal leads to bands of one-electron states, usually separated by finite energy gaps. A state in which each band is either completely full or completely empty corresponds to a state of springcom in which all the bits are equally excited; we note that our simple model, being one-dimensional, can only reproduce the translational symmetry of a crystal. Hence in such a state, which is the ground state of insulating and semi-conducting crystals, molecular orbital and equivalent orbital descriptions are equally valid. In the corresponding states of springcom, it is not more meaningful to state that there is a unit excitation in each normal mode than to associate a unit excitation with each bit of springium. Similarly, in the ground states of these crystals it is a matter of choice whether we prefer to think of each electron as localised in a bond or as spread over the whole

crystal, and our choice will usually depend on what properties of the system are being considered. Thus, if, as is often the case in physics, we are interested in the effect of non-localised perturbations such as electromagnetic fields, molecular orbitals with electrons spread over the whole crystal are probably to be preferred, and this is also the more natural scheme for considering such properties as electrical conductivity. If on the other hand we are concerned with the charge distribution in the ground state or the relationship between molecules and crystals, as is often the case in chemistry, equivalent orbitals with localised electrons would seem to provide the more useful approach.

2.2 A VARIATIONAL TREATMENT OF EQUIVALENT ORBITALS.

Derivation of the Equivalent Orbital Equation.

It is apparent that equation 2.12 is just the equation that would be obtained by requiring the functional

$$X' = \int a_i^*(\underset{\sim}{r})\, H(\underset{\sim}{r})\, a_i(\underset{\sim}{r})\, d\underset{\sim}{r} \qquad (2.23)$$

to be stationary subject to the auxiliary conditions

$$\int a_i^*(\underset{\sim}{r})\, b_j\, a_i(\underset{\sim}{r})\, d\underset{\sim}{r} = \delta_{ji} \quad , \quad j = 1, 2, \ldots L, \qquad (2.24)$$

where $H(\underset{\sim}{r})$ is assumed to be independent of $a_i(\underset{\sim}{r})$. This fact is the basis of the method proppesed by Koster for treating generalised Wannier functions or equivalent orbitals.[23] However, it is not necessary to assume that $H(\underset{\sim}{r})$ is independent of $a_i(\underset{\sim}{r})$, and in practice it is not desirable to do so if a self-consistent H is required. We will now show that equation 2.12 is the one that would be derived from minimising the total energy of the system, as in the derivation of the Hartreee-Fock equations, provided it is assumed that the density matrix

$q(\underset{\sim}{r}^1 | \underset{\sim}{r})$ is invariant under G (or at least under B, the subgroup of G which generates the set of equivalent orbitals being considered.)

As previously, we have to make stationary the functional

$$\langle X \rangle = \langle H \rangle - \sum_{ij} \varepsilon_{ij} \int \phi_j^*(\xi)\, \phi_i(\xi)\, d\xi \quad . \tag{2.25}$$

On re-writing equation 1.29 by making use of $q(\underset{\sim}{r}^1 | \underset{\sim}{r})$, we have that

$$\begin{aligned}
\delta X = {} & 2 \sum_i \left[\int \delta\phi_i^*(\xi)\, H_1(\xi)\, \phi_i(\xi)\, d\xi + \int \phi_i^*(\xi)\, H_1(\xi)\, \delta\phi(\xi)\, d\xi \right] \\
& + 4 \sum_i \iint \left[\delta\phi_i^*(\xi)\, \phi_i(\xi) + \phi_i^*(\xi)\, \delta\phi_i(\xi) \right] \frac{q(\xi'|\xi')}{|\xi - \xi'|}\, d\xi'\, d\xi \\
& - 2 \sum_i \iint \left[\delta\phi_i^*(\xi)\, \phi_i(\xi')\, \frac{q(\xi'|\xi)}{|\xi - \xi'|} + \delta\phi_i(\xi)\, \phi_i^*(\xi')\, \frac{q(\xi|\xi')}{|\xi - \xi'|} \right] d\xi'\, d\xi \\
& - 2 \sum_{ij} \varepsilon_{ij} \left[\int \delta\phi_i^*(\xi)\, \phi_j(\xi)\, d\xi + \int \phi_j^*(\xi)\, \delta\phi_i(\xi)\, d\xi \right] .
\end{aligned} \tag{2.26}$$

Previously all the $\delta\phi_j$ were independent, so that the coefficient of each could separately be put equal to zero. Now, however, for the first L orbitals ϕ_j which are a set of equivalent orbitals

$$\phi_j(\xi) = b_j\, \phi_1(\xi) , \tag{2.27}$$

and so

$$\delta\phi_j(\xi) = b_j\, \delta\phi_1(\xi) . \tag{2.28}$$

For convenience, we write $a_j(\underset{\sim}{r})$ instead of $\phi_j(\xi)$ for $j = 1,2, \ldots L$.

However, since $H_1(\underset{\sim}{r})$ and $q(\underset{\sim}{r}^1 | \underset{\sim}{r})$ are invariant under all the elements of B, we can write for the change in X when only the orbitals $a_j(r)$ are varied

$$\delta X_a = 2L \left[\int \delta a_1^*(\xi)\, H_1(\xi)\, a_1(\xi)\, d\xi + \int a_1^*(\xi)\, H_1(\xi)\, \delta a_1(\xi)\, d\xi \right]$$

$$+ 4L \iint \left[\delta a_1^*(\xi)\, a_1(\xi) + \delta a_1(\xi)\, a_1^*(\xi) \right] \frac{(\xi'|\xi')}{|\xi - \xi'|}\, d\xi'\, d\xi$$

$$- 2L \iint \left[\delta a_1^*(\xi)\, a_1(\xi') \frac{q(\xi'|\xi)}{|\xi - \xi'|} + \delta a_1(\xi)\, a_1^*(\xi') \frac{q(\xi'|\xi')}{|\xi - \xi'|} \right] d\xi'\, d\xi$$

$$- 2L \sum_{j=1}^{L} \varepsilon_{j1} \left[\int \delta a_1^*(\xi)\, a_j(\xi)\, d\xi + \int a_j^*(\xi)\, \delta a_1(\xi)\, d\xi \right]$$

$$- 2L \sum_{i=1}^{L} \sum_{j=L+1}^{\frac{1}{2}N} \varepsilon_{ji} \left[\int \delta a_i^*(\xi)\, \phi_j(\nu, \xi)\, d\xi + \int \delta a_i(\xi)\, \phi_j^*(\nu, \xi)\, d\xi \right] \quad (2.29)$$

The penultimate term is of the form shown since all the elements can be generated by multiplying any given element, b_p say, by each of the b_i in turn, while if $b_t = b_j^{-1}$ and ε_{j1} is real then $\varepsilon_{t1} = \varepsilon_{j1}$.

The last term in δX_a will be zero if our functions are so chosen that

$$\varepsilon_{ji} = 0 \quad \text{for } i \leq L, \quad j \geq L+1 \quad (2.30)$$

In such a case, on equating the coefficient of $\delta a_1^*(\xi)$ to zero, we find that

$$H_1(\xi)\, a_1(\xi) + 2 \int \frac{q(\xi'|\xi')}{|\xi - \xi'|}\, d\xi'\, a_1(\xi) - \int \frac{q(\xi'|\xi)}{|\xi - \xi'|}\, a_1(\xi')\, d\xi' = \sum_{j=1}^{L} \varepsilon_{j1}\, a_j(\xi) \quad (2.31)$$

i.e.
$$H(\xi)\, a_1(\xi) = \sum_{j=1}^{L} \varepsilon_{j1}\, a_j(\xi) \quad (2.32)$$

where $H(\underset{\sim}{r})$ is as defined by equation 2.30. Equation 2.32 is equivalent to equation 2.9, and provides an additional justification for the latter.

If equation 2.32 is satisfied, the requirements of equation 2.30 will also be satisfied if

$$\langle a_i | \phi_j \rangle = 0 \quad \text{for } 1 \leq i \leq L, \; L+1 \leq j \leq \tfrac{1}{2}N \quad (2.33)$$

which is an orthogonality condition which is always assumed for solutions of the Hartree-Fock equations; for in such a case

$$\varepsilon_{ji} = \langle \phi_j | H | a_i \rangle = \langle \phi_j | b_i | a_i \rangle = \langle \phi_j | b_i | \sum_{s=1}^{L} \varepsilon_{si} a_s \rangle$$

$$= \langle \phi_j | \sum \varepsilon_{si} a_s \rangle \text{ where } b_r = b_i \delta_{rs} = 0$$

in view of equation 2.30.

In many practical applications, the functions $\phi_j(\underset{\sim}{r})$ for $j \geq L+1$ will just be atomic core functions, so that if one restricts oneself to the Hilbert space of functions orthogonal to these core functions equation 2.32 can be used to find the valence band orbitals. It is worth noting that the above analysis can be used either to find $a_1(\underset{\sim}{r})$ if all the ϕ_j are known or to simplify $H(\underset{\sim}{r})$ and the orthogonality conditions for the ϕ_j if $a_1(\underset{\sim}{r})$ is known. This latter case could occur, for instance, if one wished to find the conduction bands of tetrahedral semiconductors after finding the equivalent orbitals for the valence bands, and is similar to the method that Redei attempted to use for diamond.[4]

It is sometimes convenient to split the Hamiltonian $H(\underset{\sim}{r})$ into parts depending on the nuclei, the equivalent orbitals being considered, and all the other occupied orbitals. For this purpose we write

$$q(\underset{\sim}{r}'|\underset{\sim}{r}) = q_a(\underset{\sim}{r}'|\underset{\sim}{r}) + q_b(\underset{\sim}{r}'|\underset{\sim}{r}) \qquad (2.34)$$

where

$$q_a(\underset{\sim}{r}'|\underset{\sim}{r}) = \sum_{i=1}^{L} a_i^*(\underset{\sim}{r}') a_i(\underset{\sim}{r}) \qquad (2.35)$$

and

$$q_b(\underset{\sim}{r}'|\underset{\sim}{r}) = \sum_{j=L+1}^{N} \phi_j^*(\underset{\sim}{r}') \phi_j(\underset{\sim}{r}) \qquad (2.36)$$

We can then write

$$H(\underset{\sim}{r}) = H_1(\underset{\sim}{r}) + H_a(\underset{\sim}{r}) + H_b(\underset{\sim}{r}) \qquad (2.37)$$

where $H_1(\underset{\sim}{r}) = -\frac{1}{2}\nabla^2 - \sum_{A=1}^{L'} \frac{Z_A}{|\underset{\sim}{r} - \underset{\sim}{r}_A|}$ as previously, (2.38)

$$H_a(\underline{r})\,\phi(\underline{r}) = 2\int \frac{\rho_a(\underline{r}'|\underline{r})}{|\underline{r}-\underline{r}'|}\,d\underline{r}'\,\phi(\underline{r}) - \int \frac{\rho_a(\underline{r}'|\underline{r})}{|\underline{r}-\underline{r}'|}\,\phi(\underline{r}')\,d\underline{r}' \qquad (2.39)$$

and $$H_b(\underline{r})\,\phi(\underline{r}) = 2\int \frac{\rho_b(\underline{r}'|\underline{r})}{|\underline{r}-\underline{r}'|}\,d\underline{r}'\,\phi(\underline{r}) - \int \frac{\rho_b(\underline{r}'|\underline{r})}{|\underline{r}-\underline{r}'|}\,\phi(\underline{r}')\,d\underline{r}' \qquad (2.40)$$

Then if the other orbitals are assumed known, the problem of finding equivalent orbitals reduces to that of finding a set of functions orthogonal to these known orbitals and such that the functional

$$X_i = \int a_i^*(\underline{r})\left[H_1(\underline{r}) + H_2(\underline{r})\right] a_i(\underline{r})\,d\underline{r} + 2\sum_{j=1}^{L}\iint \frac{a_i^*(\underline{r})\,a_j^*(\underline{r}')\,a_j(\underline{r}')\,a_i(\underline{r})}{|\underline{r}-\underline{r}'|}\,d\underline{r}'\,d\underline{r}$$
$$- \sum_{j=1}^{L}\iint \frac{a_i^*(\underline{r})\,a_j^*(\underline{r}')\,a_j(\underline{r})\,a_i(\underline{r}')}{|\underline{r}-\underline{r}'|}\,d\underline{r}'\,d\underline{r} - \sum_{j=1}^{L}\varepsilon_{ji}\int a_j^*(\underline{r})\,a_i(\underline{r})\,d\underline{r} \qquad (2.41)$$

is stationary, where

$$\varepsilon_{ji} = \langle a_j | H | a_i \rangle . \qquad (2.42)$$

Derivation of Molecular Orbitals from Equivalent Orbitals.

Before we describe how equations 2.41 and 2.42 can be used to find the equivalent orbitals in tetrahedral crystals, it is convenient to complete the general theory by showing how the molecular orbitals can be found once the equivalent orbitals are known. As previously, we assume that G is the space group of the crystal, and that the orbitals are equivalent under the subgroup B which is of order L. Our treatment is similar to that proposed by Koster,[23] but for convenience in our later work use the irreducible representations of B rather than those of G; it is obvious that the theory applies if B = G, since each group is, trivially, a subgroup of itself.

Since the Hamiltonian H is invariant under B, the set of eigenvectors of H having a given eigenvalue will span one

or more of the irreducible representations of B. Let g_l be the dimension of the l'th irreducible representation of B, and let B^{li}_{mn} be the (m, n)th element of the matrix B^{li} which is equivalent to b_i in this representation. If we operate on both sides of equation 2.32 with b_i, we find that

$$H a_j(\xi) = \sum_{j=1}^{L} \varepsilon_{jl} b_i a_j(\xi), \qquad (2.43)$$

and hence that

$$H \sum_{i=1}^{L} B^{li}_{mn} a_j(\xi) = \sum_{j=1}^{L} \varepsilon_{jl} \sum_{i=1}^{L} B^{li}_{mn} b_i a_j(\xi). \qquad (2.44)$$

Now let $$a_j(\xi) = \sum_{p,q,r} c_{pqr} \phi^{pr}_q(\xi), \qquad (2.45)$$

where the functions ϕ^{pr}_q transform according to the q'th row of the p'th irreducible representation of B, and r labels the different functions having this property, so that

$$b_i \phi^{pr}_q = \sum_{q'} B^{pi}_{q'q} \phi^{pr}_{q'}. \qquad (2.46)$$

Then $$\left(\sum_{i=1}^{L} B^{li}_{mn} b_i \right) a_j = \sum_{p,q,r} c_{pqr} \sum_{q'} B^{lj}_{q'q} \left(\sum_{q''} \sum_{i=1}^{L} B^{li}_{mn} B^{pi}_{q''q'} \phi^{pr}_{q''} \right). \qquad (2.47)$$

Now it is well known from group theory that

$$\sum_{i=1}^{L} B^{li}_{mn} B^{pi}_{q''q'} = \frac{L}{g_l g_p} \delta_{pl} \delta_{q''m} \delta_{q'n}. \qquad (2.48)$$

Hence $$\sum_{i=1}^{L} B^{li}_{mn} b_i a_j = \frac{L}{g_l} \sum_{q} B^{lj}_{nq} \sum_{r} c_{lqr} \phi^{lr}_m. \qquad (2.49)$$

Since b_1 is the identity element,

$$B^{1j}_{nq} = \delta_{nq}, \qquad (2.50)$$

and so, on putting j = 1 in equation 2.49 we find that

$$\sum_{i=1}^{L} B^{li}_{mn} a_i(\xi) = \frac{L}{g_l} \sum_{r} c_{lnr} \phi^{lr}_m(\xi). \qquad (2.51)$$

Let us now introduce the functions

$$f^{l}_{mn} = \sum_{r} c_{lnr}\, \phi^{lr}_{m} \qquad (2.52)$$

Then, using in turn equations 2.52, 2.51, 2.44 and 2.49, we find that

$$H f^{l}_{mn}(\underset{\sim}{r}) = \sum_{j=1}^{L} \varepsilon_{jl} \sum_{q} B^{lj}_{nq}\, f^{l}_{mq}(\underset{\sim}{r}) \qquad (2.53)$$

We now introduce the row vector

$$\underline{f}^{l}_{m} = (f^{l}_{m1}, f^{l}_{m2}, \ldots\ldots f^{l}_{m,g_l}) \qquad (2.54)$$

Then

$$H \underline{f}^{l\,+}_{m} = \sum_{j=1}^{L} \varepsilon_{jl}\, B^{lj}\, \underline{f}^{l\,+}_{m} \qquad (2.55)$$

Thus, if C is a unitary matrix such that

$$C \Big(\sum_{j=1}^{L} \varepsilon_{jl}\, B^{lj} \Big) \tilde{C} = E \qquad (2.56)$$

where E is diagonal, then

$$H C \underline{f}^{l\,+}_{m}(\underset{\sim}{r}) = E C \underline{f}^{l\,+}_{m}(\underset{\sim}{r}) \qquad (2.57)$$

Hence, the eigenvalues of H are just the diagonal elements of E, and these in turn are the eigenvalues of the matrix

$$S_l = \sum_{j=1}^{L} \varepsilon_{jl}\, B^{lj} \qquad (2.58)$$

while the eigenvectors of H are the vectors $C \underline{f}^{l\,+}_{m}$

Thus, once the Lagrangian multipliers ε_{jl} have been determined, the eigenvalues of H can easily be found. In some cases, the group B will be Abelian, in which case S_l will automatically be diagonal; most previous equivalent orbital calculations have been concerned with such cases. For instance, the translation group T is Abelian, so that if $a_l(\underset{\sim}{r})$ is a Wannier function and $\varepsilon_{\underset{\sim}{n}} = \langle a_l(\underset{\sim}{r}-\underset{\sim}{n}) \,|H|\, a_l(\underset{\sim}{r}) \rangle$ then the one-electron energy for the state with translational symmetry $\underline{k}$ is just $\sum_{\underset{\sim}{n}} e^{-i\underline{k}\cdot\underset{\sim}{n}} \varepsilon_{\underset{\sim}{n}}$ as stated in equation 2.22.

Uniqueness and Degeneracy.

So far, we have only assumed that B is a subgroup of G. If we now require that B shall be an invariant subgroup, then we can divide G into its cosets with respect to B. There will be s such cosets, where s is the quotient of the orders of G and B, and we take as their representatives the elements g_1 = the identity, g_2, g_3, ... g_s. Then, operating on both sides of equation 2.12 with g_1, we have that

$$g_l H a_1(\underset{\sim}{r}) = \sum_{k=1}^{L} \langle b_k a_1 | H | a_1 \rangle g_l b_k a_1(\underset{\sim}{r}), \qquad (2.59)$$

and also, since G and H commute, that

$$g_l H a_1(\underset{\sim}{r}) = H a_1(g_l^{-1}\underset{\sim}{r}) = \sum_{j=1}^{L} \langle b_j a_1 | H | a_1 \rangle b_j a_1(g_l^{-1}\underset{\sim}{r})$$

$$= \sum_{j=1}^{L} \langle b_j a_1 | H | a_1 \rangle b_j g_l a_1(\underset{\sim}{r}). \qquad (2.60)$$

Now since B is invariant, there exists for each b_j and g_1 an element b_k of B such that

$$b_j g_l = g_l b_k . \qquad (2.61)$$

Hence, from equation 2.60,

$$g_l H a_1(\underset{\sim}{r}) = \sum_{k=1}^{L} \langle b_j a_1 | H | a_1 \rangle g_l b_k a_1(\underset{\sim}{r}). \qquad (2.62)$$

Since the functions $b_j a_1(\underset{\sim}{r})$ form an orthonormal set, we have from equations 2.59 and 2.62 that

$$\langle b_k a_1 | H | a_1 \rangle = \langle b_j a_1 | H | a_1 \rangle . \qquad (2.63)$$

But

$$\langle b_k a_1 | H | a_1 \rangle = \langle g_l^{-1} b_j g_l a_1 | H | a_1 \rangle = \langle b_j g_l a_1 | H | g_l a_1 \rangle , \qquad (2.64)$$

and so, on substituting this in equation 2.60 we find that

$$H g_l a_1(\underset{\sim}{r}) = \sum_{j=1}^{L} \langle b_j g_l a_1 | H | g_l a_1 \rangle b_j g_l a_1(\underset{\sim}{r}). \qquad (2.65)$$

Thus, $g_1 a_1(\underset{\sim}{r})$ satisfies equation 2.12 if $a_1(\underset{\sim}{r})$ does, and also has the same Lagrangian parameters ε_{j1}. Hence, unless the set of equivalent orbitals $a_j(\underset{\sim}{r})$ is degenerate, there must exist for each 1 a j such that

$$g_i a_1(\underset{\sim}{r}) = e^{im_i} a_j(\underset{\sim}{r}), \qquad (2.66)$$

where e^{im_i} is just a phase factor and can normally be omitted if H is real. In such a case, we can choose our elements g_1 in such a way that, in the absence of degeneracy,

$$g_i a_1(\underset{\sim}{r}) = a_i(\underset{\sim}{r}), \quad i=1,2,\ldots s. \qquad (2.67)$$

Hence if G be expressed as the direct product of an invariant subgroup B and some other group Q, and if there is reason to believe that a set of orbitals $a_j(\underset{\sim}{r})$ equivalent under B is non-degenerate and satisfies equation 2.12, then we can require that $a_1(\underset{\sim}{r})$ be invariant under Q. For a crystal having a symmorphic space group, B will normally contain the translation group T and so Q will be a subgroup of the point group P.

2.3 THE FORMALISM FOR TETRAHEDRAL SEMICONDUCTORS.

Formal Problem.

As mentioned in section 2 of Chapter One, the zincblende structure has a symmorphic space group G which is the direct product of the translation group T and the tetrahedral group T_d. G has an invariant subgroup B which is the direct product of the dihedral group D_{2h} and T, while G is the direct product of this

subgroup B with C_{3v}. If we label the elements of D_{2h}

$$\beta_0 = \varepsilon, \quad \beta_1 = \delta_{1x}, \quad \beta_2 = \delta_{1y}, \quad \beta_3 = \delta_{1z}, \tag{2.68}$$

then a general element of B is

$$b_j = (\beta_i | \underset{\sim}{n}) \tag{2.69}$$

Since each element β_i of D_{2h} is its own inverse, we have from equation 1.46 that

$$(\beta_i | \underset{\sim}{n})^{-1} = (\beta_i | -\beta_i \underset{\sim}{n}) \tag{2.70}$$

This subgroup B is such that if we start with a given band $a_1(\underset{\sim}{r})$ the set of operations b_j will generate all the bonds in the crystal. The equivalent orbitals $a_j(\underset{\sim}{r})$ generated in this way will henceforth be called bond orbitals, a name suggested by Professor Coulson in a paper[29] in which he considered L.C.A.O. approximations to these orbitals. If we write

$$\varepsilon_{j1} = \langle a_j | H | a_1 \rangle = \langle (\beta_i | \underset{\sim}{n}) a_1 | H | a_1 \rangle = \varepsilon_i(\underset{\sim}{n}) \tag{2.71}$$

then equation 2.32 for the orbital $a_1(\underset{\sim}{r})$ can be written

$$H a_1(\underset{\sim}{r}) = \sum_{i=0}^{3} \sum_{\underset{\sim}{n}} \varepsilon_i(\underset{\sim}{n}) (\beta_i | \underset{\sim}{n}) a_1(\underset{\sim}{r}) \tag{2.72}$$

and, as shown in the last section, a_1 will make stationary the functional

$$X_1 = \langle a_1 | H | a_1 \rangle - \sum_{i,\underset{\sim}{n}} \varepsilon_i(\underset{\sim}{n}) \langle (\beta_i | \underset{\sim}{n}) a_1 | a_1 \rangle \tag{2.73}$$

We assume that the bond orbitals are not degenerate, in which case the results of the last section show that they will be invariant under the group C_{3v} about a bond.

This solves, in principle, the problem of finding bond orbitals for the valence bands of tetrahedral semiconductors. In the next chapter we describe how we

attempted to apply this theory to a whole range of such crystals, and in chapter four we present and discuss the results we obtained. Practical calculations must involve a number of approximations, such as the expansion of $a_i(\underset{\sim}{r})$ in terms of a finite number of basis functions, but the theory we have developed so far is an exact one. As larger and faster computers become available, it should therefore be possible to perform much more accurate calculations without any alterations in the basic method.

One-electron Energies.

The irreducible representations of B are mostly of order four, since those of T are one-dimensional and D_{2h} contains four elements. For a given translational symmetry $\underset{\sim}{k}$, the matrices appropriate to the elements of B are as follows:

for $(\beta_0|\underset{\sim}{n}) \equiv (\varepsilon|\underset{\sim}{n})$,
$$\begin{pmatrix} e^{-i\underset{\sim}{k}\cdot\underset{\sim}{n}} & 0 & 0 & 0 \\ 0 & e^{-i\beta_1\underset{\sim}{k}\cdot\underset{\sim}{n}} & 0 & 0 \\ 0 & 0 & e^{-i\beta_2\underset{\sim}{k}\cdot\underset{\sim}{n}} & 0 \\ 0 & 0 & 0 & e^{-i\beta_3\underset{\sim}{k}\cdot\underset{\sim}{n}} \end{pmatrix};$$

for $(\beta_1|0)$,
$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$
and for $(\beta_1|\underset{\sim}{n})$
$$\begin{pmatrix} 0 & e^{-i\underset{\sim}{k}\cdot\underset{\sim}{n}} & 0 & 0 \\ e^{-i\beta_1\underset{\sim}{k}\cdot\underset{\sim}{n}} & 0 & 0 & 0 \\ 0 & 0 & 0 & e^{-i\beta_2\underset{\sim}{k}\cdot\underset{\sim}{n}} \\ 0 & 0 & e^{-i\beta_3\underset{\sim}{k}\cdot\underset{\sim}{n}} & 0 \end{pmatrix};$$

for $(\beta_2|\underline{0})$, $\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$ and for $(\beta_2|\underline{n})$ $\begin{pmatrix} 0 & 0 & e^{-i\underline{k}\cdot\underline{n}} & 0 \\ 0 & 0 & 0 & e^{-i\beta_1\underline{k}\cdot\underline{n}} \\ e^{-i\beta_2\underline{k}\cdot\underline{n}} & 0 & 0 & 0 \\ 0 & e^{-i\beta_3\underline{k}\cdot\underline{n}} & 0 & 0 \end{pmatrix}$;

for $(\beta_3|\underline{0})$, $\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$ and for $(\beta_3|\underline{n})$ $\begin{pmatrix} 0 & 0 & 0 & e^{-i\underline{k}\cdot\underline{n}} \\ 0 & 0 & e^{-i\beta_1\underline{k}\cdot\underline{n}} & 0 \\ 0 & e^{-i\beta_2\underline{k}\cdot\underline{n}} & 0 & 0 \\ e^{-i\beta_3\underline{k}\cdot\underline{n}} & 0 & 0 & 0 \end{pmatrix}$.

Thus, on substituting in equation 2.58 we find that the eigenvalues of H associated with a given $\underset{\sim}{k}$ are just the eigenvalues of

$$S_{\underline{k}} = \begin{pmatrix} \xi_0(\underline{k}) & \xi_1(\underline{k}) & \xi_2(\underline{k}) & \xi_3(\underline{k}) \\ \xi_1(\beta_1\underline{k}) & \xi_0(\beta_1\underline{k}) & \xi_3(\beta_1\underline{k}) & \xi_2(\beta_1\underline{k}) \\ \xi_2(\beta_2\underline{k}) & \xi_3(\beta_2\underline{k}) & \xi_0(\beta_2\underline{k}) & \xi_1(\beta_2\underline{k}) \\ \xi_3(\beta_3\underline{k}) & \xi_2(\beta_3\underline{k}) & \xi_1(\beta_3\underline{k}) & \xi_0(\beta_3\underline{k}) \end{pmatrix} \qquad (2.74)$$

where

$$\xi_j(\underline{k}') = \sum_{\underline{n}} e^{-i\underline{k}'\cdot\underline{n}} \, \xi_j(\underline{n}) \,, \quad j = 0, 1, 2, 3. \qquad (2.75)$$

There are four linearly independent eigenfunctions, and so four different molecular orbitals, for each value of $\underset{\sim}{k}$, giving a total of $4M$ orbitals for a crystal containing M unit cells; if each of these orbitals is doubly occupied in the ground state of the crystal, all the $8M$ valence electrons in the crystal will be accommodated.

It is interesting to note that the elements of $S_{\underset{\sim}{k}}$ are the same as the matrix elements of H between the functions

$$\psi_j^{\underset{\sim}{k}} = \frac{1}{\sqrt{M}} \sum_{\underset{\sim}{n}} e^{i\underset{\sim}{k}\cdot\underset{\sim}{n}} (\beta_j|\underset{\sim}{n})\, a_1(\underset{\sim}{r}), \quad j = 0,1,2,3. \qquad (2.76)$$

This follows since, by analogy with equation 1.121

$$\langle \psi_i^{\underset{\sim}{k}} | \psi_j^{\underset{\sim}{k}} \rangle = \sum_{\underset{\sim}{l}} e^{i\underset{\sim}{k}\cdot\underset{\sim}{l}} \langle \beta_i a_1 | H | (\beta_j|\underset{\sim}{l}) a_1 \rangle \qquad (2.77)$$

$$= \sum_{\underset{\sim}{l}} e^{i\underset{\sim}{k}\cdot\underset{\sim}{l}} \langle (\beta_j\beta_i|-\beta_j\underset{\sim}{l}) a_1 | H | a_1 \rangle = \sum_{\hat{\underset{\sim}{s}}} e^{-i\beta_j\underset{\sim}{k}\cdot\hat{\underset{\sim}{s}}} \varepsilon_1(\hat{\underset{\sim}{s}}) \text{ if } \beta_i\beta_j = \beta_t$$

Hence the molecular orbitals and their energies can be obtained from the bond-orbitals $a_j(\underset{\sim}{r})$ by forming Bloch waves from these orbitals and mixing those waves having the same translational symmetry $\underset{\sim}{k}$. This is very similar to the procedure adopted by Stocker[18] and by Cohan[19]; it is strictly valid only if the bond orbitals are mutually orthogonal (a condition which Cohan's orbitals nearly satisfied), when it is justified by the theory contained in this chapter. However, it may well give reasonable results in other cases provided that allowance is made for overlap when the eigenvalues are calculated, since allowing for overlap should have the same effect as deriving mutually orthogonal orbitals from the approximate bond orbitals by means of Lowdin's[30] symmetric orthogonalisation procedure.

CHAPTER 3

OUR METHOD OF FINDING THE EQUIVALENT ORBITALS

3.1 THE PROBLEM OF CONVERGENCE.

It was shown in the last chapter that a set of equivalent orbitals $a_j(\underset{\sim}{r})$ must make stationary the functional X_1 defined by equations 2.41 and 2.42. These can be written in the form

$$X_1 = \langle a_1 | H | a_1 \rangle - \sum_{j=1}^{L} \varepsilon_{j1} \langle a_1 | a_j \rangle \qquad (3.1)$$

where $$\varepsilon_{j1} = \langle a_j | H | a_1 \rangle \qquad (3.2)$$

In practice, it is convenient to eliminate ε_{11} by only considering normalised functions when varying X_1; with this restriction, the functional which has to be made stationary is

$$X' = \langle a_1 | H | a_1 \rangle - \sum_{j=2}^{L} \varepsilon_{j1} S_{j1} \qquad (3.3)$$

where we have written

$$S_{j1} = \langle a_1 | a_j \rangle = \langle a_1 | b_j a_1 \rangle \qquad (3.4)$$

The way in which equations 3.1 and 3.2 were derived makes it clear that in varying X_1 or X' the Lagrangian parameters ε_{j1} must be treated as fixed, and only after a_1 has been found can their values be calculated from equation 3.2. However, such a method may easily fail to converge. In crystals, for instance, the lowest

normalised eigenfunction of H may well resemble a plane wave; for such a function, both s_{j1} and ε_{j1} will in general be quite large. On the other hand, if these large values of ε_{j1} are used in equations 3.1 or 3.3, the function which now makes X^1 stationary will be well localised and will lead to very small values for s_{j1} and ε_{j1}. The use of these new, small, ε_{j1} in equation 3.3, however, will again lead to a function closely resembling a plane wave, and this oscillation will go on indefinitely. In order to overcome this difficulty, it is simplest to consider the variations of the functional

$$Y = \langle a_1|H|a_1\rangle - \sum_{j=2}^{L} \langle a_j|H|a_1\rangle\langle a_1|a_j\rangle \qquad (3.5)$$

where a_1 is normalised so that

$$\langle a_1|a_1\rangle = 1 \qquad (3.6)$$

Since

$$\delta Y = \delta X' - \sum_{j=2}^{L} \langle a_1|a_j\rangle \delta[\langle a_j|H|a_1\rangle] \qquad (3.7)$$

while for exact equivalent orbitals

$$\langle a_1|a_j\rangle = \delta_{j1} \qquad (3.8)$$

it is apparent that for such orbitals $\delta Y = \delta X'$. Thus, if we were to restrict the variations of a_1 in such a way that all the orthonormalisation conditions of equation 3.8 were satisfied, it would make no difference whether we found stationary values of Y or of X^1. Although in practice equation 3.8 cannot be exactly satisfied if only a few basis functions are used, we felt justified in using the variation of Y rather than that of X^1 in our calculations, since the difference between δY and $\delta X'$ is only $\sum_{j=2}^{L} s_{j1}(\frac{1}{2}Ha_1 + \frac{1}{2}Ha_j)$ and this should be negligible for a good approximation to

an equivalent orbital.

Thus, finally, with the notation $s_{j1} = s_i(\underset{\sim}{n})$ and $\epsilon_{j1} = \epsilon_i(\underset{\sim}{n})$ introduced in equation 2.71, the problem of finding equivalent orbitals in tetrahedral semiconductors is reduced to the mathematical problem of finding a normalised function $a_1(\underset{\sim}{r})$ which makes stationary

$$Y = \langle a_1 | H_1 + H_b | a_1 \rangle + 2 \sum_{i \underset{\sim}{n}} \iint \frac{a_i^*(\underset{\sim}{r}) [(\beta_1 | \underset{\sim}{n}) a_i(\underset{\sim}{r}') a_1(\underset{\sim}{r}')] a_1(\underset{\sim}{r})}{|\underset{\sim}{r} - \underset{\sim}{r}'|} d\underset{\sim}{r}' d\underset{\sim}{r}$$

$$- \sum_{i \underset{\sim}{n}} \iint \frac{a_i^*(\underset{\sim}{r}) [(\beta_1 | \underset{\sim}{n}) a_i(\underset{\sim}{r}')] a_1(\underset{\sim}{r}') [(\beta_1 | \underset{\sim}{n}) a_1(\underset{\sim}{r})]}{|\underset{\sim}{r} - \underset{\sim}{r}'|} d\underset{\sim}{r}' d\underset{\sim}{r} - \sum_{i \underset{\sim}{n}} s_1(\underset{\sim}{n}) \epsilon_1(\underset{\sim}{n}) \quad (3.9)$$

$i \neq 1$

The Localisation Potential.

For systems such as tetrahedral semiconductors in their ground states, in which all the valence electrons are contained in the equivalent orbitals $a_j(\underset{\sim}{r})$, the operator $H_b(\underset{\sim}{r})$ arises solely from the core electrons. In such a case, the first term on the right-hand side of equation 3.9 represents the kinetic energy of an electron plus its potential energy in the field of all the atomic nuclei and core electrons. The second term represents the Coulomb energy due to all the valence electrons, including the one being considered, while the third term is the valence-valence exchange energy. The advantage of including $(i, \underset{\sim}{n}) = (o, \underset{\sim}{o})$, i.e. the orbital being considered, in these two terms is that each term separately then arises from a potential having the full crystal symmetry, which greatly simplifies the representation of the crystal potential.

The last term in equation 3.9 is zero for exact equivalent orbitals, in view of equation 3.8, but plays a

very important role in determining the function which minimises Y in a given Hilbert space. It is convenient to define, for a general normalised function a, an operator $\hat{D}$ such that

$$\hat{D}\,a(r) = -\sum_{j=2}^{L} \langle b_j a | H | a \rangle b_j a(r) . \qquad (3.10)$$

we can then write

$$Y = \langle a_1 | H | a_1 \rangle + \langle a_1 | D | a_1 \rangle . \qquad (3.11)$$

The function $\langle a|D|a\rangle$ is not, in general, independent of the choice of origin from which energy is measured; for if we write $H' = H + W$, where W is a constant, then

$$\langle a|D'|a\rangle - \langle a|D|a\rangle = -W \sum_{j=2}^{L} |\langle b_j a | a \rangle|^2 . \qquad (3.12)$$

The right-hand side of equation 3.12 is never less than zero if W is negative. Hence it is usually possible to choose the origin of energy in such a way that $\langle a|D|a\rangle$ is always positive or zero in a given finite-dimensional Hilbert space. Such a choice of origin is desirable since it ensures that at a minimum of Y $\langle a|D|a\rangle \simeq 0$ and so $\langle b_j a | a\rangle \simeq 0$ for all j; if $\langle a|D|a\rangle$ were negative, a minimum of Y could occur with $\langle b_j a|a\rangle$ very large, which would not correspond to an equivalent orbital. On the other hand, it is important that $\langle a|D|a\rangle$ should not be too large compared with $\langle a|H|a\rangle$ initially, if only a finite number of basis functions are used, since otherwise the minimisation of $\langle a|D|a\rangle$ may dominate the calculation; this is undesirable, since it is unlikely that $\sum_{j=2}^{L} |\langle a_j | a_1 \rangle|^2$ can ever be exactly zero with a finite basis set, and the function for which it is an absolute minimum is not necessarily a good approximation to the desired equivalent orbital.

We call the operator ρ a localisation potential, since its expectation value tends to be least for functions which are well localised. Thus for non-overlapping core orbitals its value is zero, while for plane waves $\langle a|\rho|a\rangle$ is very large. Hence, even if the lowest eigenfunction of H resembles a plane wave, as in a crystal, Y will not be minimised for such a function. The localisation potential ρ ensures that the functions which minimise Y are fairly localised, and is the source of the difference between molecular and equivalent orbitals.

3.2 THE CHOICE OF BASIS FUNCTIONS.

Having described the mathematical problem we are trying to solve, we can now turn to the problem of choosing a suitable set of basis functions in terms of which to expand the bond orbital a_1. We note at once that, although there are 4M bonds in the crystal, a complete set of functions for a_1 will not lead to over-completeness when the whole crystal is considered. This is to be contrasted with the calculation of molecular orbitals by the L.C.A.O. method, where a complete set of functions in each unit cell would lead to a very large degree of over-completeness in the total basis set.

If we wish to ensure that the expectation value of the localisation potential is fairly small for the first few of our basis functions, it is advisable to use functions centred about one or two points. Since the use of bipolar co-ordinates, although very suitable for bond orbitals, is not very convenient arithmetically, we restrict our choice

to one of two simple sets of functions, viz. Slater functions of the form $x^l y^m z^n e^{-cr}$ and Gaussian functions of the form $x^l y^m z^n e^{-cr^2}$. In either case, only the first few members of the set would be used in practice for a calculation, and these are quite well localised.

The Disadvantages of Slater Functions.

Since the bond orbitals we are considering are expected to be quite closely connected with chemical bonds, and these latter are frequently described by linear combinations of atomic orbitals, the most natural choice for a set of basis functions would be a few Slater orbitals centred about the end points of our reference bond. Such orbitals could easily be made approximately orthogonal to the core orbitals on the same atom, since these latter can also be adequately represented by a few Slater-type orbitals. If only one orbital per atom is used, we have what is essentially the L.C.B.O. method used by Stocker[18] and by Cohan;[19] however it is most improbable that the overlap integrals s_{j1} can all be made small with such a limited basis set, and so the functions derived in this way cannot be regarded as very good approximations to equivalent orbitals. This is not, however, a fundamental objection, since larger basis sets can easily be chosen, as in Roothaan's[2] self-consistent field method for atoms and molecules.

A much more serious objection to Slater orbitals, and the one which finally made us decide not to use them, is the difficulty of calculating accurately the three and four centre integrals which would arise in computing Y using a

self-consistent Hamiltonian H; one advantage of the L.C.B.O. method is that these integrals can be estimated rather than calculated. Even if one uses the Fourier analysis of the potential which we derive in the next section, it is not easy to calculate integrals of $e^{i\underset{\sim}{k}.\underset{\sim}{r}}$ between Slater functions centred on different atoms, and for a reasonable number of basis functions, say six per atom, a prohibitive amount of work is involved.

Gaussian Functions.

One big advantage of Gaussian functions over Slater functions is that, as shown by Boys[31], all Coulomb and exchange integrals involving Gaussian functions about different centres can easily be calculated explicitly in terms of standard functions, without any need for numerical integration. With the plane wave representation of the potential that we use the position is even simpler, since the integral of $e^{i\underset{\sim}{k}.\underset{\sim}{r}}$ between two Gaussian functions on different centres involves only the sums of products of one-dimensional integrals. Thus, Gaussian functions enable us to avoid most of the computational problems associated with Slater functions.

The usual problem that arises with Gaussian functions is that a large number of them are needed in order to approximate to a single atomic orbital. Orbitals which can be described quite well by two or three Slater functions may require a dozen or more Gaussian functions to provide a satisfactory approximation to them. As a result, most atomic and molecular calculations using Gaussian functions involve a secular determinant of quite high order. However,

all this applies to atomic orbitals rather than to bond orbitals. Unless we are interested in the L.C.A.O. approximation, there seems to be no reason to centre our Gaussian functions about the nuclei at the ends of the bond in a bond-orbital calculation. Considerations of symmetry suggest that a bond is best associated with its mid-point rather than with its ends[32] (at least in tetrahedral semi-conductors), and the usual L.C.A.O. picture for a diatomic molecule also suggests that Gaussian functions centred about the mid-point of a bond might be quite a good basis set. A few such functions might well be enough to represent the smooth part of the bond orbital, while in the regions near the atomic nuclei orthogonalisation of these Gaussian functions to the core orbitals should represent the behaviour of the valence electrons as adequately as it does in the O.P.W. method. Since on the L.C.A.O. picture the bonds involve sp^3 orbitals, this method cannot be expected to work well for crystals such as diamond or boron nitride in which there are no p-type core orbitals.[7] However, in principle there is no reason why localised auxiliary functions should not be introduced, just as they can be in the O.P.W.[8] and modified plane wave[12] methods, to deal with such cases.

Our Choice of Functions.

Since no calculations of the type outlined above seem to have been reported previously, we decided to carry out a series of such calculations. Our aim was to see what could be learned about a whole family of crystals using this method, rather than to carry out an extremely accurate calculation

for just one substance, and we therefore made quite a number of approximations in order to reduce the amount of computer time needed per compound. These approximations are outlined in the remainder of this chapter, and mean that our results are cetainly not definitive: our method was designed to take under two hours per compound on the Mercury computer on which we carried out the calculations, and many of the approximations would be unnecessary on a larger and faster computer such as Atlas.

One major problem in this work, as in most calculations using Gaussian orbitals, is how to choose the parameter C which appears in the factor e^{-cr^2}; if a complete set of functions is used the choice of C is not important, but if only a few functions are used the value of C is much more critical. Since the dimensions of C are $(\text{length})^{-2}$, we assumed that C was inversely proportional to the square of the bond length R for the whole range of compounds we considered. We did not make C depend on any other factors, such as the atomic numbers or electronegativities of the elements of the compound, since it was not apparent a priori how to allow for them. The effects of our choice, which unfortunately turned out to be none too good, are considered in the next chapter, where we present and examine our results.

The limited size of the fast and auxiliary stores in the Mercury computer led to our using only twelve basis functions for the heteronuclear compounds and six for the homonuclear ones. The functions we used are

$$e^{-cr^2},\ xe^{-cr^2},\ \zeta^2 e^{-cr^2},\ r^4 e^{-cr^2},\ x^2\zeta^2 e^{-cr^2},\ \zeta^4 e^{-cr^2}$$

and, for heteronuclear compounds ζ times these, where

$\zeta = x + y + z$ and the axes are as described in section 2 of Chapter 1. All these functions have axial symmetry about the bond, rather than the C_{3v} symmetry required by group theory, as this somewhat simplified the calculations and in any case our basis set was very small. This approximation should not be too bad, since if we write $H = H_{ax} + H_{v}$, where H_{ax} is the part of H having axial symmetry, then in general H_{v} will be much smaller than H_{ax}. The lowest eigenfunction of H_{ax} would normally have axial symmetry, and it should be possible to treat H_{v} as a fairly small perturbation, although in view of the localisation potential D which depends very much on the wave function more extensive calculations would be needed to confirm this.

These functions were all made orthogonal to the core orbitals of the atoms at the ends of the reference bond by means of the Schmidt orthonormalisation process, and the matrix routine described in section 3 of Chapter 1 was used; only core functions having the required C_{3v} symmetry needed to be included in this process. In principle our basis functions should also have been made orthogonal to all the other core orbitals in the crystal, as mentioned in section 4 of Chapter 1, but the other atoms are so much farther away from the mid-point of the reference bond that it seemed a justifiable approximation to save computer time by not doing so. The error this introduced should be small, since in any case there is little likelihood of an orbital built up from our basis functions tending to a core orbital on another

atom. The core functions we used were the self-consistent Hartree-Fock wave functions for a free neutral atom kindly supplied to us by Dr.Mayers. These will not be the exact crystal core functions, i.e. eigenfunctions of the crystal Hamiltonian H, but seemed to be the best available approximations to them. It is known that if an O.P.W. type method is used, rather than an M.P.W. one, the choice of core functions can make a crucial difference to the results of calculations on crystals,[33] but it is reasonable to expect that the effect on the valence bands will be much less than the effect on the energy gap and conduction bands.

3.3 THE CRYSTAL POTENTIAL.

General Considerations.

Ideally, the crystal potential should be calculated self-consistently allowing for all the electrons in the valence band. One advantage of the bond orbital method is that in principle this can be achieved quite simply, but in practice the calculation of an exact self-consistent exchange potential is somewhat laborious and requires a lot of storage space in the computer, and we therefore decided to use an approximate exchange potential.

As previously, we divide H_a into two parts, writing

$$H_a\, a_1(\underset{\sim}{r}) = \left(2\sum_{j=1}^{L}\int \frac{a_j^*(\underset{\sim}{r}')a_j(\underset{\sim}{r}')}{|\underset{\sim}{r}-\underset{\sim}{r}'|}\,d\underset{\sim}{r}'\; a_1(\underset{\sim}{r})\right) - \left(\sum_{j=1}^{L}\int \frac{a_j^*(\underset{\sim}{r}')a_1(\underset{\sim}{r}')a_j(\underset{\sim}{r})}{|\underset{\sim}{r}-\underset{\sim}{r}'|}\,d\underset{\sim}{r}'\right) \quad (3.13)$$

The first term is of the form $V(\underset{\sim}{r})a_1(\underset{\sim}{r})$, and the second term is often replaced b y an approximation of this form so as to avoid the problems associated with a non-local potential. We decided to combine this exchange term with that arising

from the valence-core exchange interaction, and replace them both by Slater's approximation[34] to the exchange potential. Thus we write

$$H\phi_i = -\tfrac{1}{2}\nabla^2\phi_i + V_c\phi_i + V_E\phi_i \qquad (3.14)$$

where V_c is the Coulomb potential due to the nuclei and all the electrons in the crystal and V_E is the exchange potential arising from all the core and valence electrons having the same spin as the one being considered.

Fourier Analysis of the Crystal Potential.

If $V(\underset{\sim}{r})$ is a potential function having the full crystal symmetry, we can write

$$V(\underset{\sim}{r}) = \sum_j V_j e^{i\underset{\sim}{K}_j\cdot\underset{\sim}{r}} \qquad (3.15)$$

where $\underset{\sim}{K}_j$ is a reciprocal lattice vector. On multiplying both sides of this equation by $e^{-i\underset{\sim}{K}_j\cdot\underset{\sim}{r}}$ and integrating over the whole crystal, we find that

$$V_j = \frac{1}{M\Lambda}\int V(\underset{\sim}{r})\, e^{-i\underset{\sim}{K}_j\cdot\underset{\sim}{r}}\, d\underset{\sim}{r} \qquad (3.16)$$

If we now assume that $V(\underset{\sim}{r})$ can be split into components from each unit cell, so that it can be expressed in the form

$$V(\underset{\sim}{r}) = \sum_n v(\underset{\sim}{r}-\underset{\sim}{n}) \qquad (3.17)$$

then, since $e^{i\underset{\sim}{K}_j\cdot\underset{\sim}{n}} = 1$, we find on substituting in equation 3.16 that

$$V_j = \frac{1}{\Lambda}\int v(\underset{\sim}{r})\, e^{-i\underset{\sim}{K}_j\cdot\underset{\sim}{r}}\, d\underset{\sim}{r} \qquad (3.18)$$

which is a most convenient form to use for the Coulomb potential V_c.

It was found to be much more convenient, especially for heteronuclear compounds, to take the origin of co-ordinates

at an A-atom rather than at the mid-point of a bond. Such a choice of origin is usual in L.C.A.O. calculations[18], but not in O.P.W. ones;[22] it is more convenient for the application of symmetry operators, but does mean that, even in homonuclear compounds, the Fourier components of the potential are not all real. With this choice of origin, the valence charge density $\rho^{v}(\underline{r})$ has the form

$$\rho^{v}(\underline{r}) = 2 \sum_{i} \sum_{l=0}^{3} (\beta_{l}|\underline{r}) [a_{i}^{*}(\underline{r})\, a_{i}(\underline{r})] \qquad (3.19)$$

so that, by analogy with equations 3.17 and 3.18

$$\rho_{j}^{v} = \frac{2}{\Omega} \int a^{*}(\underline{r}) \left[\sum_{l=0}^{3} e^{-i\beta_{l}\underline{R}_{j}\cdot\underline{r}} \right] a(\underline{r})\, d\underline{r} \qquad (3.20)$$

Approximations to the Potential.

For the calculation of the Coulomb potential V_{c}, it makes no difference how the electrons are divided up, since the electrostatic potential due to different charge distributions is additive. However, this is not the case for the Slater exchange potential, which is proportional to the cube root of the total charge density $\rho(\underline{r})$ It was not considered practicable to evaluate $\rho^{\frac{1}{3}}(\underline{r})$ exactly, and so the following approximation, based on a suggestion by Kleinman and Phillips[35] was used. Since $\rho(\underline{r})$ is periodic, we can write

$$\rho(\underline{r}) = \sum_{j} \rho_{j} e^{i\underline{K}_{j}\cdot\underline{r}} \qquad (3.21)$$

by analogy with equation 3.15. We now write

$$\rho(\underline{r}) = \rho_{0}(\underline{r}) + \rho_{1}(\underline{r}) \qquad (3.22)$$

where $\rho_{0}(\underline{r})$ is the charge density that would arise in the crystal if each atom were in its ground state and unaffected

by its neighbours. Then, provided that $\rho_1 << \rho_0$ everywhere, we can write

$$\rho^{\frac{1}{3}}(\underline{r}) \simeq \rho_0^{\frac{1}{3}}(\underline{r}) \left[1 + \frac{\rho_1(\underline{r})}{3\rho_0(\underline{r})}\right] \qquad (3.23)$$

and so, as a further approximation

$$\rho^{\frac{1}{3}}(\underline{r}) \simeq \rho_0^{\frac{1}{3}}(\underline{r}) + \tfrac{1}{3}\rho_0^{-\frac{2}{3}}(\underline{r}) \sum_j \rho_j' e^{-i\underline{K}_j \cdot \underline{r}} \qquad (3.24)$$

where

$$\rho_j' = \frac{1}{\Omega} \int \rho_1(\underline{r})\, e^{-i\underline{K}_j \cdot \underline{r}}\, d\underline{r} \qquad (3.25)$$

In our calculations, we replaced the factor $\rho_0^{-\frac{2}{3}}(\underline{r})$ by the value appropriate for a uniform distribution of the valence electrons, viz. $\left(\frac{8}{\Omega}\right)^{-\frac{2}{3}}$. The Fourier components of ρ_0 were calculated using values of ρ_0 derived from Dr.Mayers' wave functions, assuming a spherically symmetric charge distribution about each atom, and with a cut-off in the integrations when $\rho(\underline{r})$ became less than $\frac{1}{2\Omega}$ so as to prevent too large a contribution arising from the outer parts of the atom. This cut-off led to a sensible value for the radius of the atom, and gave a value of V_{111} for silicon in fair agreement with that calculated by other authors.[36] $\rho_1(\underline{r})$ was easily calculated, and its Fourier components obtained without difficulty, so that a nominally self-consistent potential could be found without too much trouble.

The Fourier series for $V(\underline{r})$ was cut off at $|\underline{K}_j|^2 = 20(\pi\sqrt{3}/2R)^2$, chiefly because the storage capacity of the Mercury computer would not allow us to go much farther for heteronuclear compounds. This cut-off is quite reasonable, since the matrix element of $e^{i\underline{k}\cdot\underline{r}}$ between two Gaussian functions decreases very rapidly as $|\underline{K}|$ increases. In

germanium, for instance, $\int e^{-2cr^2} \sum_{l=0}^{3} e^{-\beta_l \underline{K}\cdot\underline{r}}\, d\underline{r}$ has for $K^2 = 19\,(\pi\sqrt{3}/2R)^2$ about one fourtieth of its value for $K^2 = 3\,(\pi\sqrt{3}/2R)^2$; in addition, $|V_j|$ decreases as $|\underline{K}_j|$ increases.

The potential described above was only used for the smooth part of the wave function, i.e. for that part which is spanned by the Gaussian orbitals. For the core functions, where our cut-off of the Fourier series of the potential would involve a much greater error, we preferred to use the energy of the orbital in a free atom, so that if u_i is a core function we wrote

$$Hu_i = E_i u_i. \qquad (3.26)$$

This does not mean that two different Hamiltonians were used for the smooth and oscillating parts of the bond orbital, but rather that two different approximations were used in calculating the matrix elements of the true crystal Hamiltonian H. In order to allow for the difference of local environment of the two atoms in a heteronuclear crystal,[37] the positions of their zeroes of energy were so adjusted that the average energy of the valence electrons in the neutral free atoms should be the same; this had only a slight effect on the results.

Finally, the component V_o of the crystal potential, which also does not affect the results much, was chosen so that the average energy of a valence electron in a crystal, i.e. $\epsilon_i(\underline{0})$ should equal its average energy in the free atoms plus the approximate binding energy per electron of the

crystal. We note in passing that the method used by Kleinman and Phillips[36] to find this average valence energy, which involved averaging the first n ionisation potentials of an atom with n valence electrons, is not valid. The quantity we require is the average energy needed to remove one valence electron while all the other electrons are still on the atom, and according to Koopnan's theorem this is just the average of the Hartree-Fock energies of the valence orbitals. This leads for silicon to V_o = 0.95 a.u., which is in much better agreement with Woodruff's value[38] of V_o = -1.0 a.u. calculated for free atoms than is the value V_o = -1.5 a.u. given by Kleinman and Phillips.[36]

3.4 OTHER APPROXIMATIONS MADE IN THE CALCULATIONS.

The Number of Neighbouring Bonds Considered.

Since it is obviously not practicable to deal with all the bonds in the crystal, it is necessary to decide how far out from the central bond to go. We decided to classify the bonds by the distances of their mid-points from the mid-point of our reference bond a_1, and to ignore all bonds for which this distance was greater than twice the side of our unit cube. ~~cell.~~ As a result, we considered a total of 72 neighbouring bonds, including all first and second neighbours, the third and fourth neighbours parallel to a_1, and most of the other third neighbours. This leads to eleven independent parameters $\xi_i(n)$ for a heteronuclear compound, which we label as follows; a similar notation is adopted for the $S_i(n)$.

$E_0 = \xi_0(000)$ the self-energy of a bond.

$E_1 = \xi_1(000)$ for first neighbours having an A atom in common with a_1.

$E_2 = \xi_1(011)$ for first neighbours having a B atom in common with a_1.

$E_3 = \xi_1(101)$ for second neighbours not parallel to a_1.

$E_4 = \xi_0(011)$ for second neighbours parallel to a_1.

$E_5 = \xi_0(0\bar{1}1)$ for third neighbours parallel to a_1.

$E_6 = \xi_1(01\bar{1})$ for third neighbours not parallel to a_1.

$E_7 = \xi_1(020)$ " " " " " " "

$E_8 = \xi_1(\bar{1}0\bar{1})$ " " " " " " "

$E_9 = \xi_1(121)$ " " " " " " "

$E_{10} = \xi_0(200)$ for fourth neighbours parallel to a_1.

Of these, E_0 plays no part in determining the shape of the energy bands, while in a homonuclear crystal $E_2 = E_1$, $E_7 = E_6$, and $E_9 = E_8$. Our cut-off seems fairly reasonable, since E_9 and E_{10} tend to be considerably smaller in magnitude than E_1 and E_2.

An Approximation to $\langle(\beta_i - l_\alpha)a_1|H|a_1\rangle$

Although the matrix elements of H between different bond orbitals are not difficult to calculate, and our programmes would only need a slight modification in order to give them, it is not possible to store on the Mercury computer

all the matrices needed in order to form these elements reasonably quickly for a given function a_1. On the other hand, since we were using an iterative procedure to find H and a_1 self-consistently, it would have taken a prohibitively long time to re-calculate these matrices whenever they were needed. We therefore decided to make the rather drastic approximation that the matrix $TH^{in}\tilde{T}$ of the Hamiltonian H between our orthonormalised basis functions on different atoms could be replaced by $TS^{in}\tilde{T}TH^{oo}\tilde{T}$, where $TS^{in}\tilde{T}$ is the overlap matrix between these orthonormal functions. This approximation would involve no error if Ha_1 could be expressed exactly in terms of our original set of functions, and was expected not to involve too much error for our type of bond orbital. However, it can easily introduce spurious minima of Y i.e. ones which would not appear if this approximation were not used; this will happen if Ha_1 has a large component lying outside our original basis set, and we modified our convergence procedure slightly so as to reduce the risk of this happening.

The Convergence Procedure.

The convergence procedure we used is essentially that devised by Powell,[39] for which a library programme in Mercury Autocode was available; several other possible procedures were tested before we finally decided to use this one. This procedure is a very rapid and stable method for finding the minimum of a differentiable function of many variables nearest to the original starting point, and has second order convergence. We decided to start with the

approximation that $a_1 = Ae^{-cr^2}$, since this is fairly similar in shape to the bonding orbital in H_2. Since a_1 is normalised, if it is expressed in terms of n basis functions only n-1 of the coefficients will be indepentent, so that the minima of Y^4 will lie on lines, rather than at discreet points, in n-dimensional co-ordinate space; on the other hand, trial calculations suggested that with our convergence procedure it was better to work in n-dimensional space rather than in an (n-1)-dimensional space. However, the existence of a line of minima means that rounding errors can be very important, and additional difficulties arose because exact differentials of Y^4 could not be found very conveniently. For these reasons, as well as for the one mentioned in the last paragraph, we modified the convergence procedure slightly so that if necessary it would stop on finding a region in which Y^4 was effectively constant.

As mentioned in section one of this chapter, it is necessary to lower the energy by a constant in order to ensure that the expectation value of D is not negative for all the orbitals we consider. It was found that subtracting 4 a.u. of energy from H was sufficient to ensure that at a minimum of Y^4 all the overlap integrals S_{j1} were small, while this energy shift did not upset the balance between $\langle a_1|D|a_1\rangle$ and $\langle a_1|H|a_1\rangle$ in Y^4.

An Approximation to $TS_K T$.

Let $S_{\underset{\sim}{K}}$ be the matrix of $\sum_i e^{-i\beta_i \underset{\sim}{K}.\underset{\sim}{r}}$ between our original basis functions, so that $TS_{\underset{\sim}{K}}\tilde{T}$ is the corresponding

matrix for the orthonormalised functions. If we write

$$S_{\underset{\sim}{K}} = \left(\begin{array}{c|c} P_{\underset{\sim}{K}} & Q_{\underset{\sim}{K}} \\ \hline \tilde{Q}_{\underset{\sim}{K}} & R_{\underset{\sim}{K}} \end{array} \right), \qquad T = \left(\begin{array}{c|c} 1 & 2 \\ \hline A & B \end{array} \right), \tag{3.26}$$

in accordance with equations 1.92 and 1.93, then by analogy with equation 1.96 the lower right-hand part of $TS_{\underset{\sim}{K}}\tilde{T}$ has the form

$$Z_{\underset{\sim}{K}} = A P_{\underset{\sim}{K}} \tilde{A} + A Q_{\underset{\sim}{K}} \tilde{A} + A R_{\underset{\sim}{K}} \tilde{B} + B R_{\underset{\sim}{K}} \tilde{B} \tag{3.27}$$

The matrix $TS_{\underset{\sim}{K}}\tilde{T}$ is only used in calculating the valence charge density, from which is derived the contribution of the valence electrons to the crystal potential. For the important values of $\underset{\sim}{K}$ this contribution is about 10% of that due to the core electrons and nuclei, so that an error of even 20% in it would produce an error of only about 2% in the crystal potential. On the other hand, the calculation of $Q_{\underset{\sim}{K}}$ involves numerical integration in two dimensions, and took about as long as all the rest of the calculation. We therefore decided to use the approximation

$$Z_{\underset{\sim}{K}} = B R_{\underset{\sim}{K}} \tilde{B} - A R_{\underset{\sim}{K}} A P_{\underset{\sim}{K}} \tilde{A} \tag{3.28}$$

which is exact for $\underset{\sim}{K} = \underset{\sim}{o}$ as shown in equation 1.97. The error involved in this approximation increases with $|\underset{\sim}{K}|$, and was about 10% for $|\underset{\sim}{K}|^2 = 3(\pi\sqrt{3}/2R)^2$. Since the matrix elements of $BR_{\underset{\sim}{K}}\tilde{B}$ decrease in magnitude as $|\underset{\sim}{K}|$ increases, this approximation should not introduce a greater error into H than was introduced by our approximation to the valence-valence exchange interaction.

The Computer Programmes.

A summary of the computer programmes which we used, ~~written in Mercury Autocode,~~ together with brief explanations of the mathematical techniques they embody, is contained in Appendix III.

CHAPTER 4

THE RESULTS OF OUR CALCULATIONS

4.1 THE ONE-ELECTRON ENERGY BANDS.

The Factorisation of $S_{\underline{k}}$

As was shown in section 3 of chapter 2, once the equivalent orbitals are known and the parameters have been found, one can construct a 4 x 4 matrix $S_{\underline{k}}$, the eigenvalues of which for different values of $\underline{k}$ give the one-electron energies of the four valence bands of a tetrahedral semi-conductor. Since a quartic equation can be solved algebraically, it is in principle possible to write down explicit formulae for these energies $E^{\underline{k}}$ in terms of the elements of $S_{\underline{k}}$. This may be very useful if really accurate calculations have been performed and one wants to derive a density of states curve or effective masses, but is not very helpful if we just want to obtain a general picture of the results of our calculations. For such purposes, it is much simpler and more enlightening to plot the variations of $E^{\underline{k}}$ with $\underline{k}$ along a few lines in reciprocal space. The determinant of $S_{\underline{k}}$ factorises if $\underline{k}$ has components of the form (l, m, m), and it is usual to consider in particular the directions (l, o, o) and (l, l, l), since in these directions the symmetry of $\underline{k}$ contributes most to simplifying calculations carried out by the standard molecular orbital methods.[40]

If we write $\underline{k} = {}^{1}\!/\!_{L}\ (21,\ o,\ o)$ and ${}^{1}\!/\!_{L}(1,\ 1,\ 1)$ respectively for these two directions, where L is the length of the side of our basic cube inscribed in the unit cell, then the formulae for $E^{\underline{k}}$ are as given below. We use the notation of section 4 of chapter 3, and write

$F_0=E_0$, $F_1=E_1+E_2$, $F_2=E_1-E_2$, $F_3=2E_3$, $F_4=2E_4$, $F_5=2E_5$, $F_6=E_6+E_7$,

$F_7=E_6-E_7$, $F_8=E_8+E_9$, $F_9=E_8-E_9$, $F_{10}=2E_{10}$. (4.1)

We note that for a homonuclear compound, $F_2 = F_7 = F_9 = 0$

k = 1/L (21, 0, 0)

Let $p = F_0 + (F_4+F_5)(1+2\cos 21) + F_{10}(2 + \cos 41)$,

$q = F_1 + 2F_3 \cos 21 + 2F_6 + 4F_8 \cos 21$,

$r = (F_1+ 2F_3+ 2F_6 \cos 21 + 4F_8 \cos 21) \cos 1$,

and $s = F_2 \sin 1 + 2F_7 \cos 21 \sin 1 + 4F_9 \sin 21 \cos 1$.

Then $E_1 = E_2 = p - q$,

$E_3 = p + q + 2(r^2 + s^2)^{\frac{1}{2}}$

and $E_4 = p + q - 2(r^2 + s^2)^{\frac{1}{2}}$ (4.2)

For a homonuclear compound, s = o.

k = 1/L (1, 1, 1)

Let $a = F_0 + 3F_4 \cos 21 + 3F_5 + 3F_{10} \cos t\ 21$,

$b = F_0 + F_4\ (2 + \cos 21) + F_5(1+2\cos 21) + 3F_{10} \cos 21$,

$c = (F_1+ 2F_3 + 2F_6 + 4F_8 \cos 21) \cos 1$,

$d = F_2 \sin 1 + 2F_7 \sin 1 + 4F_9 \cos 1 \sin 21$,

and $e = F_1+ 2F_3 \cos^2 1 + 2F_6 \cos 21 + 4F_8 \cos^2 1$.

Then $E_1 = E_2 = b - e,$

$$E_3 = \tfrac{1}{2}(a+b) + e + \left[(e+\tfrac{1}{2}a-\tfrac{1}{2}b)^2 + 3c^2 + 3d^2 \right]^{\frac{1}{2}}$$

and $$E_4 = \tfrac{1}{2}(a+b) + e - \left[(e+\tfrac{1}{2}a-\tfrac{1}{2}b)^2 + 3c^2 + 3d^2 \right]^{\frac{1}{2}} \qquad (4.3)$$

For a homonuclear compound, d = 0.

The Effect of Non-Zero Overlap Integrals.

The overlap integrals S_j are typically of order 0.01 or less. Thus, the energies are scarcely changed if one allows for the fact that the Bloch waves with a given $\underline{k}$ formed from the equivalent orbitals are not exactly orthogonal by using Lowdin's symmetric orthogonalisation procedure[30] and introducing the matrix $(1 + S)^{-\frac{1}{2}}$. On the other hand, since the parameters E_j for j greater than 2 are typically also of order 0.01 or less, it does make a considerable difference whether one uses E_j or $E_j - E_o S_j$ in calculating the one-electron energies. The parameters used in the derivation of equation 2.58 for the matrix S_1 are essentially defined by

$$H a_1 = \sum_{j=1}^{L} \varepsilon_{j1} a_j \qquad (4.4)$$

and not by

$$\varepsilon_{j1} = \langle a_j | H | a_1 \rangle \; ; \qquad (4.5)$$

equation 4.5 can only be derived from equation 4.4 if the functions a_j are strictly orthogonal. In our case, although S_{j1} is small ε_{j1} is of the same order of magnitude, and so in forming $S_{\underline{k}}$ we must use the parameters ε'_{11} defined by

$$\varepsilon'_{11} = \langle a_1 | H | a_1 \rangle - \sum_{j \neq 1} \varepsilon'_{j1} \langle a_1 | a_j \rangle \qquad (4.6)$$

The largest term in the sum on the right arises from j = 1 since ε_{11} in our work was an order of magnitude larger than any of the other ε_{j1} . Hence as a first approximation we can write

$$\varepsilon'_{11} \simeq \langle a_1|H|a_1\rangle - \varepsilon'_{11}\langle a_1|a_1\rangle \qquad (4.7)$$

i.e. $E'_1 = E_1 - E_0 S_1$. (4.8)

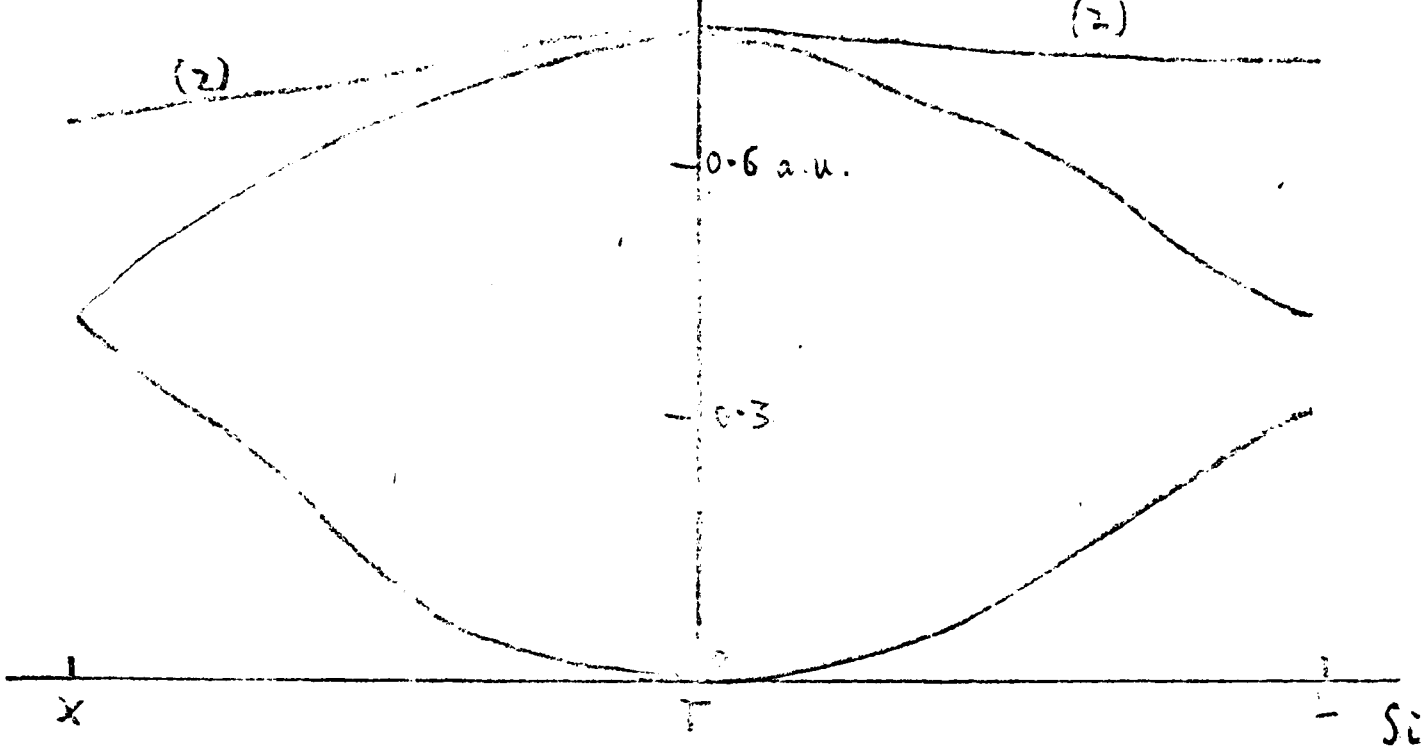
This corresponds to the form used by Mulliken et al.[41] in molecular calculations and by Stocker[18] in the calculations in which he included overlap, and ensures that the shapes of the energy bands are not affected by changes in the zero of energy. In fact, we used a rapidly convergent iterative procedure, starting with the approximation 4.8, to find self-consistent values of the E'_j, using all 72 bonds which we considered in the sum on the right of equation 4.6. We found that the resulting E'_j differed only slightly from those given by equation 4.8.

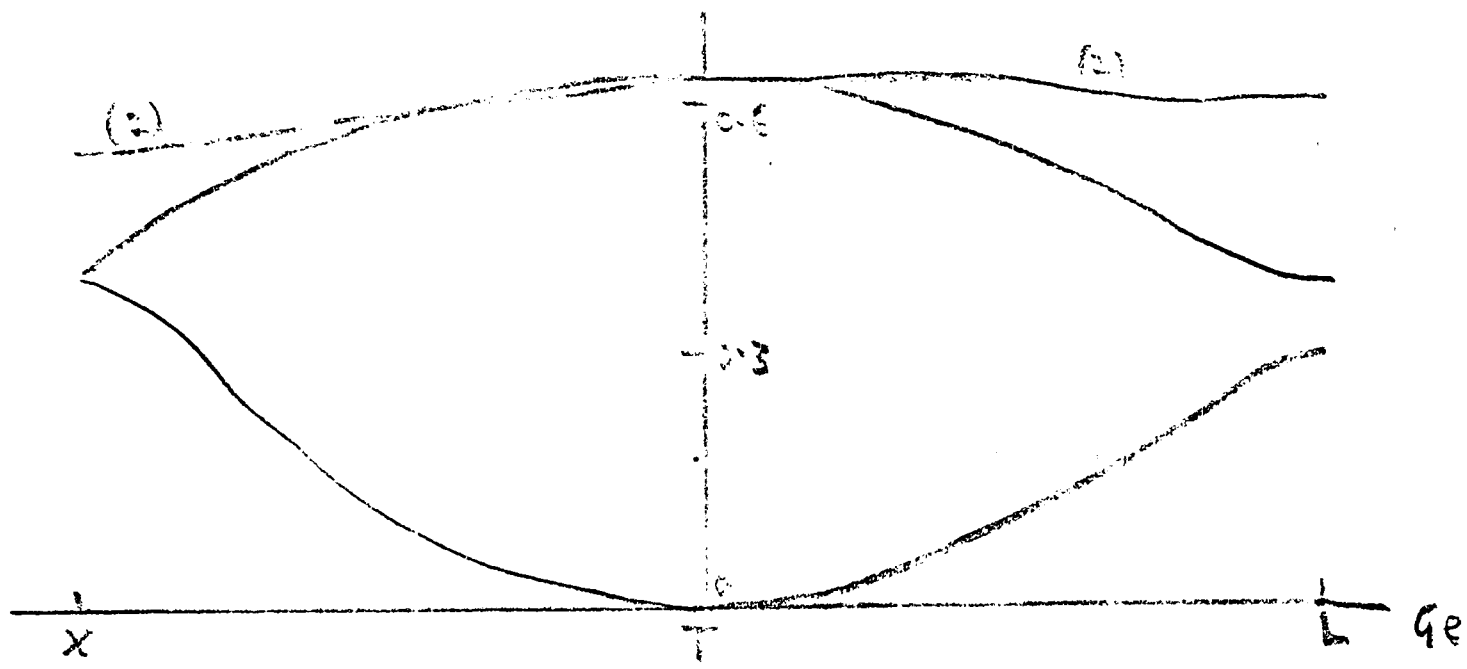
Energy Bands Derived from Equivalent Orbitals.

Our calculations showed that the parameters E'_j, because of their small size, were very sensitive to the crystal potential; a change of five or ten per cent in this potential could even sometimes lead to a change in the sign of some of the E'_j. This may not be quite as disastrous as it seems, since the orbital which makes Y stationary for a given choice of potential will not in general do so for a different type of potential, and the change in a_1 would probably tend to stabilise thevalues of the E'_j. Nevertheless, since our exchange potential, and especially that part of it due to the valence electrons, was only a rough approximation,

we cannot expect our values for the parameters E'_j to be particularly accurate. On the other hand, the detailed shape of the energy bands is quite sensitive to these parameters, as can be seen from equations 4.2 and 4.3, except perhaps at $\underline{k} = 0$. Hence, the results we obtained for the energy levels cannot be regarded as very reliable.

The energy bands we obtained for the diamond-type crystals, shown in figure 2, all have fairly reasonable shapes; in fact we found that almost any reasonable set of E'_j, and even such an improbable set as those obtained for tin with $CR^2 = 3.5$ (see table 1) produced bands of the same general shape as those derived by other authors.[42] For the heteronuclear zincblende type compounds, however, the position is very different. If these compounds could be treated by applying perturbation theory to the results for the corresponding homonuclear crystals, as suggested by Kane[43] and by Callaway[44] among others, one would expect F_2, F_7, and F_9 as defined by equation 4.1 to be much smaller in magnitude than F_1, F_6, and F_8 respectively. As a result, S and d would be fairly small in equations 4.2 and 4.3, so that at most points in reciprocal space the energy bands would be quite similar to those for the homonuclear crystals. Unfortunately, our calculations seem to have over-emphasised the anti-symmetric part of the equivalent orbital, and the resulting band shapes were sometimes very different from those predicted by other authors and found experimentally; for instance, in several cases we found that at $\underline{k} = 0$ there was a local minimum instead of a maximum in the upper valence band, which is most improbable. Because we do not feel that a vast amount of significance can be attached to these





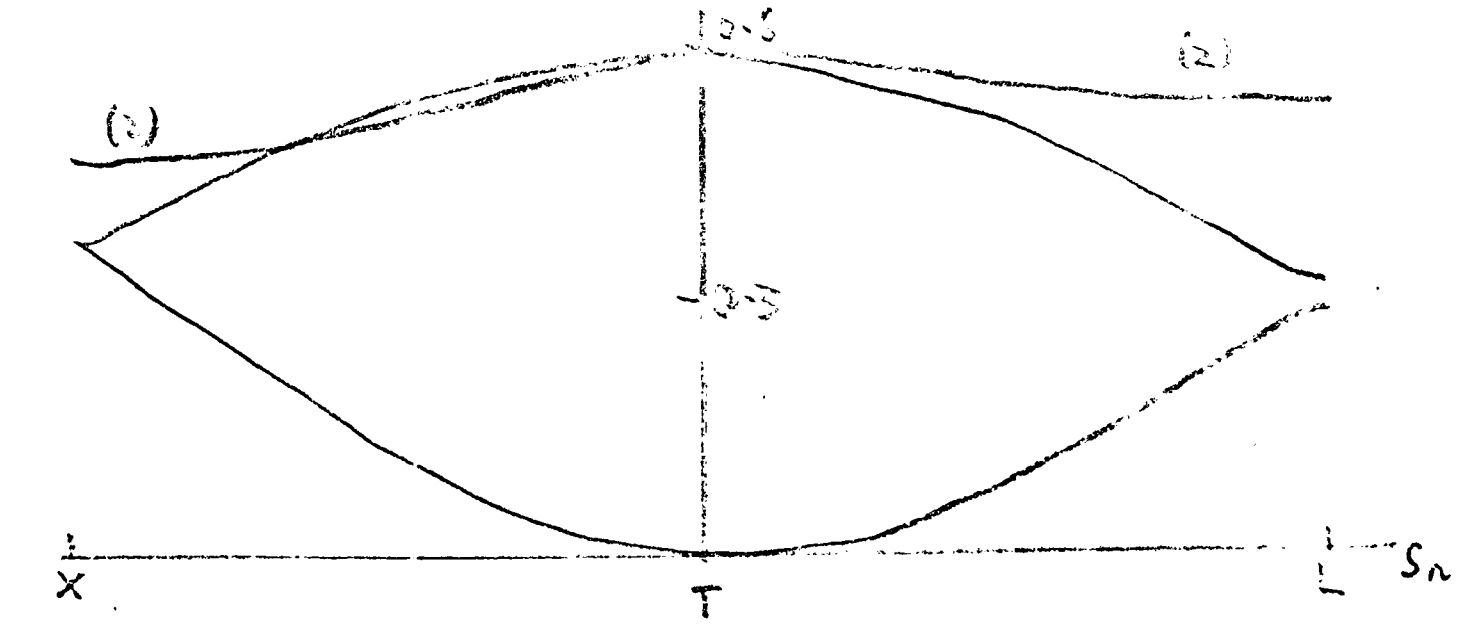


Figure 2. Valence bands of Si, Ge, and Sn. $X = \frac{\pi}{L}(1,0,0)$, $L = \frac{\pi}{L}(\frac{1}{2},\frac{1}{2},\frac{1}{2})$.

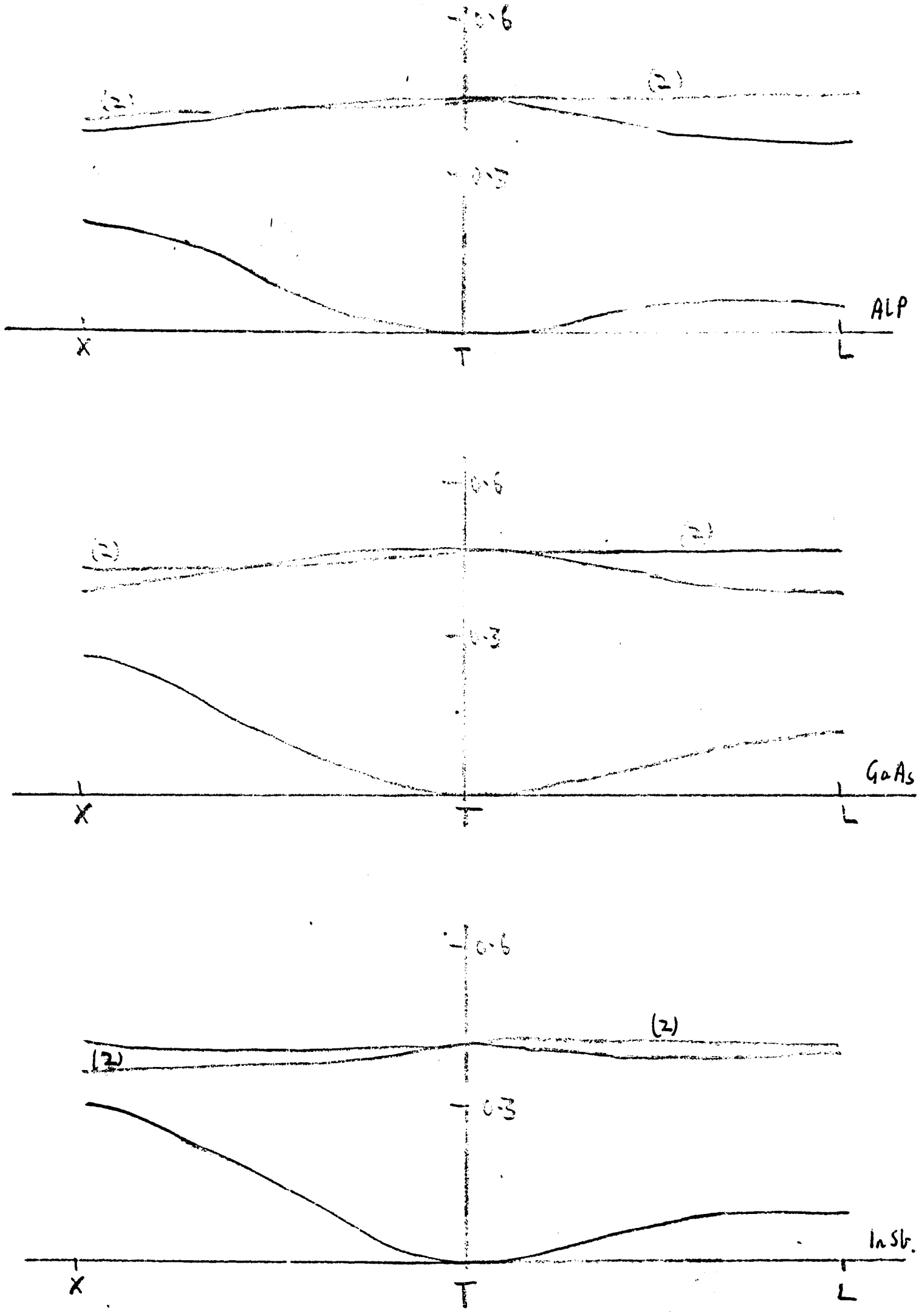


Figure 3. Valence bands of ALP, GaAs, and InSb.

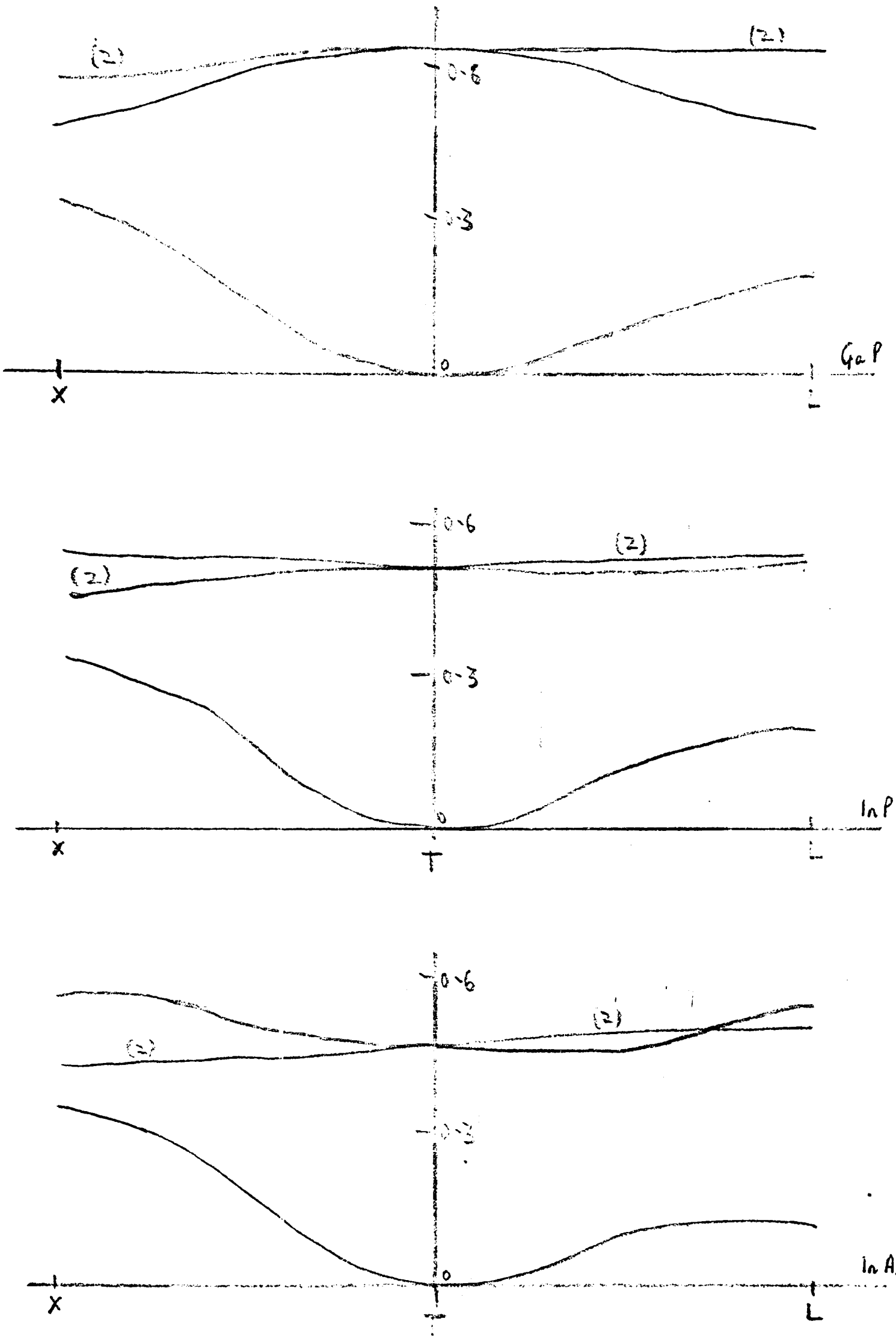


Figure 4. Valence bands of GaP, InP, and InAs.

results, we only present in figures 3 and 4 the energy bands for a couple of sets of III - V compounds for which our results seem not too improbable. They are discussed further in section 3, where we also present the parameters E_j^{ℓ} for those crystals whose energy bands we have not drawn. We note in passing that a calculation on CdTe in which the anti-symmetric part of the wave function was omitted led to a very reasonable set of energy bands.

If the energy bands were to be found by an accurate calculation, the effective masses could in principle be found from them quite easily. However, differentiation of equations 4.2 and 4.3 with respect to 1 shows that the parameters F_8, F_9, and F_{10}, and perhaps also parameters for bonds which we have not even considered, would have to be known very accurately (say to within 0.01 e.v.) for reliable values to be obtained. This problem is considered in more detail in Appendix II. Finally, we must point out that in all our work spin-orbit interactions were ignored, although the equivalent orbital formalism might well provide a very convenient way for treating them.

4.2 EQUIVALENT ORBITALS IN DIAMOND-TYPE CRYSTALS.

Since for the homonuclear crystals the valence-band orbitals are symmetric about the mid-point of the bond, we only used six expansion functions in calculations on them. As a result, these calculations took much less time on the computer than did those for the heteronuclear compounds, and so we were able to carry them out for a number of different values of C, to test various convergence procedures on them, etc. Another reason for experimenting with these substances

was that a fair number of calculations on silicon and germanium have been reported, while until recently[42] very little had been published on the heteronuclear compounds which interested us.

The results of some of our calculations are contained in table 1, where we list the wave functions, energy parameters E_j^1, ~~overlap integrals S_j, and energy levels at Γ~~, etc., ~~X, and L~~ obtained for silicon with $CR^2 = 3$ and 3.3, for germanium with $CR^2 = 3$, 3.5, and 4, and for grey tin with $CR^2 = 3$ and 3.5; as previously, R is the bond-length in atomic units. It ~~will be seen~~ was found that C had some effect on the energies but tended to leave the order of levels ~~at these points~~ unchanged. The value of C to be used in our subsequent calculations was eventually decided on the basis of our results for tin, where the parameters E_j^1 obtained with $CR^2 = 3$ looked far more reasonable than those obtained with $CR^2 = 3.5$, even though the shape of the bands is not too sensitive to the change; the results of numberous calculations on silicon made us unwilling to use a lower value of CR^2. However, the fact that $CR^2 = 4$ gives sensible results for silicon while $CR^2 = 3$ is needed for tin suggests that the bond orbital in tin is in fact less localised than in silicon. This would be in accord with the suggestion of Mooser and Pearson[45] that the directionality of a bond decreases as the principal quantum number n of the valence electrons in it increases, since n = 3 for silicon while for tin n = 5. In spite of this, we decided to keep CR^2 constant in all our calculations, since otherwise we would have had to assume a somewhat arbitrary variation in it, especially when considering the heteronuclear

	Si	Si	Ge	Ge	Ge	Sn	Sn
CR^2	3	3.3	3	3.5	4	3	3.5
$E1'$	-0.0360	-0.0333	-0.0268	-0.0235	-0.0131	-0.0228	-0.0115
$E3'$	-0.0111	-0.0125	-0.0134	-0.0137	-0.0125	-0.0135	-0.0151
$E4'$	0.0101	0.0076	0.0074	0.0068	0.0066	0.0091	0.0075
$E5'$	-0.0082	-0.0077	-0.0079	-0.0076	-0.0063	-0.0073	-0.0073
$E6'$	-0.0019	-0.0019	-0.0018	-0.0015	-0.0013	-0.0021	-0.0014
$E8'$	0.0012	0.0012	0.0014	0.0015	0.0014	0.0011	0.0016
$E10'$	0	0	0.0001	0.0001	0.0001	0.0001	0.0002
$S1$	-0.0455	-0.0412	-0.0341	-0.0294	-0.0151	-0.0256	-0.0130
$S3$	-0.0118	-0.0166	-0.0237	-0.0346	-0.0557	-0.0304	-0.0166
$S4$	0.1225	0.1270	0.1201	0.1186	0.1122	0.0942	0.0678
$S5$	-0.0048	-0.0043	0.0104	0.0083	-0.0033	0.0152	+0.0052
$S6$	0.0041	0.0047	0.0001	-0.0002	-0.0024	-0.0044	-0.0031
$S8$	0.0040	0.0034	0.0028	0.0026	0.0034	0.0025	0.0018
$S10$	0.0003	0.0003	-0.0003	-0.0002	0.0001	-0.0007	-0.0008
χ	0.2407	0.2348	0.2315	0.2287	0.2244	0.2033	0.2056
	-0.0124	-0.0087	-0.0127	-0.0117	-0.0119	-0.0107	-0.0117
	-0.0033	-0.0006	0.0016	0.0039	0.0119	-0.0013	0.0029
	0.0001	-0.0001	0	-0.0001	-0.0002	0.0001	0
	0.0002	0.0002	0.0006	0.0007	0.0012	0.0003	0.0006
	-0.0002	-0.0003	-0.0004	-0.0006	-0.0012	-0.0001	-0.0004
χ_A	-0.0150	-0.0163	-0.0043	-0.0045	-0.0046	-0.0016	-0.0017
	0.1277	0.1351	0.0302	0.0311	0.0317	0.0104	0.0115
	-0.1225	-0.1267	-0.0100	-0.0106	-0.0127	-0.0019	-0.0023
			-0.1362	-0.1390	-0.1412	-0.0398	-0.0438
			-0.0479	-0.0439	-0.0364	-0.0212	-0.0165
			0.1472	0.1536	0.1753	0.0220	0.0260
						0.1460	0.1568
						0.0344	0.0681
						-0.1885	-0.2133

TABLE 1

For key, see table 2.

$$S_j \simeq 0.23(E_j' - E_j)$$

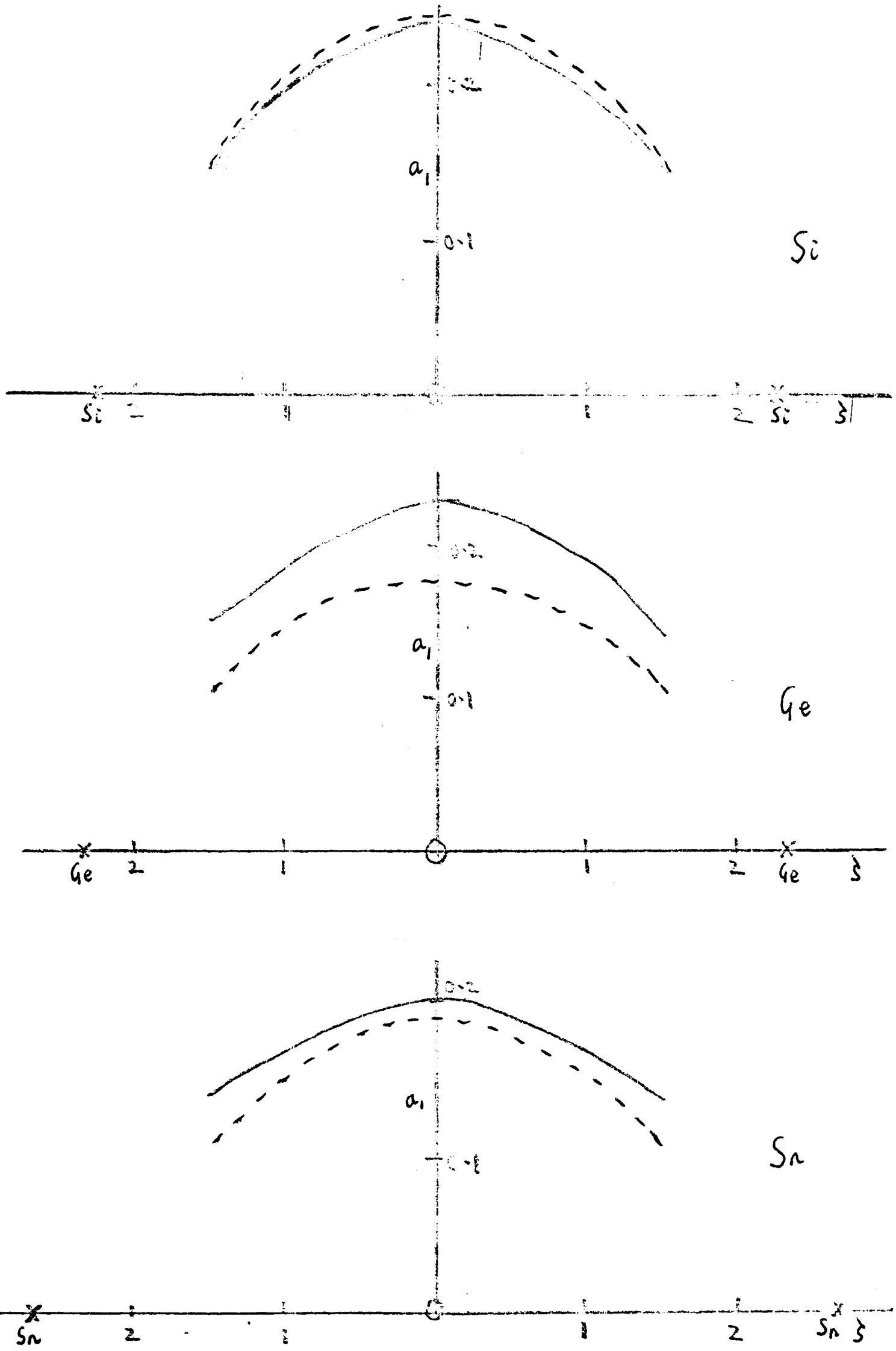


Figure 5. Equivalent orbitals (continuous line) and L.C.A.O. bond orbitals (broken line) in homonuclear semiconductors.

compounds with elements from different rows of the periodic table.

In figure 5 we have drawn the smooth parts of the equivalent orbitals for Si, Ge, and Sn with $CR^2 = 3$, and for comparison the orbitals obtained by a simple L.C.B.O. treatment (ignoring overlap). It is to be noted that our orbitals have a shallow maximum at the mid-point of the bond, although classically the electrostatic potential would be less attractive there than nearer the two atoms. In all our diagrams of equivalent orbitals we go out to a distance of 1.5 atomic units from the mid-point of the bond, since for almost all the crystals we consider the effect of the orthogonalisation of the bond orbitals to the atomic core functions is small in this region but tends to become appreciable as we get closer to the ends of the bond. In section 4 of this chapter we consider farther the shapes of the equivalent orbitals and the experimental evidence in connection with them.

4.3 EQUIVALENT ORBITALS IN ZINCBLENDE-TYPE CRYSTALS.

With one or two exceptions, from which we were unable to make any useful deductions, all our calculations on the heteronuclear compounds were performed with $CR^2 = 3$. As remarked previously, the resulting energy bands, expecially for the II - VI compounds, seem highly improbable, but this does not necessarily mean that our bond orbitals are not good approximations to the correct ones. However, the results are sensitive to the choice of C, and if Mooser and Pearson[45] are correct in suggesting that the directionality of the bonds decreases as the difference in the electronegatives of the atoms increases (as well as decreasing as the principal quantum number of the valence electrons

increases), smaller values of C would probably be better for most of the compounds we considered. In addition, it seems probable in view of the way in which Y sometimes approached its minimum that the valence-valence exchange potential we used was such as to over-emphasise the anti-symmetric part of the orbital. More extensive calculations would be needed to test these possibilities.

Another, rather more fundamental, question is whether for heteronuclear compounds the orbitals should be expanded in terms of functions centred about the mid-point of the bond. In view of the differences in the sizes of the constituent atoms and in the number of electrons they contribute to the bond, it might well be prefereable to centre the expansion functions about some other point of the bond. Support for this idea comes from the fact that our most reasonable results were obtained for the series of compounds in which both these differences were least, namely the III - V compounds AlP, GaAs, and InSb in which both atoms come from the same row of the periodic table; also, of the II - VI compounds CdTe gave the least unexpected results while our calculation for ZnSe is not reliable. Such a change in the expansion functions would require very little alteration in the theory we have developed or in our computer programmes, but it might not be easy to find the best point about which to centre the functions; alternatively, a much larger set of basis functions centred about the mid-point could be used, though this would probably not be so convenient.

In figures 6, 7, and 8 we present respectively the equivalent orbitals for AlP, GaAs, and InSb in which both elements are about the same size, for AlAs, AlSb, and GaSb

in which the group V atom is much heavier, and for InP, InAs, and GaP in which the group III atom is the heavier. Figures 9 and 10 contain the bond orbitals for the zinc and cadmium compounds respectively. Tables 2 to 6 contain the relevant parameters for all these compounds. For comparison, we have also drawn in figure 6 the smooth parts of the L.C.A.O. bond orbitals obtained by Coulson, Redei, and Stocker[20] for AlP, GaAs, and InSb.

In spite of the uncertainty as to the reliability of our results, it is perhaps worthwhile to consider briefly their significance; we consider general trends rather than the results for any particular substances. Our results for the III - V compounds show that in these the charge density is least just outside the core of the group III atom and increases steadily as one approaches the group V atom; whether this is equivalent to covalent or ionic bonding depends largely on how these terms are defined.[20] However our results for the II - VI compounds, if correct, suggest that here the bonding is definitely ionic, with charge tending to accumulate near the two atoms and with a node at some point of the bond nearer to the group II than to the group VI atom. In such a case the binding would presumably arise from electrostatic forces due to a transfer of charge from the group II to the group VI atom. This is almost certainly the position in I - VII compounds such as CuCl and CuBr; it is not entirely inconsistent with the open tetrahedral structure, since for some ratios of the radii of the atoms (or rather the ions) this is classically the most favourable structure energetically for an ionic compound.

AlP		GaAs		InSb.	
E'_j	S_j	E'_j	S_j	E'_j	S_j
0.0149	0.0061	0.0081	0.0054	0.0159	0.0024
-0.0529	-0.0003	-0.0465	-0.0116	-0.0358	-0.0349
-0.0075	-0.0018	-0.0093	0.0003	-0.0074	0.0074
0.0127	-0.0166	0.0088	-0.0003	0.0100	-0.0174
-0.0156	0.0054	-0.0124	-0.0014	-0.0167	0.0112
0.0035	-0.0133	0.0045	-0.0065	0.0042	-0.0043
-0.0074	0.0143	-0.0079	0.0081	-0.0085	0.0118
-0.0010	0.0028	0.0001	0.0037	-0.0020	0.0045
0.0016	-0.0017	0.0005	0.0014	-0.0036	0.0060
0.0009	-0.0062	0.0010	-0.0014	0.0024	-0.0018
x	y	x	y	x	y
0.2053	-0.0150	0.2226	-0.0046	0.1987	-0.0011
-0.0103	0.1334	-0.0161	0.0317	-0.0160	0.0075
-0.0007	-0.0287	0.0013	-0.0021	0.0005	0.0003
-0.0003		0	-0.1403	0.0002	-0.0291
0.0010	-0.0150	0.0006	-0.0595	0.0004	-0.0009
-0.0004	0.1217	-0.0004	0.0559	-0.0002	-0.0020
	-0.1629				0.1118
0.0624		0.0563	-0.0045	0.0628	0.0810
-0.0083		-0.0067	0.0311	-0.0064	-0.0308
-0.0013	-0.0017	-0.0150	-0.0011		
0.0003	0.0002	-0.1379	0.0001	-0.0018	
0.0003	0.0002	-0.0407	0.0001	0.0120	
-0.0001	-0.0001	0.2012	0	-0.0038	
				-0.0457	
				-0.0031	
				0.0422	
				0.1594	
				0.0861	
				-0.3115	

TABLE 2

Key to Tables 2 to 6.

The E'_j and S_j are as defined on p. 87, in serial order, for $j = 1, 2, \cdots 10$. Under x are listed the coefficients of the expansion functions, in the order in which they are listed on p.79, with a gap between the symmetric and anti-symmetric ones. Under y are listed the coefficients of the core orbitals, in the order 1s, 2s, 2p, 3s, 3d, 3p, 4s, 4d, 4p, for atom A, followed after a gap by those for atom B. Our compounds are written in the form AB.

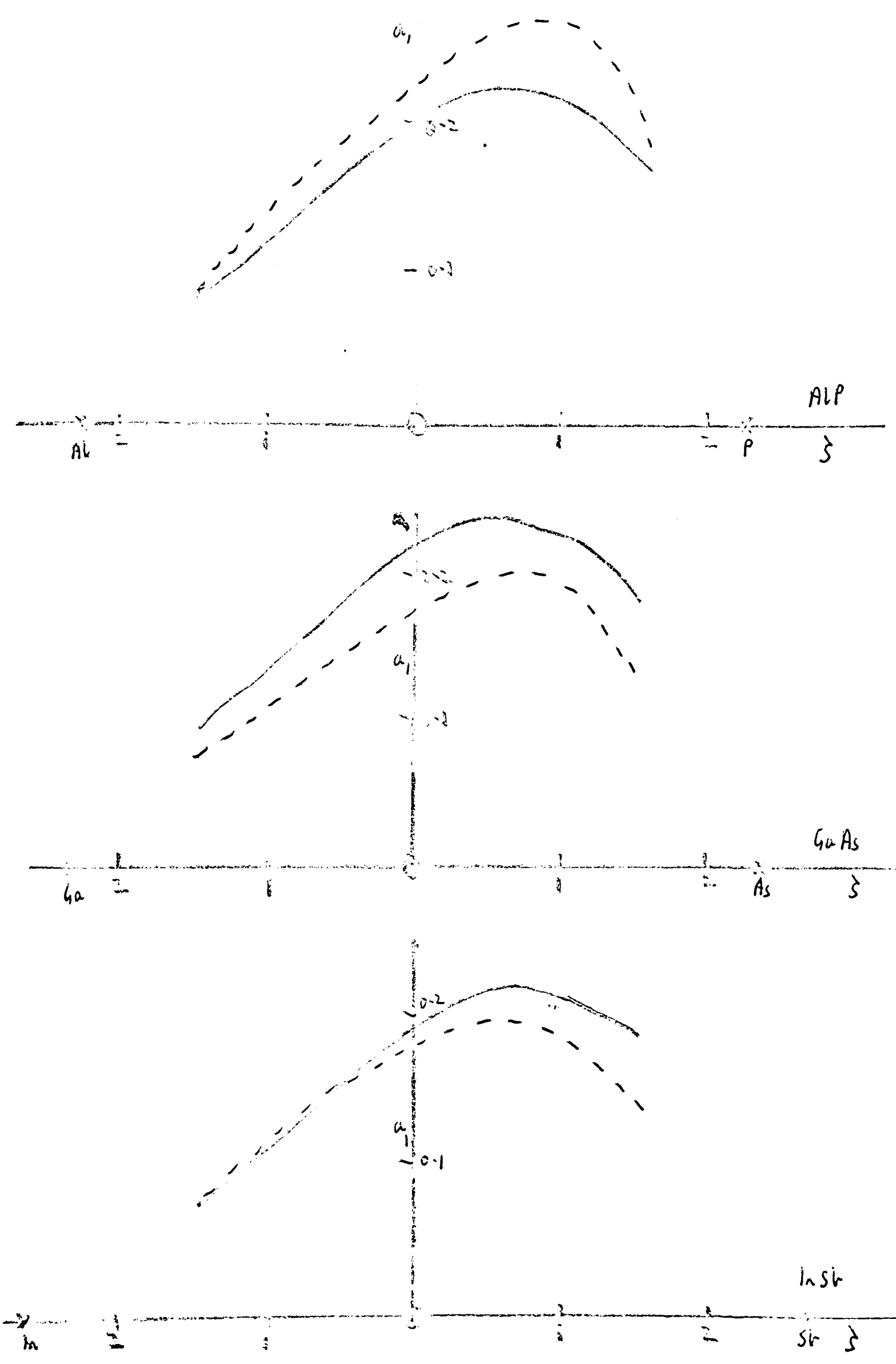
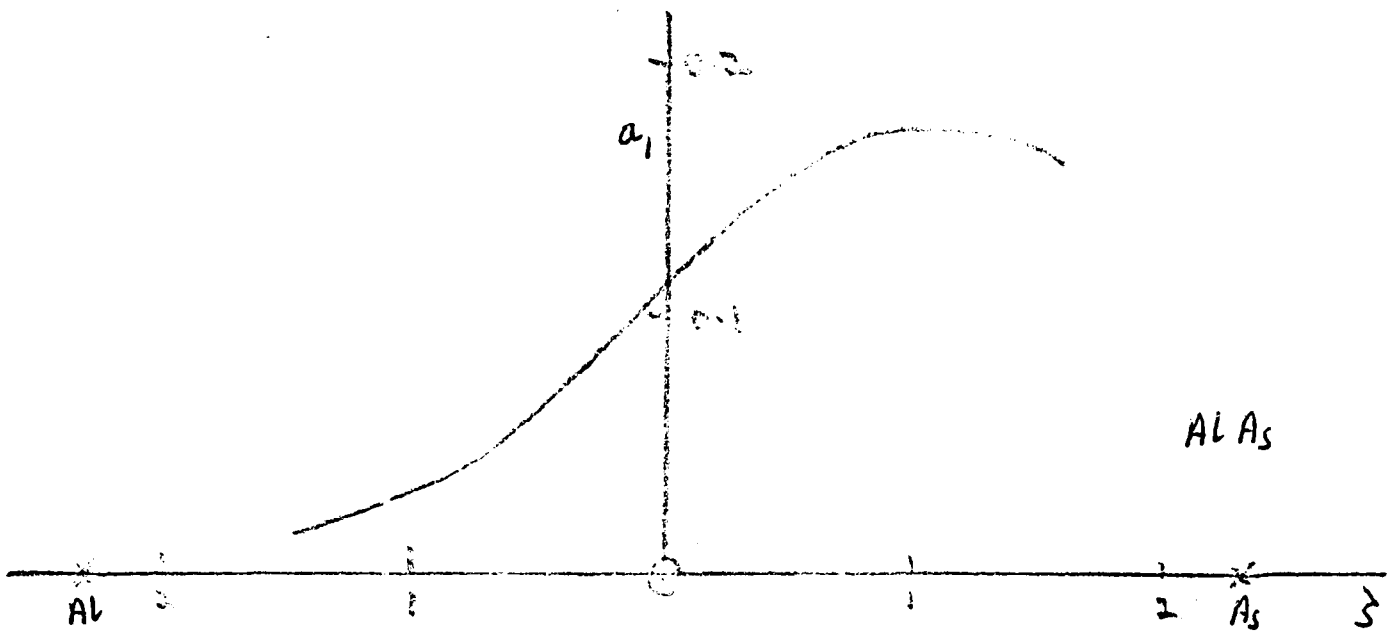


Figure 6. Equivalent orbitals (continuous line) and L.C.A.O. bond orbitals in III − V compounds with $Z_B = Z_A + 2$.

AlAs.		AlSb.		GaSb.	
E'_j	S_j	E'_j	S_j	E'_j	S_j
0.0431	0.0136	0.0223	0.0141	0.0235	0.0040
-0.0330	-0.0049	-0.0238	-0.0097	-0.0399	-0.0220
-0.0029	-0.0027	-0.0058	-0.0018	-0.0105	-0.0078
0.0142	-0.0138	0.0166	-0.0172	0.0062	-0.0263
-0.0299	0.0277	-0.0404	0.0230	-0.0381	0.0278
0.0032	-0.0132	0.0035	-0.0040	0.0033	-0.0035
-0.0060	+0.0105	-0.0050	0.0084	-0.0082	0.0103
-0.0038	0.0069	-0.0156	0.0191	-0.0112	0.0228
0.0016	0.0060	0.0029	0.0033	-0.0027	0.0104
0.0064	-0.0149	0.0915	-0.0144	0.0091	-0.0095
X	Y_A	X	Y_A	X	Y_A
0.1127	-0.0181	0.1315	-0.0169	0.1398	-0.0038
0.0058	0.1475	0.0022	0.1394	-0.0010	0.0261
-0.0011	0.1088	-0.0023	0.0463	-0.0014	0.0089
-0.0010		-0.0006		-0.0004	-0.1131
0.0006	-0.0035	0.0003	-0.0011	0.0002	-0.0255
-0.0002	0.0746	-0.0001	0.0072	-0.0001	-0.0951
	-0.0169		-0.0078		
0.0968	-0.1053	0.0611	-0.0278	0.0980	-0.0014
-0.0135	-0.0037	-0.0082	-0.0010	-0.0117	0.0089
-0.0021	0.2172	-0.0009	0.0417	-0.0023	- 0.0047
0.0004		0.0002	0.1037	0.0002	-0.0342
0.0005		0.0003	0.0186	0.0004	-0.0025
-0.0002		-0.0001	-0.2991	-0.0001	0.0516
					0.1241
					0.0597
					-0.3576

TABLE 3



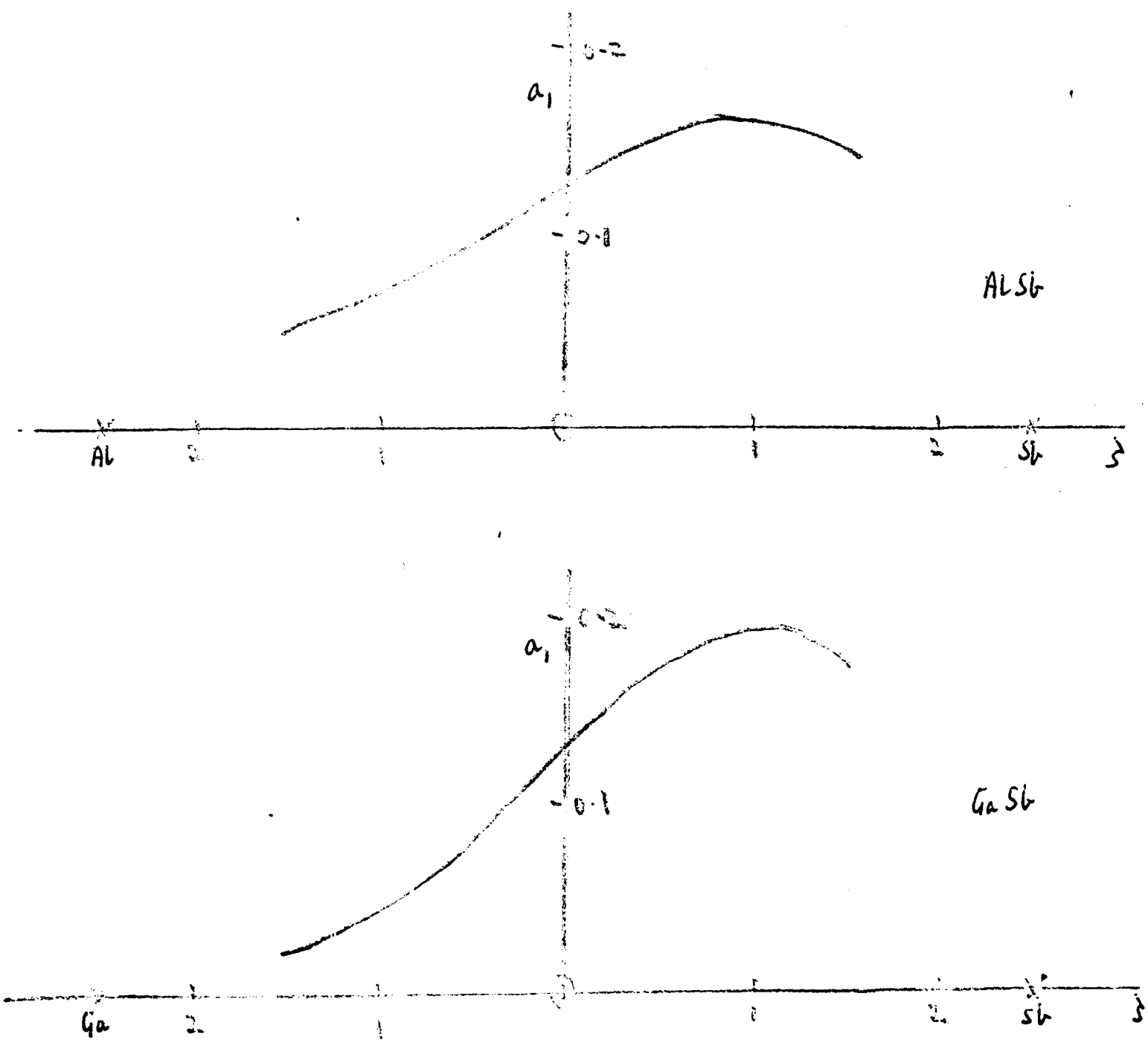
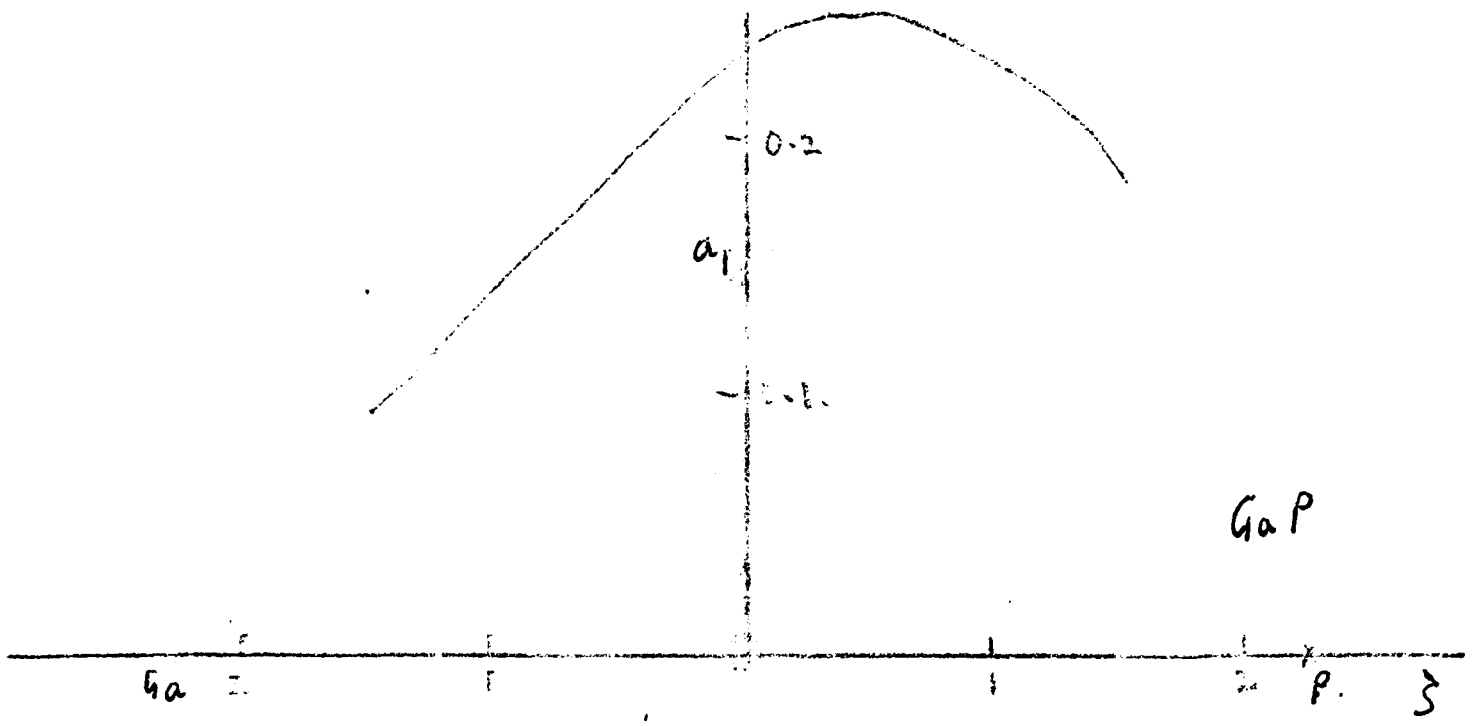
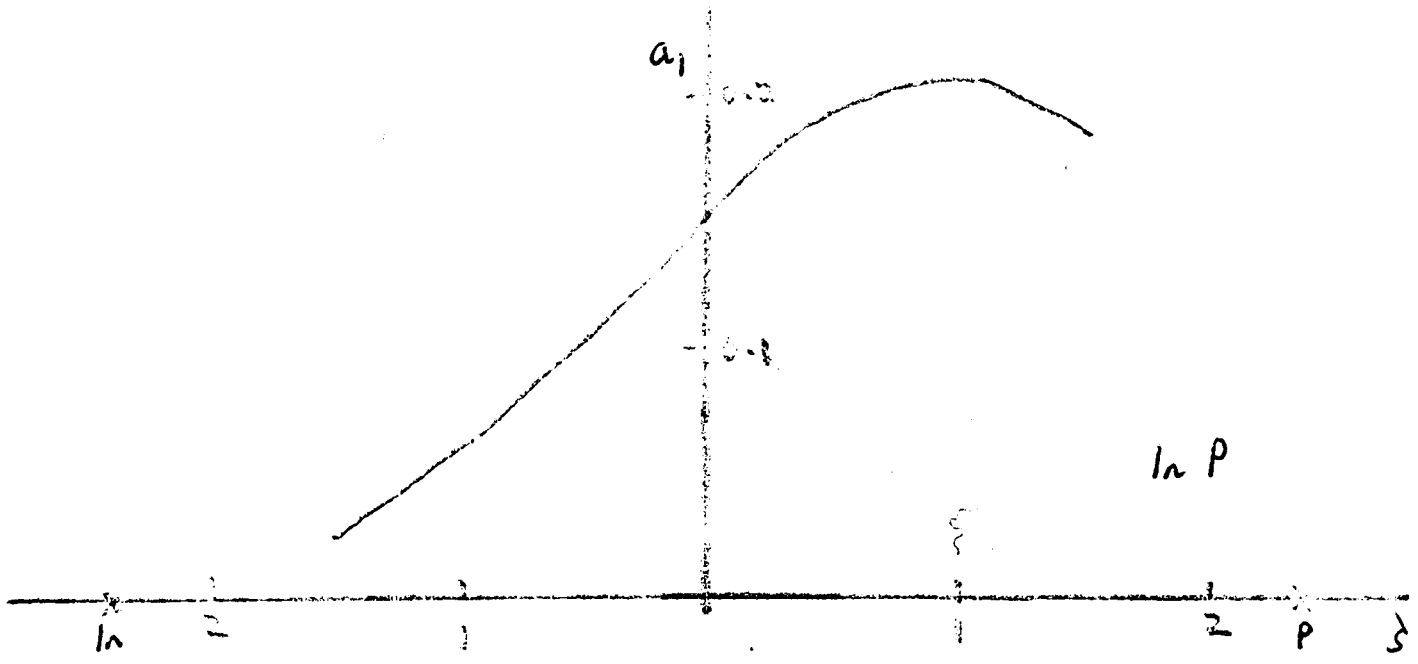


Figure 8. Equivalent orbitals in III-V compounds with $Z_B > Z_A + 2$.

GaP		InP		InAs	
E'_j	S_j	E'_j	S_j	E'_j	S_j
0.0009	0.0015	0.0227	-0.0023	0.0315	-0.0011
-0.0577	-0.0145	-0.0486	-0.0175	-0.0490	-0.0234
-0.0104	0.0005	-0.0119	-0.0038	-0.0117	-0.0044
0. 0099	-0.0224	0. 0090	-0.0257	0.0112	-0.0253
-0.1255	0.0006	-0.0183	0.0033	-0.0236	0.0031
0.0039	-0.0087	-0.0005	-0.0125	0.0012	-0.0100
-0.0030	0.0102	-0.0093	0.0147	-0.0103	-0.0129
0.0006	0.0029	0	0.0017	-0.0011	0.0050
0.0003	0.0003	-0.0025	0.0006	-0.0035	0.0054
0.0007	-0.0024	0.0006	-0.0039	0.0013	-0.0035
X	Y	X	Y	X	Y
0.2330	-0.0040	0.1586	-0.0002	0.1767	-0.0004
-0.0170	0. 0280	-0.0074	0.0011	-0.0091	0.0025
-0.0003	-0.0020	0.0005	0.0013	-0.0017	0.0010
0	-0.1296	-0.0002	-0.0062	-0.0002	-0.0113
0.0007	-0.0721	0.0008	0.0002	0.0007	-0.0003
-0.0004	0.0567	-0.0004	-0.0138	-0.0003	-0.0101
			0.0535		0.0693
0.0612	-0.0153	0.0959	0.0665	0.0846	0.0731
-0.0077	0.1252	-0.0097	0.0427	-0.0095	0.0157
-0.0018	-0.1624	-0.0018		-0.0014	
0 .0002	0.0003	-0.0192	0.0002	-0.0041	
0.0003	0.0002	0.1500	0.0002	0.0284	
-0.0001	-0.0001	-0.1683	-0.0001	-0.0155	
				-0.1269	
				-0.0363	
				0.2086	

TABLE 4





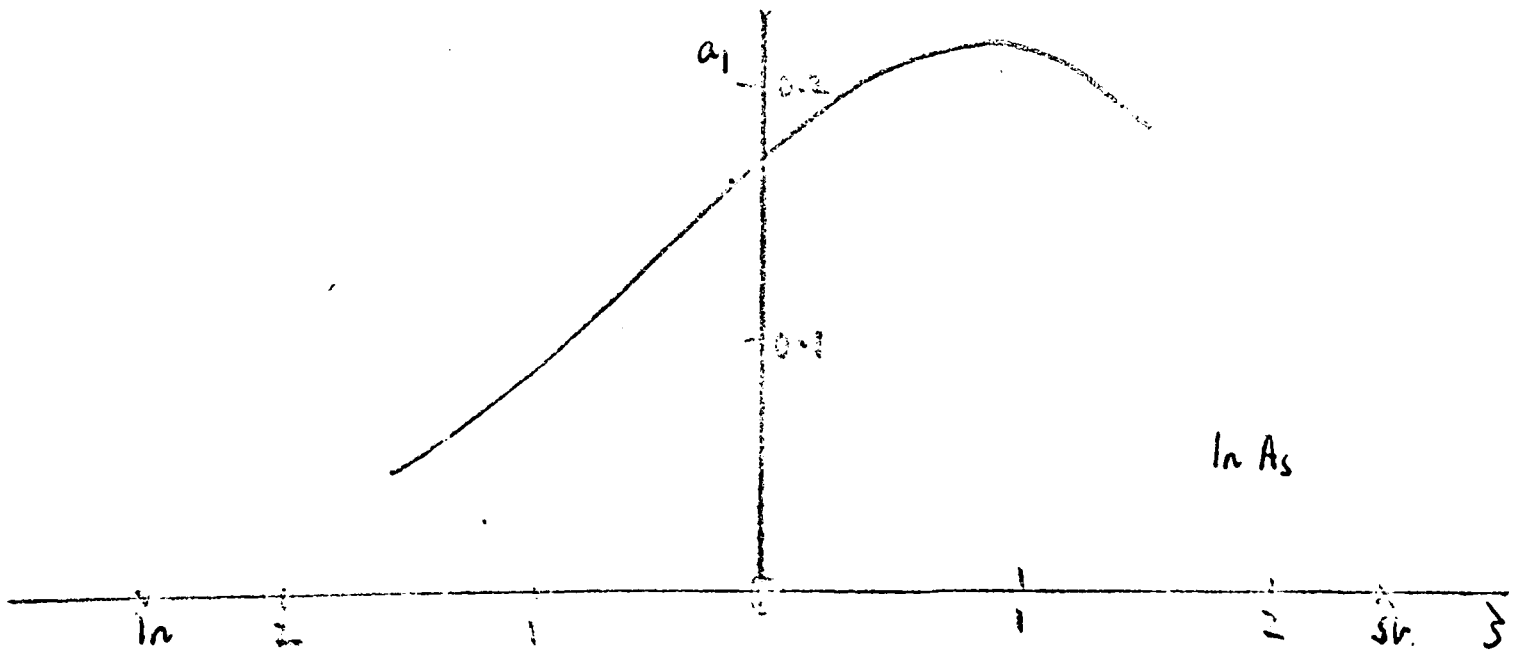
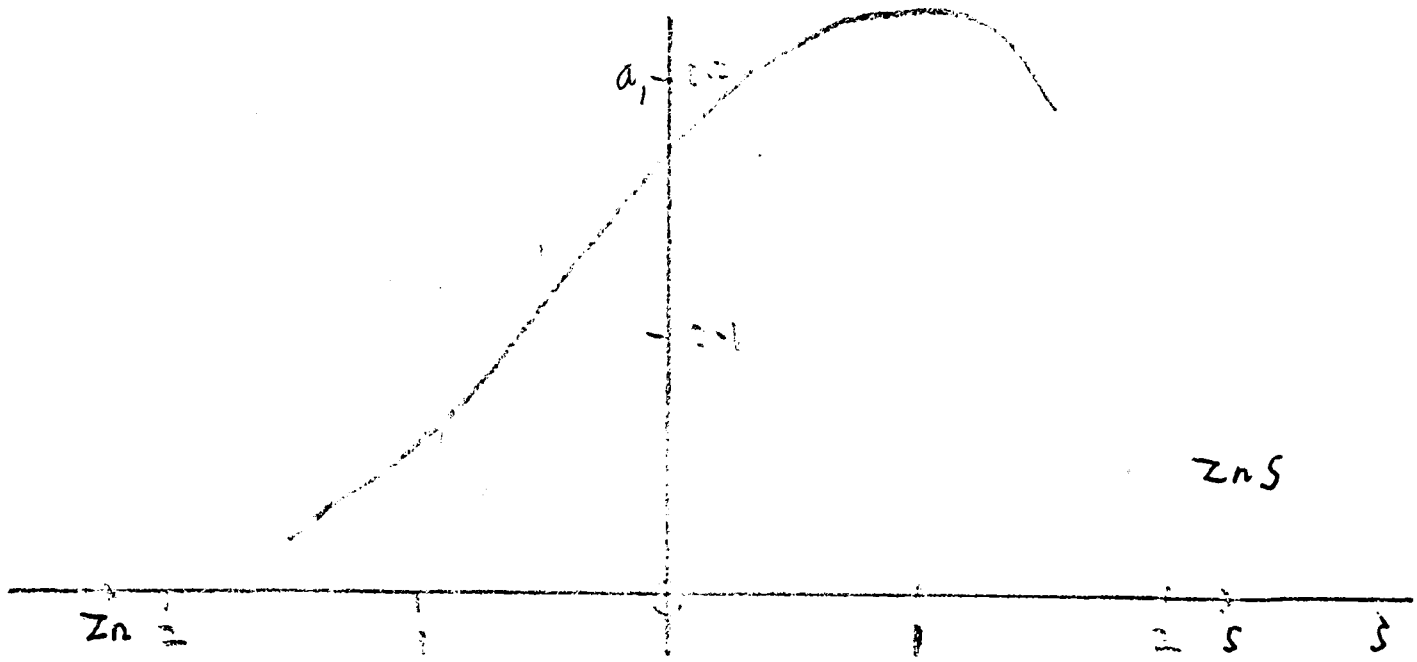


Figure 8. Equivalent orbitals in III − V compounds with $Z_A > Z_B$.

ZnS		ZnSe		ZnTe	
E'_j	S_j	E'_j	S_j	E'_j	S_j
0.0505	0.0083	0.0641	0.0133	0.0561	0.0144
-0.0577	-0.0021	-0.0342	-0.0067	-0.0038	-0.0057
-0.0057	-0.0006	-0.0046	0.0002	-0.0070	0.0016
0.0167	-0.0169	0.0130	-0.0219	0.0094	-0.0186
-0.0229	0.0151	-0.0292	0.0257	-0.0430	0.0326
0.0026	-0.0114	-0.0008	-0.0133	-0.0075	-0.0141
-0.1000	0.0150	-0.0080	0.0133	-0.0038	0.0054
-0.0027	0.0042	-0.0079	0.0075	-0.0192	0.0112
0.0004	0.0070	-0.0029	0.0106	-0.0038	0.0097
0.0019	-0.0091	0.0049	-0.0132	0.0097	-0.0166
X	Y	X	Y	X	Y
0.1720	-0.0021	0.1203	-0.0027	0.0469	-0.0032
-0.0101	0.0151	-0.0020	0.0189	0.0112	0.0222
0.0004	0.0086	0.0008	0.0137	-0.0030	0.0176
-0.0006	-0.0858	-0.0006	-0.0949	-0.0008	-0.1015
0.0016	-0.0773	0.0009	-0.0471	0.0002	-0.0170
-0.0006	-0.0703	-0.0003	-0.1438	0	-0.2145
0.1158	-0.0153	0.1252	-0.0046	0.1170	-0.0014
-0.0159	0.1210	-0.0174	0.0314	-0.0145	0.0093
-0.0025	-0.1647	-0.0022	-0.0177	-0.0021	-0.0042
0.0005		0.0005	-0.1317	0.0004	-0.0345
0.0005		0.0005	-0.0153	0.0003	-0.0016
-0.0002		-0.0002	0.2228	-0.0001	0.0460
					0.1114
					0.0193
					-0.3076

TABLE 5



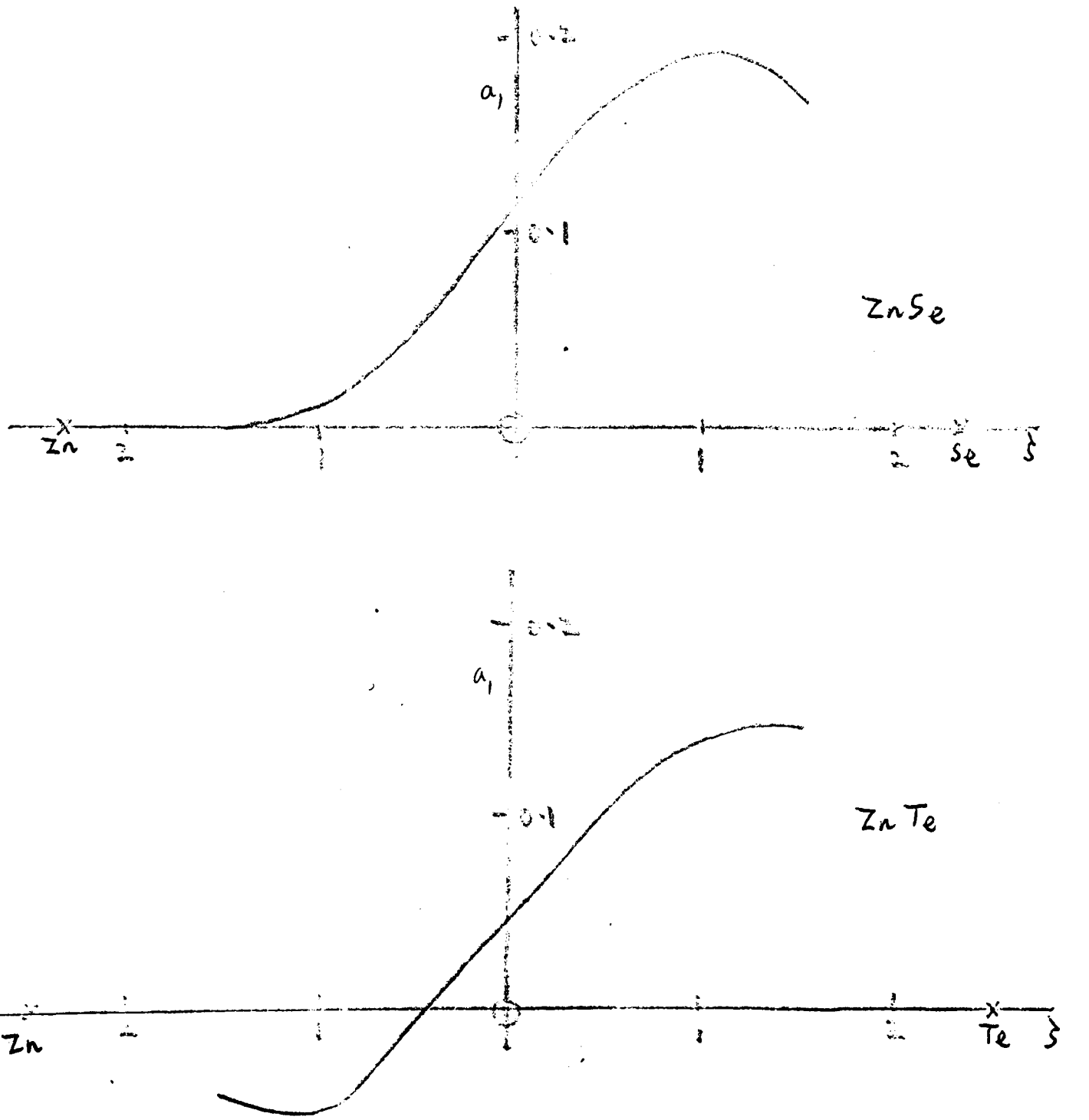
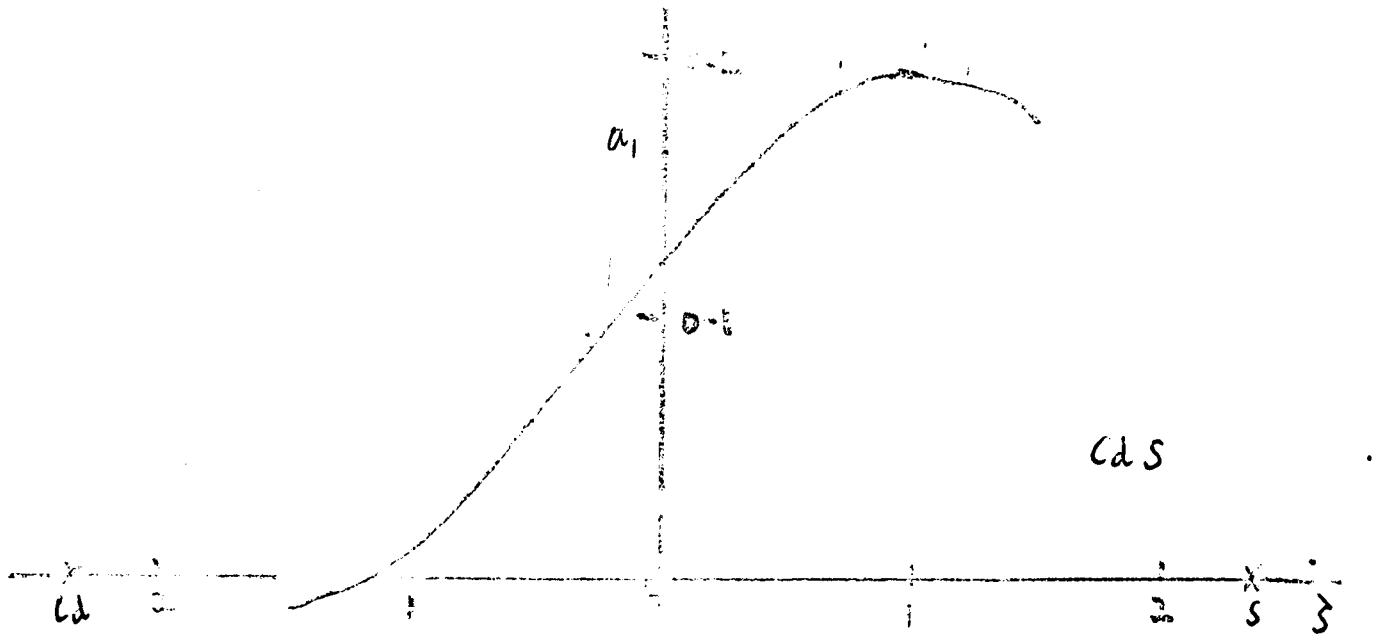


Figure 9. Equivalent orbitals in II-VI zinc compounds.

CdS		CdSe		ZnTe	
E'_j	S_j	E'_j	S_j	E'_j	S_j
0.0537	0.0131	0.0623	-0.0054	0.0404	-0.0055
-0.0363	-0.0048	-0.0344	0.0094	-0.0264	-0.0185
-0.0074	-0.0002	-0.0064	-0.0036	-0.0032	-0.0027
0.0130	-0.0210	0.0105	0.1320	0.0112	-0.0267
-0.0320	0.0297	-0.0320	-0.1739	-0.0246	0.0105
0.0035	-0.0106	-0.0002	0.0610	0.0012	-0.0084
-0.0134	0.0146	-0.0107	-0.0738	-0.0101	0.0130
-0.0042	x0.0042	-0.0068	-0.0329	-0.0345	0.0077
-0.0054	0.0161	-0.0071	-0.0823	-0.0044	0.0030
0.0037	-0.0107	0.0047	0.0621	0.0026	-0.0029
X	Y	X	Y	X	Y
0.1229	-0.0001	0.1017	-0.0004	0.1675	-0.0008
-0.0044	0.0010	-0.0025	0.0030	-0.0091	0.0056
-0.0017	0.0034	-0.0002	0.0037	-0.0004	0.0009
-0.0005	-0.0061	-0.0004	-0.0128	0	-0.0222
0.0013	0.0010	0.0008	0.0017	0 .0004	-0.0001
-0.0005	-0.0362	-0.0003	-0.0401	-0.0002	-0.0096
	0.0615		0.0735		0.0353
0.1220	0.0816	0.1216	0.0537	0.0735	0.0770
-0.0148	0.1317	-0.0145	0.2224	-0.0077	0.0147
-0.0023		-0.0019		-0.0011	
0.0003	-0.0115	0.0003	-0.0040	0.0001	-0.0013
0 .0004	0.0942	0.0003	0.0273	0.0002	0.0120
-0.0001	-0.1305	-0.0001	-0.0159	0	-0.0037
			-0.1173		-0.0452
			-0.0178		-0.0024
			0.2060		-0.0041
					0.1530
					0.0626
					-0.2936

TABLE 6



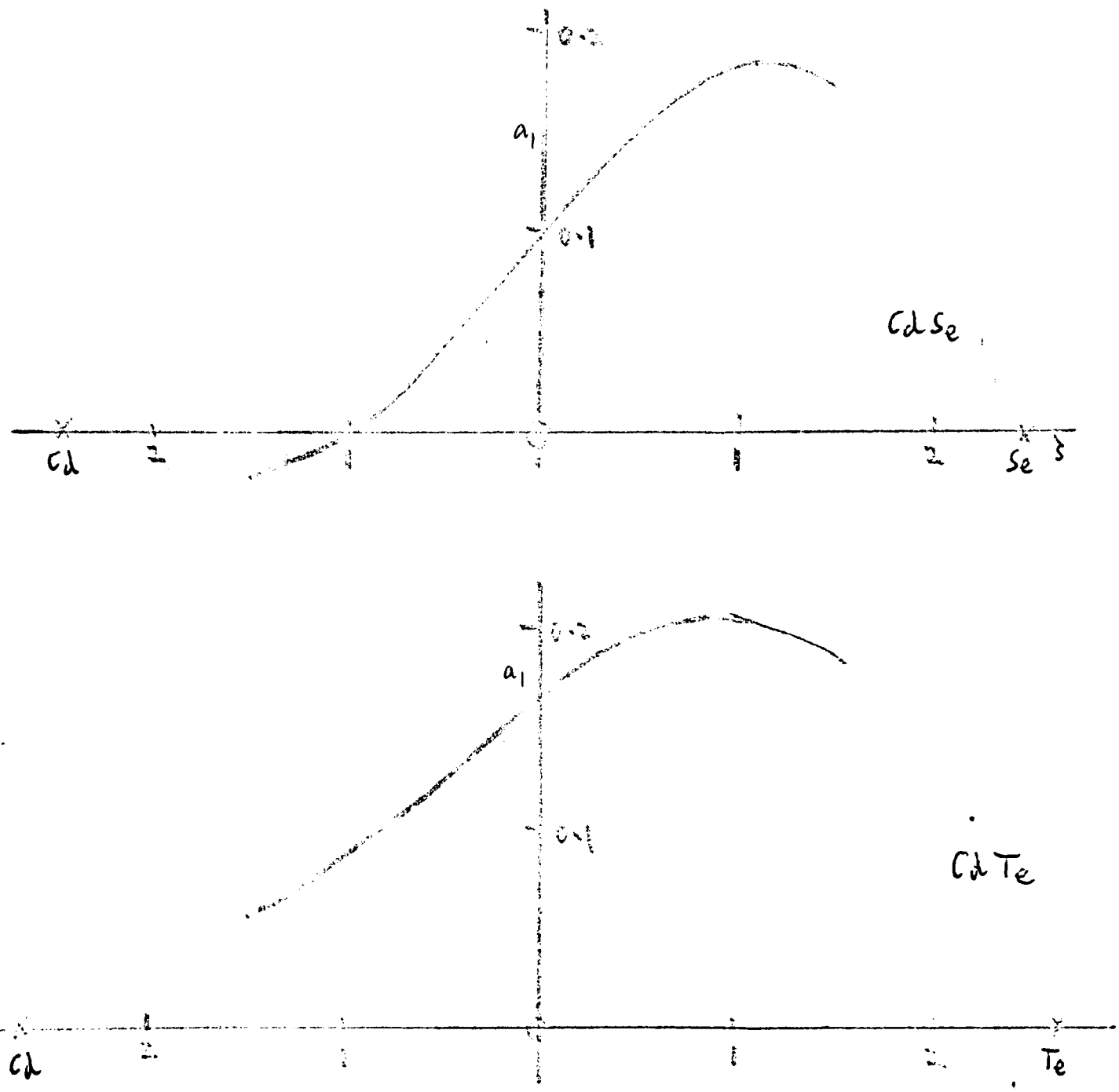


Figure 10. Equivalent orbitals in II – VI Cadmium compounds.

4.4 FURTHER DISCUSSION OF OUR RESULTS.

Experimental Results.

Most of the experimental evidence available[46] for comparison with the results of theoretical calculations on crystals concerns the general shape of the energy bands, the gap between different bands at various points of symmetry in reciprocal space, effective masses at these points, and similar quantities.[47] In view of our comments at the beginning of this chapter, these do not provide a very suitable test of the accuracy of our equivalent orbitals, since the correct orbitals and a slightly erroneous crystal potential could lead to large errors in the one-electron energies. The most natural test for our orbitals would concern such localised properties as effective charges, ionicity parameters, and charge density along a bond; only the last of these can be easily defined quantitatively for calculations such as ours in which the L.C.A.O. approximation is not employed.

Unfortunately, the amount of experimental evidence in these matters is very limited and often difficult to interpret. For instance, Wolff and Broder[48] attempt to estimate the percentage ionicities of a bond, and so the effective charges, by examining the proportions of different surfaces that appear when a crystal is cleaved into small parts. Picus et al.,[49] on the other hand, give a table of effective charges derived from infra-red (restrahlen) absorption. Such measurements are difficult to interpret on our scheme, since they involve associating with the individual atoms charges which we tend rather to associate with bonds. They would perhaps be most useful for the II - VI compounds if our bond

orbitals were reliable, since we would then naturally associate with each atom the charge lying between it and the node in the charge density; however, it would not be safe to draw any such conclusions about these compounds without further calculations. Incidentally, most of these methods give the magnitude, but not the sign, of the effective charges.

A somewhat more promising method might be that reported recently by Ugai and Domashevskaya,[50] who attempt to relate the charge distribution in III - V compounds to the L-lines in the emission spectra of the crystal. This method has an advantage over those previously mentioned in that it concerns only the equilibrium state of the crystal, without excessive lattice vibrations or surface effects, and could presumably be used with crystals at a very low temperature. It also reveals the sign of the effective charge, and they find that the antimony atom has a negative charge in Alsb and GaSb but is neutral in InSb. Our results, unlike those of Coulson, Redei, and Stocker,[20] would also suggest that there are less electrons, i.e. less negative charge, near the antimony atom in InSb than in the other two antimonides.

X-ray Scattering Measurements.

The most direct experimental evidence in connection with bond orbitals would come from observations of the variation in the charge density along a bond. It is unfortunate that X-ray scattering measurements only give some of the terms in a Fourier analysis of the charge density in the crystal, so that truncation errors are liable to arise on converting this into a charge distribution in real space. On the other hand, a comparison of the Fourier components

expected from our results with those found experimentally, although probably feasible, would not reveal whether any differences were due to the core electrons, the tails of the bond orbitals, or the behaviour of the bond orbitals in the region between adjacent atoms. In spite of the problems, Sirota and his co-workers,[51] using a new method of analysing the results, have published an extensive series of papers presenting the charge densities in real space for diamond, silicon, germanium, the arsenides and the antimonides. Unfortunately most of their measurements were made at 20°C, and only towards the end of the series did they discover that a lowering of the temperature led to an appreciable enhancement in the charge density along the bond; our caluculations, of course, refer to a perfect crystal at 0°K, so that our results would be expected to lead to higher charge densities in this region than they report, and in fact this happens quite consistently.

In silicon and germanium, they found that the charge density reached a minimum of about three fifths of an electron per cubic Angstrom at the mid-point of the bond (it is much higher in diamond) while our orbitals, perhaps because of the nature of the basis functions we used, give a shallow local maximum at this point with a charge density some 30% higher. For the arsenides and antimonides they find a lower charge density than this, on average about 0.4 electrons per cubic Angstrom, at the mid-point of the bond, and in general that most of the valence charge is nearer to the group V than to the group III atom. This is what would be expected intuitively, and agrees with our results. They find that the charge density at the mid-point of a bond in the antimonides is highest in AlSb and lowest in

InSb. This is to be expected, since a lighter atom such as Al should occupy a smaller proportion of the bond than does a heavy element like In, so that the mid-point of the bond is relatively farther from the group V atom in InSb than in GaSb and AlSb; however our results did not give this. On the other hand, in connection with our most reliable results, we agree with their finding that the charge density at the mid-point of the bond in GaAs is somewhat higher than it is in InSb. Our value for InSb is about 0.55 electrons per cubic Angstrom, while theirs is 0.30, but this difference could again be due to a temperature effect.

The Nature of our Wave Functions.

As mentioned at the end of Chapter 1, we attempted to compare the wave functions for silicon obtained by the L.C.A.O. method and by our bond orbital method with those obtained by Kleinman and Phillips,[52] for certain symmetry points, in their O.P.W. calculations.[36] This attempt was unfortunately not very successful, perhaps because of the different amounts of core functions contained in the different wave functions. However, it did throw a good deal of light on the differences between the different wave functions. As can be seen from figure 2, the L.C.A.O. wave functions for the valence band of silicon closely resemble those obtained by our equivalent orbital method in the region between adjacent atoms, i.e. along a bond, which is gratifying but not unexpected. However, the L.C.A.O. wave functions cannot be expected to have the correct behaviour at large distances from the bond, and thus we found on analysing them that their Fourier components for non-zero $\underline{K}$ were much smaller than

those obtained by the O.P.W. method. This is quite similar to what Parmenter[25] found, although his use of Gaussian functions for the atomic orbitals made his results rather suspect. Because an L.C.A.O. wave function is only really designed for molecules, it cannot allow for the translational symmetry of the crystal, and so its Fourier components decrease fairly smoothly as $\underline{K}$ increases instead of showing ~~sharp~~ peaks for those values of $\underline{K}$ compatible with the crystal symmetry.

The initial approximation to our bond orbitals, e^{-cr^2} behaves similarly, but the extra terms in the wave function, which are needed more to minimise the localisation potential than to minimise the energy of the orbital, lead to Fourier components of the correct order of magnitude. Although the coefficient of r^2e^{-cr2} in our bond orbitals is typically only a tenth that of e^{-cr2}, its Fourier component for $\underline{K} = (2, 0, 0)\ \pi/L$ is about five times as great. Thus these seemingly small terms in the bond orbital, which have only a little effect on the shape of the wave function in the bond region, have a considerable effect on the Fourier components of Bloch waves formed from these orbitals. It therefore seems that accurate calculations of bond orbitals by our method could well lead to orbitals having on the one hand a clearly recognisable and understandable shape in the bond regions and, on the other hand, the correct behaviour in other parts of the crystal, so that they could be used, for instance, in the calculation of an accurate, self-consistent, one-electron Hamiltonian.

Possible Improvements to the Calculations.

We feel that the results presented in this chapter, together with the preceding theory, certainly justify further calculations to investigate the value of our methods. It would be best to perform these on the new, larger, computers now coming into use, or at least on one in which unlimited additional storage space on magnetic tape is available, so that as many matrices as needed can be formed and stored. The first approximation which should be dropped in such a calculation is that of putting $TH^{in}\tilde{T} = TS^{in}\tilde{T}TH^{oo}\tilde{T}$, as this was only introduced because of the limited storage facilities on the Mercury computer; it is otherwise both unnecessary and undesirable, and omitting it would not lengthen the calculation very much. It would also be nice, and not too impracticable, to replace the Slater approximation for the valence-valence exchange potential by a more exact potential, perhaps only considering exchange between bonds no further apart than first or second neighbours. An increase in the number of basis functions, although very desirable, is probably not as essential as the above changes, but it would certainly be worthwhile to experiment with different values of C, and for the heteronuclear compounds with a different point on the bond about which to centre the basis functions.

If exact equivalent orbitals could be found, they would certainly increase further our understanding of what happens to the electrons when atoms combine to form a crystal, as well as providing us with valuable information on the density of states curve for the valence band. They would also

enable us to construct a convenient self-consistent potential for use in O.P.W. and similar calculations on the valence and conduction bands of these crystals. The equivalent orbital method itself cannot, of course, readily be used for calculations on conduction bands.

CHAPTER 5

OVERLAP FUNCTIONS IN PLANAR CRYSTALS.

5.1 THE METHOD OF CALCULATION.

Introduction.

It has been shown by Beattie and Landsberg[53] that the transition probability for the Auger effect, which is believed to be of importance in semiconductors such as indium antimonide, depends inter alia on the inter-band overlap integral of Bloch functions:-

$$I_n^{n'}(\underset{\sim}{k}, \underset{\sim}{k}') = \int u^*_{n,\underset{\sim}{k}}(\underset{\sim}{r})\, u_{n',\underset{\sim}{k}'}(\underset{\sim}{r})\, d\underset{\sim}{r} \qquad (5.1)$$

where, as in equation 1.62, $u_{n,\underset{\sim}{k}}(\underset{\sim}{r})$ is the translationally invariant part of the molecular orbital

$$\psi_{n,\underset{\sim}{k}}(\underset{\sim}{r}) = e^{i\underset{\sim}{k}.\underset{\sim}{r}}\, u_{n,\underset{\sim}{k}}(\underset{\sim}{r}) \qquad (5.2)$$

The volume of integration is one unit cell, n is a band label, and as previously $\underset{\sim}{k}$ is the reduced wave vector.

This integral is also involved in the probability of electron emission from a crystal.[54] It has been studied by a number of approximate methods, including first-order $\underset{\sim}{k}.\underset{\sim}{p}$ perturbation theory,[55] and also on the basis of a one-dimensional Kronig-Penney band model.[56] However, no previous attempt appears to have been made to study this integral using wave functions obtained from eigenvalue calculations on a real crystal. This is not too surprising, since such wave functions are usually complicated, of

unknown accuracy, and only available at a few points of special symmetry in reciprocal space. However, Coulson and Taylor[16] have found approximate L.C.A.O. wave functions of a simple analytical form at all points in reciprocal space for the valence and conduction bands of graphite and planar boron nitride, when these crystals are regarded as containing a single isolated plane of atoms. It was therefore thought of interest to evaluate $I(\underset{\sim}{k}, \underset{\sim}{k}')$ for these substances, and to compare its behaviour with the previous predictions. There are often considerable differences between the results of quantum mechanical calculations in two or three dimensions and those in one dimension. This, and the probable effects of the presence of two different atoms in a unit cell, suggested that our results might well differ from those of Beattie and Landsberg.

The Band Structure of Graphite and Planar Boron Nitride.

Both graphite and planar boron nitride contain, in addition to the core electrons, σ and π electrons which to a first approximation can be treated as independent of each other. In that case, it is the π electrons which give rise to the highest valence band and lowest conduction band, and these are the only ones we consider. It is to be expected that the presence of the σ electrons will cause some modification to the bottom of the valence and top of the conduction bands, but their effects on the parts of the bands nearest to the energy gap should be small. It is these regions that are of most interest to us. In boron nitride there is an appreciable energy gap, but in graphite this gap vanishes and the valence and conduction bands tough.

We note in passing that a more sophisticated treatment, which allows for the presence of other planes of atoms, leads to an overlap between these bands in graphite, and so accounts for its good electrical conductivity. Such a treatment makes the wave functions more complicated, and since it also does not correspond to any situation of interest for semi-conductors having the zincblende structure we decided to ignore it. In both graphite and planar boron nitride, the adjacent band extrema are not at $\underset{\sim}{k} = \underset{\sim}{0}$, but lie at the edge of the first Brillouin zone; this is not the case for most tetrahedral semi-conductors.

The graphite or planar boron nitride lattice, regarded as a two-dimensional array of unit cells, is as shown in figure M11. The unit cell is a rhombus of side a, with angles of 120° and 60° at its vertices, as shown in figure 12. We choose a set of right-handed orthogonal axes, with the x-axis parallel to the long diagonal and the y-axis parallel to the short diagonal of the rhombus. We assume that the crystal contains M unit cells, with respect to which the usual Born-von Karman cyclic boundary conditions are applied.

All the unit cells can be obtained from any given one through translations by the lattice vectors

$$\underset{\sim}{R}_n = n_1 \underset{\sim}{a}_1 + n_2 \underset{\sim}{a}_2 \qquad (5.3)$$

where n_1 and n_2 are positive or negative integers or zero and the $\underset{\sim}{a}_j$ are the basis translational vectors. With our system of co-ordinates, we can take

$$\underset{\sim}{a}_1 = a(\sqrt{3}/2, -\tfrac{1}{2}) \qquad (5.4)$$

and

$$\underset{\sim}{a}_2 = a(\sqrt{3}/2, \tfrac{1}{2}) \qquad (5.4a)$$

It is convenient to define in addition a vector

$$\underset{\sim}{h} = \tfrac{1}{6}(\underset{\sim}{a}_1 + \underset{\sim}{a}_2) \qquad (5.5)$$

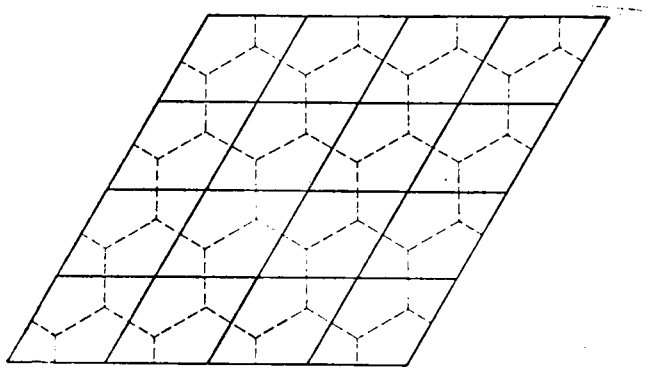


FIG. 1. The graphite lattice.

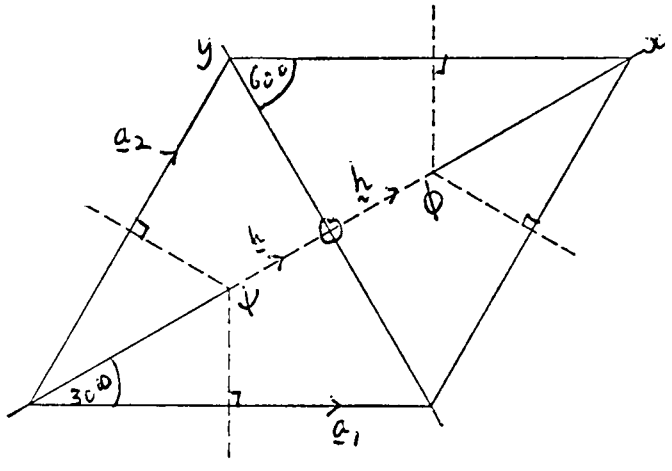


FIG. 2. The unit cell for graphite.

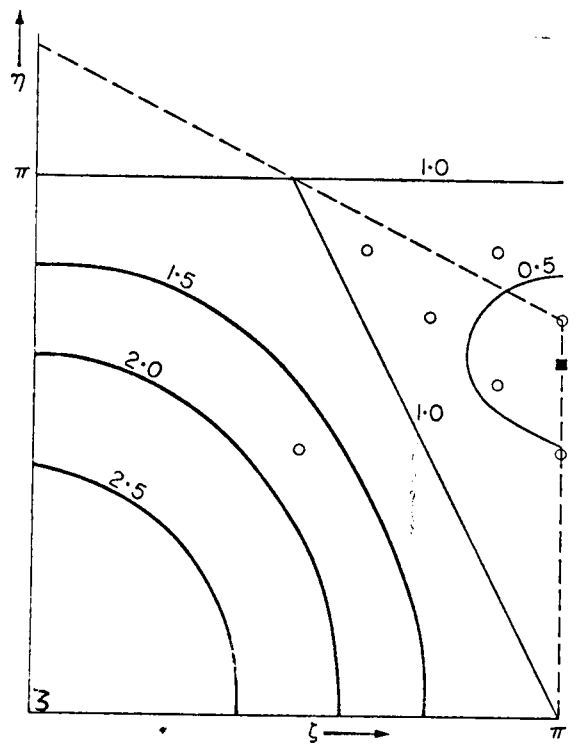


FIG. 3. $g(k)$ as a function of ζ and η. The full lines are the contours $g = 0{\cdot}5$, $1{\cdot}0$, $1{\cdot}5$, $2{\cdot}0$, and $2{\cdot}5$; $g = 3$ at $(0, 0)$ and $g = 0$ at $(\pi, 2\pi/3)$, the band edge, which is marked by a black square. It should be noted that $g(\pm\zeta, \pm\eta) = g(\zeta, \eta)$. The broken lines mark the edge of the first Brillouin zone, and the circles indicate the values of k for which the F_j^2 were evaluated.

equal to half the distance between the two atoms in the same unit cell.

In reciprocal space, we choose orthogonal axes ξ and η, and write

$$\underline{k} = \frac{1}{a}(\xi, \eta) \qquad (5.6)$$

It is convenient to introduce a change of scale by writing

$$\xi = 2\xi/\sqrt{3} \qquad (5.7)$$

and all co-ordinates in reciprocal space will henceforth be expressed in terms of ξ and η. The first Brillouin zone, part of which is shown in figure 13, is the hexagon formed by the perpendicular bisectors of the joins of (0, 0) to $(\pm\pi, \pm 2\pi)$ and $(\pm 2\pi, 0)$. Adjacent band extrema lie at $(\pm\pi, \pm 2\pi/3)$, and the bottom of the valence band and the top of the conduction band lie at (0, 0).

Figure 12 shows that there are two atoms in each unit cell. We call the corresponding $2p\pi$ orbitals $\psi(\underline{r} + \underline{h})$ and $\phi(\underline{r} - \underline{h})$, where $\underline{r}$ is measured from the centre of the cell. Let

$$S = \int \psi(\underline{r}+\underline{h})\,\phi(\underline{r}-\underline{h})\,d\underline{r} \qquad (5.8)$$

and let $E_{\underline{k}}$, the energy of the state with translational symmetry $\underline{k}$, be expressed in the form

$$E_{\underline{k}} = Z\gamma \qquad (5.9)$$

where $E_{\underline{k}}$ is measured from the centre of the energy gap and γ, defined by

$$\gamma = \beta - E_0 S \qquad (5.10)$$

is a unit of energy associated with the resonance integral β and the mean energy E_0 of the atomic orbitals ϕ and ψ.[41.] If we let t denote the ratio γ/β of the difference between the energy of these orbitals and β, and write

$$f = 1 + 2e^{i\xi}\cos\tfrac{1}{2}\eta \qquad (5.11)$$

and

$$g^2 = ff^* = 3 + 2\cos\eta + 4\cos\xi\cos\tfrac{1}{2}\eta \qquad (5.12)$$

then $$Z = (1 - g^2 S^2)^{-1} \left[-g^2 S \pm \left(g^2 + t^2 - t^2 g^2 S^2 \right)^{\frac{1}{2}} \right] \qquad (5.13)$$

The molecular orbitals $\psi_{\underset{\sim}{k}}$ are given by

$$\psi_{\underset{\sim}{k}} = \frac{1}{N_{\underset{\sim}{k}} \sqrt{M}} \sum_n e^{i \underset{\sim}{k} \cdot \underset{\sim}{\beta}_n} \left[\psi(\underset{\sim}{r} - \underset{\sim}{\beta}_n + \underset{\sim}{b}) + c_{\underset{\sim}{k}} \phi(\underset{\sim}{r} - \underset{\sim}{\beta}_n - \underset{\sim}{b}) \right] \qquad (5.14)$$

where

$$c_{\underset{\sim}{k}} = \frac{1 - ZS}{Z - t} f \equiv q f, \text{ say,} \qquad (5.15)$$

and

$$N_{\underset{\sim}{k}}^2 = 1 + g^2 (2qS + q^2) \qquad (5.16)$$

According to Coulson and Taylor, t = -1 for planar boron nitride, while for graphite $\phi = \psi$ and so t = 0. In this latter case, equation 5.13 reduces to

$$Z = \pm g / (1 \pm gS) \qquad (5.17)$$

while $$q = \pm 1/g \qquad (5.18)$$

In each case, the + sign corresponds to the valence band and the - sign to the conduction band. It should be noted that, unless t = 0, $C_{\underset{\sim}{k}} = 0$ at $(\pi, 2\pi/3)$, while if t = 0 $C_{\underset{\sim}{k}}$ has an essential singularity there since the two bands are then degenerate at this point.

For both substances, if Slater orbitals[21] are used S lies between 0.2 and 0.25, while slightly larger values might be appropriate if self-consistent-field orbitals were used.

Method of Calculation.

Let $$\underset{\sim}{l} = \underset{\sim}{k} - \underset{\sim}{k}^1. \qquad (5.19)$$

If $\underset{\sim}{k}$ and $\underset{\sim}{k}^1$ are both in the first Brillouin zone then

$$\int u_{\underset{\sim}{k}}^*(\underset{\sim}{r}) u_{\underset{\sim}{k}^1}(\underset{\sim}{r}) d\underset{\sim}{r} = \int e^{i \underset{\sim}{l} \cdot \underset{\sim}{r}} \psi_{\underset{\sim}{k}}^*(\underset{\sim}{r}) \psi_{\underset{\sim}{k}^1}(\underset{\sim}{r}) d\underset{\sim}{r} \qquad (5.20)$$

Now if $\underset{\sim}{K}$ is a basis vector of the reciprocal lattice, $\psi_{\underset{\sim}{K}+\underset{\sim}{k}}(\underset{\sim}{r}) = \psi_{\underset{\sim}{k}}(\underset{\sim}{r})$ so that $\psi_{\underset{\sim}{k}}$ is periodic in $\underset{\sim}{k}$. On the other hand,

$$U_{\underset{\sim}{K}+\underset{\sim}{k}}(\underset{\sim}{r}) = e^{-i\underset{\sim}{K}\cdot\underset{\sim}{r}}\, u_{\underset{\sim}{k}}(\underset{\sim}{r}) \qquad (5.21)$$

so that $U_{\underset{\sim}{k}}(\underset{\sim}{r})$ is not periodic in $\underset{\sim}{k}$. It was therefore decided to take as the fundamental form for $I(\underset{\sim}{k}, \underset{\sim}{k}')$ in equation 5.1

$$I_n^{n'}(\underset{\sim}{k}, \underset{\sim}{k}') = \int e^{i\underset{\sim}{k}\cdot\underset{\sim}{r}}\, \psi^*_{n,\underset{\sim}{k}}(\underset{\sim}{r})\, \psi_{n',\underset{\sim}{k}'}(\underset{\sim}{r})\, d\underset{\sim}{r} \qquad (5.22)$$

This distinction is important for planar crystals, where the adjacent band extrema lie at the edge of the first Brillouin zone, but not for tetrahedral crystals where the relevant extrema are usually at the centre of the zone.

Let us write

$$I_0(\underset{\sim}{l}, \underset{\sim}{\varrho}) = \int e^{i\underset{\sim}{l}\cdot\underset{\sim}{r}}\, \phi^*(\underset{\sim}{r}-\underset{\sim}{\varrho})\, \psi(\underset{\sim}{r}+\underset{\sim}{\varrho})\, d\underset{\sim}{r} \qquad (5.23)$$

Then if we substitute ~~form~~ from equation 5.13 into equation 5.22 and, following Coulson and Taylor, ignore all overlap integrals not on the same or adjacent atoms, we find that

$$I_n^{n'}(\underset{\sim}{k}, \underset{\sim}{k}') = \frac{1}{N_{\underset{\sim}{k}} N_{\underset{\sim}{k}'}} \left[\rho_1(\underset{\sim}{l}) + c^*_{\underset{\sim}{k}}\, \rho_2(\underset{\sim}{l}) + c_{\underset{\sim}{k}'}\, \rho_3(\underset{\sim}{l}) + c^*_{\underset{\sim}{k}}\, c_{\underset{\sim}{k}'}\, \rho_4(\underset{\sim}{l}) \right] \qquad (5.24)$$

where

$$\rho_1(\underset{\sim}{l}) = e^{-i\underset{\sim}{l}\cdot\underset{\sim}{h}} \int e^{i\underset{\sim}{l}\cdot\underset{\sim}{r}}\, \psi^*(\underset{\sim}{r})\, \psi(\underset{\sim}{r})\, d\underset{\sim}{r}$$

$$\rho_2(\underset{\sim}{l}) = I_0(\underset{\sim}{l}, \underset{\sim}{h}) + e^{i(\underset{\sim}{k}+\underset{\sim}{k}')\cdot\underset{\sim}{a}_1/2}\, I_0(\underset{\sim}{l}, \underset{\sim}{h}-\underset{\sim}{a}_1/2) + e^{i(\underset{\sim}{k}+\underset{\sim}{k}')\cdot\underset{\sim}{a}_2/2}\, I_0(\underset{\sim}{l}, \underset{\sim}{h}-\underset{\sim}{a}_2/2)$$

$$\rho_3(\underset{\sim}{l}) = I_0^*(\underset{\sim}{l}, \underset{\sim}{h}) + e^{-i(\underset{\sim}{k}+\underset{\sim}{k}')\cdot\underset{\sim}{a}_1/2}\, I_0^*(\underset{\sim}{l}, \underset{\sim}{h}-\underset{\sim}{a}_1/2) + e^{-i(\underset{\sim}{k}+\underset{\sim}{k}')\cdot\underset{\sim}{a}_2/2}\, I_0^*(\underset{\sim}{l}, \underset{\sim}{h}-\underset{\sim}{a}_2/2)$$

and $$\rho_4(\underset{\sim}{l}) = e^{i\underset{\sim}{l}\cdot\underset{\sim}{h}} \int \phi^*(\underset{\sim}{r})\, \psi(\underset{\sim}{r})\, d\underset{\sim}{r} \qquad (5.25)$$

It did not prove possible to find a closed analytical form for $I_0(\underset{\sim}{l}, \underset{\sim}{\sigma})$. We therefore expanded $e^{i\underset{\sim}{l}.\underset{\sim}{r}}$ in powers of $\underset{\sim}{l}.\underset{\sim}{r}$ and integrated term by term, using Slater orbitals for ϕ and ψ and working in bipolar co-ordinates. This led to a double power series for $I_0(\underset{\sim}{l}, \underset{\sim}{\sigma})$, viz.,

$$I_0(\underset{\sim}{l}, \underset{\sim}{\sigma}) = \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \beta_{rs} (\underset{\sim}{l}.\underset{\sim}{\sigma})^r |\underset{\sim}{l} \times \underset{\sim}{\sigma}|^s \qquad (5.26)$$

which was rapidly convergent if $|\underset{\sim}{l}|\,|\underset{\sim}{\sigma}| < 1$. Since we are most interested in small values of $\underset{\sim}{l}$, and for every $\underset{\sim}{\sigma}$ we use $|\underset{\sim}{\sigma}|\,|\underset{\sim}{l}| \leq {}^{2\pi}/_{3}$ if $\underset{\sim}{l}$ is in the first Brillouin zone, (5.27) this restriction is not too serious.

We write $$F_j = |I_n^{n'}(\underset{\sim}{k}, \underset{\sim}{k}')| \qquad (5.28)$$

where $j = 1$ if n and n^1 both refer to the conduction band,

$j = 2$ if n^1 refers to the valence band and n to the conduction band,

$j = 3$ if n refers to the valence band and n^1 to the conduction band,

and $j = 4$ if n and n^1 both refer to the valence band.

The values of F_j^2 were found for seven different values of $\underset{\sim}{k}$, shown in figure 13, and for $\underset{\sim}{l} = (l_1, l_2)$ with l_1 and l_2 lying between ${}^{\pi}/_{2}$ and ${}^{-\pi}/_{2}$. Figure 13 also shows the contours of constant g, which are lines of constant energy in reciprocal space.

5.2 THE VALUES OF THE OVERLAP INTEGRALS.

Planar Boron Nitride.

The results of the calculations on planar boron nitride

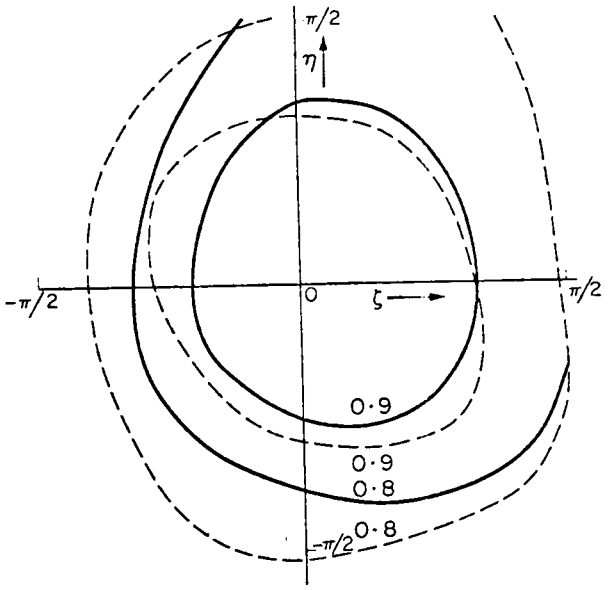


FIG. 14. Contour diagrams of F_1^2 for planar boron nitride. The contours $F_1^2 = 0{\cdot}9$ and $F_1^2 = 0{\cdot}8$ are plotted as functions of $l = k - k'$ for $k = (7\pi/8,\ 5\pi/8)$ (full lines) and $k = (5\pi/8,\ 7\pi/8)$ (broken lines).

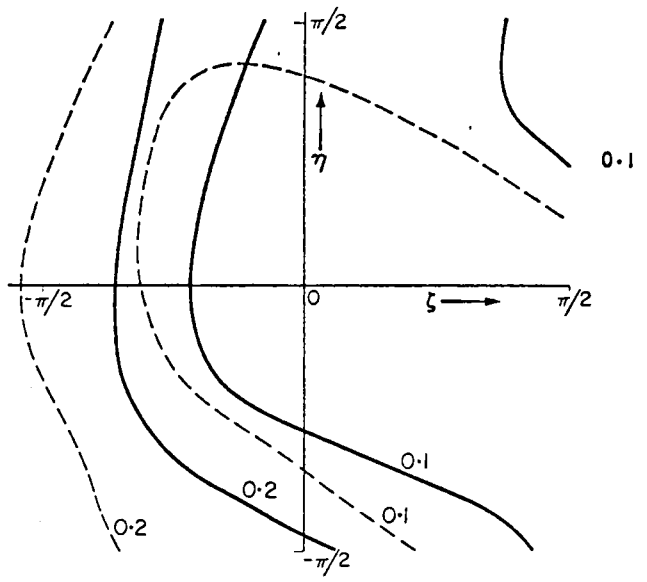


FIG. 15. Contour diagrams of F_2^2 for planar boron nitride. The contours $F_2^2 = 0{\cdot}1$ and $F_2^2 = 0{\cdot}2$ are plotted as functions of l for $k = (7\pi/8,\ 5\pi/8)$ (full lines) and $k = (5\pi/8,\ 7\pi/8)$ (broken lines).

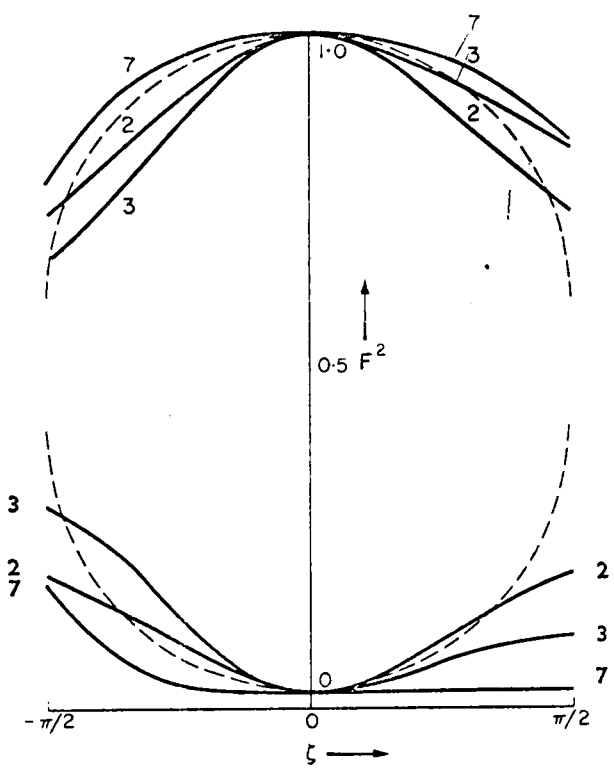


FIG. 16. F_1^2 (upper curves) and F_2^2 (lower curves) for boron nitride as functions of $l = (\zeta, 0)$ for $k = (\pi, 3\pi/4)$, $(7\pi/8,\ 5\pi/8)$, and $(\pi/2,\ \pi/2)$, numbered 2, 3, and 7 respectively. The broken lines are arcs of circles centred at $F^2 = 0{\cdot}6$ and $0{\cdot}4$, passing through these values at $\zeta = \pm\,\pi/2$, and touching the ζ-axis at $F^2 = 1$ and 0 respectively.

are shown in figures 14 to 16 and in table 7. In figure 14, the contours $F_1^2 = 0.9$ and 0.8 are plotted as functions of $\underset{\sim}{l}$ for two typical points, $\underset{\sim}{k} = (7\pi/8, 5\pi/8)$ and $\underset{\sim}{k} = (5\pi/8, 7\pi/8)$, in reciprocal space, and in figure 15 the contours $F_2^2 = 0.1$ and 0.2 are plotted similarly. The first of these points is much closer than the second to the band extrema both in energy and in reciprocal space. It will be observed that F_1^2 seems to depend more on the magnitude than on the direction of the vector $\underset{\sim}{l}$. This result was quite generally true for planar boron nitride, although directional effects tend to increase as $\underset{\sim}{k}$ approaches the band extrema at $(\pi, 2\pi/3)$ Figure 16 shows the variation of F_1^2 and F_2^2 along the line $\underset{\sim}{l} = (\zeta, 0)$ for three different values of $\underset{\sim}{k}$; similar results were found for other lines through the point $\underset{\sim}{l} = 0$ and for other values of $\underset{\sim}{k}$.

The behaviour of F_4^2 is similar to that of F_1^2, except that at almost all points F_4^2 is slightly greater than F_1^2, the difference between these quantities tending to increase somewhat irregularly as $|\underset{\sim}{l}|$ increases. This suggests that $u_{\underset{\sim}{k}}$ for the valence band varies less rapidly with $\underset{\sim}{k}$ than does $u_{\underset{\sim}{k}}$ for the conduction band, as is to be expected. Table 5 contains the mean value of each F_j^2 for $\underset{\sim}{l} = (\pm \pi/4, 0)$ and $(0, \pm \pi/4)$ for all the seven values of $\underset{\sim}{k}$ for which calculations were carried out. F_3^2 behaves in a similar manner to F_2^2, although the difference between these quantities is often as much as 0.02, or 10%.

Index no.	$\underline{k}$	ε
1	$(\pi, \pi/2)$	0.4142
2	$(\pi, 3\pi/4)$	0.2346
3	$(7\pi/8, 5\pi/8)$	0.4260
4	$(7\pi/8, 7\pi/8)$	0.6567
5	$(3\pi/4, 3\pi/4)$	0.7095
6	$(5\pi/8, 7\pi/8)$	0.9239
7	$(\pi/2, \pi/2)$	1.7321

Planar Boron Nitride

No.	F_1^2	F_2^2	F_3^2	F_4^2
1	0.909	0.073	0.073	0.917
2	0.904	0.077	0.078	0.913
3	0.908	0.074	0.074	0.917
4	0.942	0.058	0.058	0.933
5	0.942	0.058	0.058	0.933
6	0.931	0.047	0.047	0.944
7	0.965	0.010	0.010	0.980

Graphite.

For the first five values of $\underline{k}$, the deviation of F_j^2 from its mean value exceeded 0.15.

6	0.888	0.100	0.100	0.892
7	0.972	0.008	0.008	0.981.

TABLE # 7. The mean value of F_j^2, averaged over the points $\underline{l} = (\pi/4, 0), (-\pi/4, 0), (0, \pi/4)$ and $(0, -\pi/4)$.

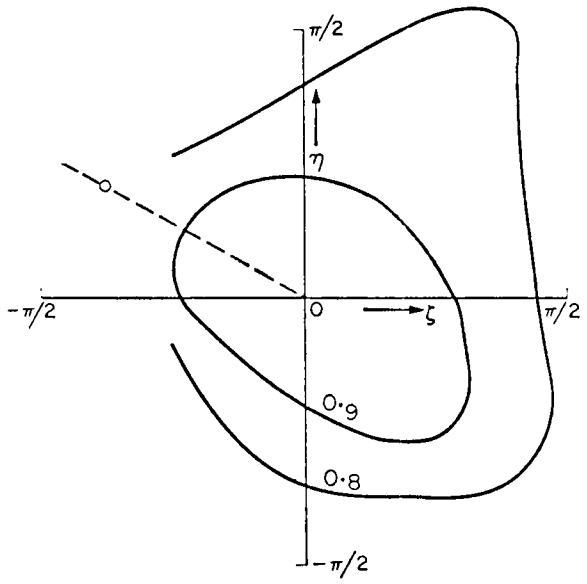


FIG. 17. Contour diagram of F_1^2 for graphite, The contours $F_1^2 = 0{\cdot}9$ and $F_1^2 = 0{\cdot}8$ are plotted as functions of l for $k = (5\pi/8, 7\pi/8)$. The broken line is the join of $k' = k$ to $k' = (\pi, 2\pi/3)$, the band edge, which is marked by a circle.

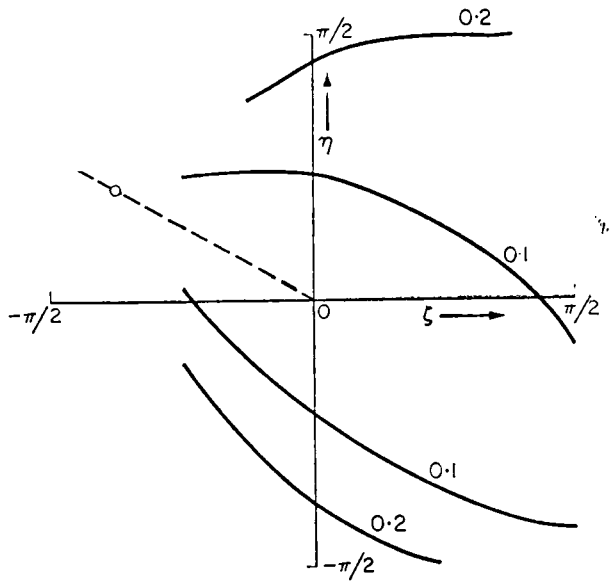


FIG. 18. Contour diagram of F_2^2 for graphite. The contours $F_2^2 = 0{\cdot}1$ and $F_2^2 = 0{\cdot}2$ are plotted as functions of l for $k = (5\pi/8, 7\pi/8)$. The broken line is the join of $k' = k$ to $k' = (\pi, 2\pi/3)$, the band edge, which is marked by a circle.

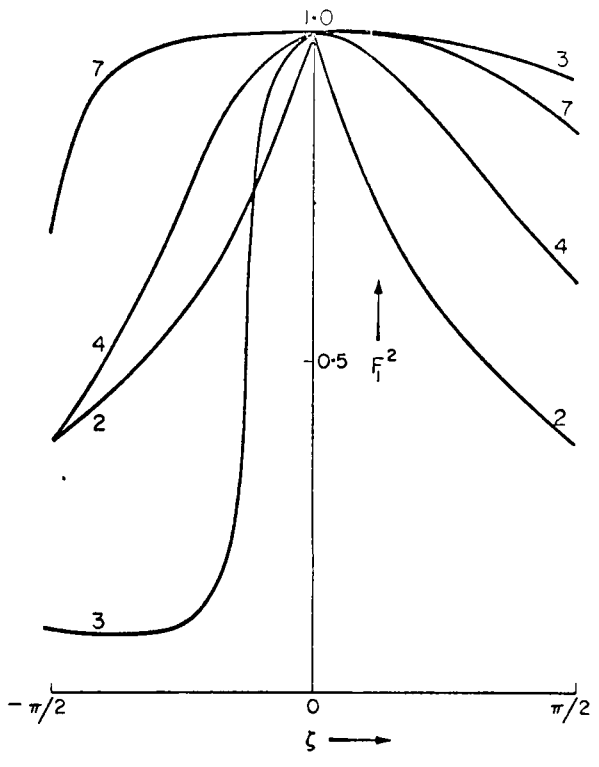


FIG. 19. F_1^2 for graphite as a function of $l = (\zeta, 0)$ for $k = (\pi, 3\pi/4)$, $(7\pi/8, 5\pi/8)$, $(7\pi/8, 7\pi/8)$, and $(\pi/2, \pi/2)$, numbered 2, 3, 4, and 7 respectively.

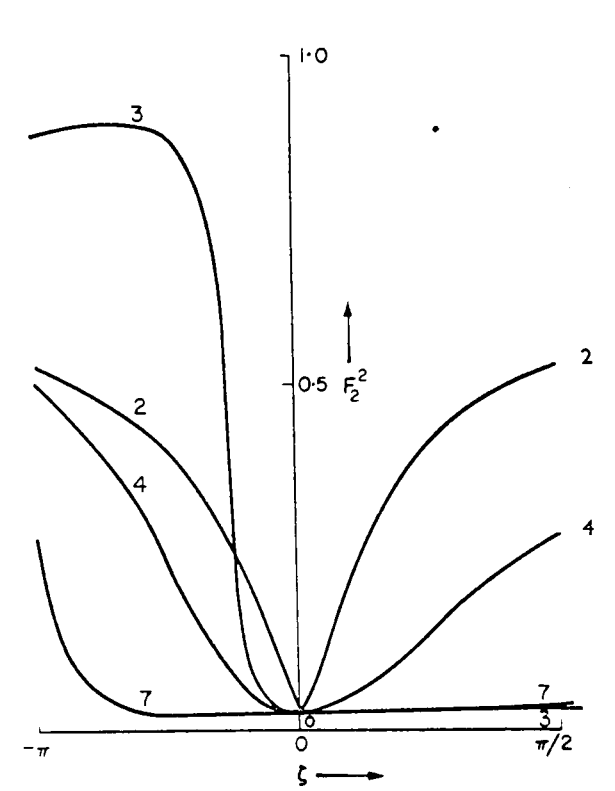


FIG. 20. F_2^2 for graphite as a function of $l = (\zeta, 0)$ for $k = (\pi, 3\pi/4)$, $(7\pi/8, 5\pi/8)$, $(7\pi/8, 7\pi/8)$ and $(\pi/2, \pi/2)$, numbered, 2, 3, 4 and 7 respectively.

Since $F_2^2 + F_4^2$ cannot exceed one, while generally F_4^2 is greater than F_1^2, the sum of F_1^2 and F_2^2 tends to be somewhat less than one, unlike the results found by Beattie and Landsberg.[56] Whether this difference is due to the error in our wave functions, and so is not significant, or whether it arises from a fundamental difference between a one-dimensional Kronig-Penney chain and a real planar crystal, is not known. It should be noted that in this heteronuclear case all the F_j^2 are continuous functions of $\underset{\sim}{k}$ and $\underset{\sim}{k}^1$, in contrast to the results found for the "homonuclear" linear chain and for graphite, which is also homonuclear.

Graphite.

The results for graphite differ markedly from those obtained for boron nitride because of the essential singularity in the wave functions for the valence and conduction bands of graphite at (π, $2\pi/3$), where on our model the two bands touch and so the staes are degenerate. It is found that on passing through this point of reciprocal space in certain directions the valence and conduction bands switch over, so that F_1^2 on one side is continuous with F_2^2 on the other. This behaviour is clearly shown in figure 22, and a theoretical explanation of it is presented in the next section. Because of this singularity, the behaviour of F_j^2 as a function of $\underset{\sim}{l}$ is very erratic if $\underset{\sim}{k}^1$ is near (π, $2\pi/3$), and the overlap functions depend strongly on the direction of $\underset{\sim}{l}$ relative to the line joining $\underset{\sim}{k}^1 = \underset{\sim}{k}$ to

this point. For this reason, it is not really practicable to draw a contour map of F_1^2 or F_2^2 for $\underset{\sim}{k} = (7\pi/8, 5\pi/8)$; however $\underset{\sim}{k} = (5\pi/8, 7\pi/8)$ is sufficiently far from the band edge for reasonable contour maps to be drawn. These are shown in figures 17 and 18, and should be contrasted with the corresponding results for boron nitride shown in figures 14 and 15. Table 7 contains the mean values of each F_j^2 for $\underset{\sim}{l} = (\pm\pi/4, 0)$ and $(0, \pm\pi/4)$ for those values of $\underset{\sim}{k}$ where the maximum deviation from the mean is less than 0.1.

For graphite, F_1^2 and F_4^2 tend to be smaller than for boron nitride unless $\underset{\sim}{k}$ is very far from the band edge; this is because $u_{\underset{\sim}{k}}$ tends to vary rapidly with $\underset{\sim}{k}$ near the point of degeneracy. This behaviour is also exhibited in figures 19 and 20, in which the variation of F_1^2 and F_2^2 along the line $\underset{\sim}{l} = (\zeta, 0)$ is plotted, and should be contrasted with the corresponding results for boron nitride shown in figure 16. Finally, in order to show quite clearly what is happening, we plot in figures 21 and 22, for planar boron nitride and graphite respectively, the values of F_1^2 and F_2^2 along the line $\underset{\sim}{k}^1 = (\pi, M'\pi/8)$ for four different values of $\underset{\sim}{k}$; the band edge lies at $M^1 = 5\frac{1}{2}$ on this line.

5.3 THEORETICAL CONSIDERATIONS AND CONCLUSIONS.

L.C.A.O. - type theory.

It is of interest to apply the methods of the usual L.C.A.O. theory for diatomic molecules to the functions $u_{\underset{\sim}{k}}$,

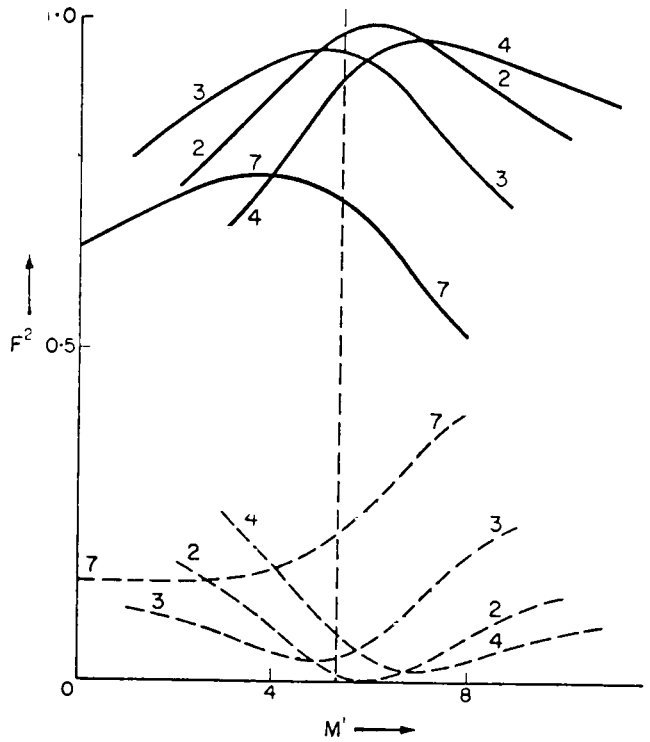


FIG. 21. F_1^2 (full line) and F_2^2 (broken line) for planar boron nitride as functions of $k' = (\pi, M'\pi/8)$ for $k = (\pi, 3\pi/4)$, $(7\pi/8, 5\pi/8)$, $(7\pi/8, 7\pi/8)$ and $(\pi/2, \pi/2)$, numbered 2, 3, 4, and 7 respectively. The vertical broken line marks the band edge.

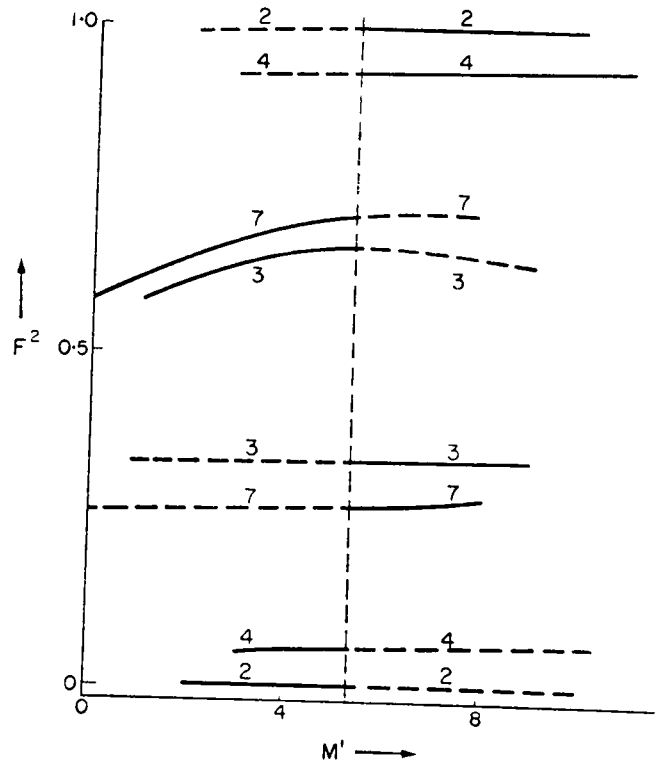


FIG. 22. F_1^2 (full line) and F_2^2 (broken line) for graphite as functions of $k' = (\pi, M'\pi/8)$ for $k = (\pi, 3\pi'4)$, $(7\pi/8, 5\pi/8)$, $(7\pi/8, 7\pi/8)$ and $(\pi/2, \pi/2)$, numbered 2, 3, 4 and 7 respectively. The vertical broken line marks the band edge.

which are the periodic parts of the molecular orbitals in the crystal. The Schrodinger equation for a molecular orbital $\psi_{\underline{k}}$ is

$$-\frac{\hbar^2}{2m}\nabla^2\psi_{\underline{k}}(\underline{r}) + V(\underline{r})\,\psi_{\underline{k}}(\underline{r}) = E_{\underline{k}}\,\psi_{\underline{k}}(\underline{r}) \qquad (5.29)$$

Consider now a crystal, the unit cell of which contains one A atom and one B atom, and let us assume that we can write

$$V(\underline{r}) = V^A(\underline{r}) + V^B(\underline{r}) \qquad (5.30)$$

where V^A and V^B are periodic potentials concentrated chiefly in the regions near the A and B atoms respectively. We assume also that $V(\underline{r})$ is a local potential, so that

$$V(\underline{r})\, e^{i\underline{k}\cdot\underline{r}}\, u_{\underline{k}}(\underline{r}) = e^{i\underline{k}\cdot\underline{r}}\, V(\underline{r})\, u_{\underline{k}}(\underline{r}) \qquad (5.31)$$

Then, since

$$\nabla\left(e^{i\underline{k}\cdot\underline{r}}\, u_{\underline{k}}(\underline{r})\right) = e^{i\underline{k}\cdot\underline{r}}\,(\nabla + i\underline{k})\, u_{\underline{k}}(\underline{r}) \qquad (5.32)$$

the substitution of equation 5.2 into equation 5.29 leads to

$$T_{\underline{k}}\, u_{\underline{k}}(\underline{r}) + V(\underline{r})\, u_{\underline{k}}(\underline{r}) = E_{\underline{k}}\, u_{\underline{k}}(\underline{r}) \qquad (5.33)$$

where $$T_{\underline{k}} = -\frac{\hbar^2}{2m}(\nabla + i\underline{k})^2 \qquad (5.34)$$

Suppose now that the functions $U^A_{\underline{k}}$ and $U^B_{\underline{k}}$ satisfy the same periodicity conditions as $u_{\underline{k}}$ and are such that

$$\left(T_{\underline{k}} + V^A(\underline{r})\right)\, U^A_{\underline{k}}(\underline{r}) = E^A_{\underline{k}}\, U^A_{\underline{k}}(\underline{r}) \qquad (5.35)$$

and $$\left(T_{\underline{k}} + V^B(r)\right)\, U^B_{\underline{k}}(\underline{r}) = E^B_{\underline{k}}\, U^B_{\underline{k}}(\underline{r}) \qquad (5.36)$$

where neither $E^A_{\underline{k}}$ nor $E^B_{\underline{k}}$ is degenerate or quasi-degenerate. Then to a first approximation we can write

$$u_{\underline{k}}(\underline{r}) = \frac{U^A_{\underline{k}}(\underline{r}) + \lambda_{\underline{k}} U^B_{\underline{k}}(\underline{r})}{(1 + \lambda_{\underline{k}} S_{\underline{k}} + \lambda^*_{\underline{k}} S^*_{\underline{k}} + \lambda_{\underline{k}} \lambda^*_{\underline{k}})^{\frac{1}{2}}} \qquad (5.37)$$

where $$S_{\underline{k}} = \int U^{A*}_{\underline{k}}(\underline{r})\, U^B_{\underline{k}}(\underline{r})\, d\underline{r} \qquad (5.38)$$

and $\lambda_{\underline{k}}$ is such as to make stationary

$$E_{\underline{k}} = \int u^*_{\underline{k}}(\underline{r}) \left(T_{\underline{k}} + V(\underline{r})\right) u_{\underline{k}}(\underline{r})\, d\underline{r} \qquad (5.39)$$

Also, let us write

$$\beta_{\underline{k}} = \tfrac{1}{2} \int \left[U^{A*}_{\underline{k}}(\underline{r}) V^A(\underline{r}) U^B_{\underline{k}}(\underline{r}) - U^{B*}_{\underline{k}}(\underline{r}) V^B(\underline{r}) U^A_{\underline{k}}(\underline{r}) \right] d\underline{r} \qquad (5.40)$$

For a homonuclear substance, in which atoms A and B are the same, $R[\lambda_{\underline{k}}] = \pm 1$. The sign of $\beta_{\underline{k}}$ determines which of these has the lower energy. However, if for some value of $\underline{k}$ (which we denote by $\underline{k}_o$) $\beta_{\underline{k}} = 0$, then $\lambda_{\underline{k}}$ is indeterminate at $\underline{k}_o$ and there are two degenerate states there. If $\beta_{\underline{k}}$ takes both positive and negative values in the neighbourhood of $\underline{k}_o$, then even if $U^A_{\underline{k}}$ and $U^B_{\underline{k}}$ are continuous functions of $\underline{k}$ in this region $U_{\underline{k}}$ (valence) will be a discontinuous function of $\underline{k}$. Thus in certain directions there will be a switch of valence and conduction band wave functions on opposite sides of $\underline{k}_o$, as we found in graphite.

For a heteronuclear substance, however, where in general $E^A_{\underline{k}}$ and $E^B_{\underline{k}}$ differ, the two different values of $\lambda_{\underline{k}}$ will everywhere correspond to states differing by a finite energy, so that no degeneracy occurs. Even if $\lambda_{\underline{k}}$ becomes infinite for some value of $\underline{k}$, $1/\lambda_{\underline{k}}$ will normally be a continuous function of $\underline{k}$ at this point. Hence, since the energies are

distinct and are continuous functions of $\lambda_{\underset{\sim}{k}}$ or $1/\lambda_{\underset{\sim}{k}}$, $u_{\underset{\sim}{k}}$ $u_{\underset{\sim}{k}}$(valence) will also be a continuous function of $\underset{\sim}{k}$ and no switching of functions is to be expected; ~~even where~~ this is the situation for planar boron nitride.

Conclusions.

The wave functions used in this work probably differ considerably from the exact ones, so that the quantitative results obtained in this chapter may not be too reliable, but the qualitative results should be reasonably accurate. Thus it seems probable that, for planar compounds in which a finite energy gap exists between the valence and conduction bands, the approximation of $F_1^2 = 1$ or $F_4^2 = 1$ should involve an error of not more than 10% provided that $|\underset{\sim}{k} - \underset{\sim}{k}'|$ is less than one quarter of the mean radius of the first Brillouin zone. A similar result was found on the one-dimensional Kronig-Penney model,[56] and so it seems quite probable that it is also true in three dimensions e.g. for non-degenerate valence and conduction bands in tetrahedral semi-conductors. It is difficult to generalise our results for F_2^2 to such crystals, since they usually have several valence and several conduction bands fairly close together in energy while the planar compounds do not. Thus, no reliable conclusions can be drawn about the range in which $\underset{\sim}{k}.\underset{\sim}{p}$ perturbation theory is likely to be valid.[55] Finally, the results for graphite and our theoretical analysis suggest that, even for small $|\underset{\sim}{k} - \underset{\sim}{k}'|$, $F_4^2 = 1$ may not be a very good approximation for degenerate bands near the point of degeneracy, e.g. for the valence bands of germanium or indium antimonide near the centre of the Brillouin zone.

REFERENCES.

1. J. LENNARD-JONES, Proc. Roy. Soc. A.198, 1 & 14 (1949).
2. C.C.J.ROOTHAAN, Rev. Mod. Phys. 23, 69 (1951).
3. G.F.KOSTER, Solid State Physics 5, 173 (1957).
4. L.B.REDEI, D. Phil. Thesis, Univ. of Oxford, 1961.
5. S. FLODMARK, Arkiv fur Physik 14, 513 (1959).
6. F.C.VON DER LAGE and H.A.BETHE, Phys. Rev. 71, 612 (1947).
7. C. HERRING, Phys. Rev. 57, 1169 (1940).
8. J.CALLAWAY, Phys. Rev. 97, 933 (1955).
9. J.C.PHILLIPS and L.KLEINMAN, Phys. Rev. 116, 287 (1959).
10. M.H.COHEN and V.HEINE, Phys. Rev. 122, 1821 (1961).
11. E. BROWN, Phys. Rev. 126, 421 (1962).
12. E. BROWN and J.A.KRUMHANSL, Phys. Rev. 109, 30 (1958).
13. J.C.SLATER, Phys. Rev. 92, 603 (1953).
14. J.H.WOOD, Phys. Rev. 126, 517 (1962).
15. C.A.COULSON and H.C.LONGUET-HIGGINS, Proc.Roy.Soc.A.191,39 (1947).
16. C.A.COULSON and R.TAYLOR, Proc.Phys.Soc. A.65, 815 & 834 (1952).
17. J.C.PHILLIPS, Proc.Internat. Conf. on Semiconductor Physics, Prague 1960, p.41.
18. D. STOCKER, Proc.Roy. Soc. A.270, 397 (1962).
19. N.V.COHAN, D.PUGH, and R.H.TREDGOLD, Proc.Phys.Soc. 82, 65 (1963).
20. C.A.COULSON, L.B.REDEI, and D.STOCKER, Proc.Roy.Soc. A.270, 357 (1962).
21. J. C. SLATER, Phys. Rev. 36, 57 (1930).
22. F. HERMAN, Phys. Rev. 93, 1214 (1954).
23. G. F. KOSTER, Phys. Rev. 89, 67 (1953).
24. J.C.SLATER and G.F.KOSTER, Phys. Rev. 94, 1498 (1954).
25. R.H.PARMENTER, Phys. Rev. 86, 552 (1952).

26. J.LENNARD-JONES and J.A.POPLE, Proc.Roy.Soc. A.202, 166 (1950)

27. G.G.HALL, Proc. Roy. Soc. A. 202, 336 (1950).

28. L.PAULING, Proc. Roy. Soc. A. 196, 343 (1949).

29. C.A.COULSON, Proc.Internat.Conf. of Theoretical Physics, Kyoto and Tokyo, 1953, p.629.

30. P.O.LOWDIN, J. Chem,Phys. 18, 365 (1950).

31. S. F. BOYS, Proc. Roy. Soc. A. 200, 542 (1950).

32. G.G.HALL, Phil. Mag. 3, 429 (1958).

33. F.W.QUELLE Jr., Proc. Internat.Conf. on Semiconductor Physics, Prague, 1960, p.48.

34. J. C. SLATER, Phys. Rev. 81, 385 (1951).

35. L.KLEINMAN and J.C.PHILLIPS, Phys.Rev. 116, 880 (1959).

36. L.KLEINMAN and J.C.PHILLIPS, Phys.Rev. 118, 1153 (1960).

37. F.HERMAN and S.SKILLMAN, Proc.Internat.Conf. on Semiconductor Physics, Prague, 1960, p.20.

38. T.O.WOODRUFF, Phys.Rev. 103, 1159 (1956).

39. M.J.D.POWELL, Computer Journal 5, 147 (1962).

40. L.B.BOUCKAERT, R.SMOLUCHOWSKI, and E.WIGNER, Phys.Rev. 50,58 (1936)

41. R.S.MULLIKEN, C.A.RIEKE, and W.G.BROWN, J.Am.Chem.Soc.63, 41 (1941)

42. F.BASSANI and Y.YOSHIMINE, Phys.Rev. 130, 20 (1963).

43. E.O.KANE, J.Phys. Chem. Sol. 1, 249 (1957).

44. J. CALLAWAY, Journal of Electronics 2, 330 (1957)

45. E.MOOSER and W.B.PEARSON, Acta Cryst. 12, 1015 (1959).

46. B. LAX, Rev. Mod. Phys. 30, 122 (1958).

47. D. LONG, J. App. Phys. 33, 1682 (1962).

48. G.A.WOLFF and J.D.BRODER, Acta Cryst. 12, 313 (1959).

49. G.PICUS, E.BURSTEIN, B.W.HENVIS, and M.HAAS, J.Phys.Chem.Sol. 8, 282 (1959).

50. Y.A.UGAI and E.P.DOMASHEVSKAYA, Dokl.Akad.Nauk. SSR. 156,430(1964)

51. N.N.SIROTA, N.M.OLEKHNOVITCH and A.U.SHELEG, Dokl.Akad.Nauk.SSR. 132, 160 (1960); 135, 1176 (1960); 136, 660 & 879 (1961); 138, 1 (1961); 143, 156, 370 & 398 (1962).

52. Dr.J.C.PHILLIPS, Private communication.

53. A.R.BEATTIE and P.T.LANDSBERG, Proc.Roy.Soc.A. 249, 16 (1959).

54. O.HOCHENBERG and W.BRAUER, Advances in Electronics 11,413 (1959)

55. E.ANTONCIK and P.T.LANDSBERG, Proc.Phys.Soc. 82, 337 (1963).

56. A.R.BEATTIE and P.T.LANDSBERG, Proc.Roy.Soc.A. 258, 486 (1960).

APPENDIX I.

THE SCHMIDT ORTHONORMALISATION PROCEDURE.

The General Problem.

Given a set of functions $V_1, V_2, \ldots, V_n$, we wish to find an orthonormal set of functions

$$W_r = \sum_{s=1}^{n} t_{rs} V_s$$

such that the set $\{W_r\}$ spans the same Hilbert space as the functions V_s.

Theorem.

If the functions V_s are linearly independent, then if we write

$$W_1 = V_1/N_1 \quad \text{where } N_1^2 = (V_1 | V_1) \tag{A 2}$$

and define recursively

$$W_r = 1/N_r \left[V_r - \sum_{s=1}^{r-1} (W_s | V_r) W_s \right] \tag{A 3}$$

where $N_r = (V_r | W_r)$ (A 4)

then the functions $\{W_r\}$ form an orthonormal set.

Proof.

By induction. Equation A 3 holds by definition for $r = 1$. Suppose now that it holds for $1 \leq r \leq m-1$. Then if $t < m$

$$(W_t | W_m) = 1/N_m \left[(W_t | V_m) - \sum_{s=1}^{m-1} (W_t | W_s)(W_s | V_m) \right] = 0$$

since by the induction hypothesis $(W_t | W_s) = \delta_{st}$. If $t = m$, we have that

$$(W_m | W_m) = 1/N_m \left[(W_m | V_m) - \sum_{s=1}^{m-1} (W_m | W_s)(W_s | V_m) \right]$$

$$= 1/N_m \left[(W_m | V_m) \right] = 1$$

Hence, by induction, the theorem holds for all m. Q.E.D.

Corollary.

$$N_r^2 = (N_r W_r | V_r) = (V_r | V_r) - \sum_{s=1}^{r-1} (W_s | V_r)^2$$

If some function V_s is linearly dependent on its predecessors, then N_s^2 as given by equation A 5 will be less than or equal to zero. In that case, we omit W_s from our final set of functions, and treat $1/N_s$ and W_s as zero in defining the remaining W_r.

Procedure

First, calculate $S_{lm} = (V_l \mid V_m)$. Then form for each value of r in turn

$$N_r^2 = \begin{cases} S_{11} & \text{if } r=1 \\ S_{rr} - \sum_{s=1}^{r-1} \left[\sum_{l=1}^{s} t_{sl} S_{lr} \right]^2 \end{cases} \qquad (A\,6)$$

and if this quantity is positive, otherwise put $1/N_r = 0$, and in either case form

$$t_{rl} = \begin{cases} -1/N_r \sum_{s=1}^{r-1} \sum_{m=1}^{s} S_{mr}\, t_{sm}\, t_{sl} & \text{if } l < r \\ 1/N_r & \text{if } l = r \\ 0 & \text{if } l > r \end{cases} \qquad (A\,7)$$

Then the functions $W_r = \sum_{s=1}^{n} t_{rs} V_s$ (A 1)

are the required set of functions.

Applications.

Suppose that the functions $u_1, u_2, \ldots, u_m$, form an orthonormal set, and that it is desired to form from the functions $\{v_s\}$ a set of normalised functions $\{w_s\}$ which are orthogonal to the functions $\{u_r\}$ and are also mutually orthogonal. Then if we write

$$V_r = \begin{cases} u_r & \text{if } r \le m \\ v_{r-m} & \text{if } r > m \end{cases} \qquad (A\,8)$$

we can apply the above analysis, and the functions W_s with $s > m$ constitute the required set. We find that

$$t_{rs} = \delta_{rs} \quad \text{if } r \le m \qquad (A\,9)$$

so that the matrix (t_{rs}) has the form

$$T = \left(\begin{array}{c|c} 1 & 0 \\ \hline A & B \end{array} \right) \qquad (A\,10)$$

which was a result used in section 3 of chapter 1.

APPENDIX II.

EFFECTIVE MASSES.

In order to show how the effective masses for the valence bands of a tetrahedral semiconductor depend on the parameters F_j, and so on the E_j, we write down expressions for certain components of the inverse effective mass tensor for that valence band which is doubly degenerate along ΓX and ΓL ; we are using the standard notation, in which

$$\Gamma = (0,0,0)\ ,\quad X = \tfrac{1}{L}(\pi,0,0) \text{ and } L = \tfrac{1}{L}\left(\tfrac{\pi}{2},\tfrac{\pi}{2},\tfrac{\pi}{2}\right) \qquad (A\ 11)$$

Along the line ΓX for this band, we find from equation 4.2 that

$$\frac{\partial^2 E_1}{\partial l^2} = -8(F_4+F_5)\cos 2l - 16F_{10}\cos 4l + 8(F_3+2F_8)\cos 2l \qquad (A\ 12)$$

while for the component in this direction

$$\frac{1}{m^*} = \frac{L^2}{4\hbar^2}\frac{\partial^2 E}{\partial l^2} \qquad (A\ 13)$$

Along the line ΓL , we have from equation 4.3 that

$$\frac{\partial^2 E_1}{\partial l^2} = -4(F_4+2F_5+3F_{10})\cos 2l + 4(F_3+2F_6+2F_8)\cos 2l \qquad (A\ 14)$$

while for the component in this direction

$$\frac{1}{m^*} = \frac{L^2}{\hbar^2}\frac{\partial^2 E}{\partial l^2} \qquad (A\ 15)$$

Our most reliable results concerned the diamond-type compounds, in which the values of F_3, F_4, F_5, F_6, F_8, and F_{10} were typically -0.025, 0.018, -0.016, -0.004, 0.602, and 0.0002 respectively; all our results are expressed in atomic units. It follows that in A 14 the terms involving F_6, F_8, and perhaps F_{10} will play an important role in determining the effective mass because of the partial cancellation of the other terms; similarly, these parameters will be of importance in A 12, especially near $l = \pi/4$.

These results could obviously be generalised to apply to other components and to other bands. The point we wish to emphasise is that the parameters associated with bonds farther apart than third-neighbours parallel, which were ignored by Stocker[18] in most of his calculations, can play a much more significant role in determining the effective masses than they do in determining the energy levels. Phillips' conclusion about the inability of

of equivalent orbital methods to give correct effective masses[17] was based mainly on the wrok of Hall[32], who aleo neglected F_5 and in any case estimated his parameters by fitting them to the energy bands calculated by other methods instead of using an ab initio calculation to find them. In our method, a much larger set of basis functions would be needed in order to determine the smaller parameters accurately, since with only a few basis functions the equivalent orbital wave function will not be flexible enough to minimise the overlaps between more distant bonds as well as those between nearer neighbours.

APPENDIX III.

THE COMPUTER PROGRAMMES.

In this appendix, we present a brief description of the computer programmes we used and the mathematical techniques that they embodied, ~~followed by the actual programmes, in Mercury autocode, used for the heteromuclear compounds~~. For convenience with regard to re-starts in the event of machine failure, the master programme was divided into seven sections, numbered M 1 to M 7, which were presented to the computer in numerical order and were preceded by a data tape with a title number between M 100 and M 199. The tapes and chapters are described in the order in which they are used in the master programme, and the actual programme is ~~given~~ then analysed in the order in which it was presented to the computer.

Tape M 100.

This tape designates in the primed special variables the various auxiliary store locations to be used in the subsequent parts of the programme, and reads in the following data in atomic units:

A = L/4, one quarter of the length of the side of our unit cube. This quantity was obtained from the determinative tables of crystal ~~data~~ data published by the American Crystallographic Association; in cases where more than one value of the lattice constant was given, the smallest value mentioned was used, but no allowance was made for thermal expansion although most measurements were made at 20°C while our calculations relate to 0°K.

C = $3/R^2$, the exponent in our Gaussian functions e^{-Cr^2}.

E = the number of core functions on atom A.

F = the number of core functions on atom B.

G = the total number of data elements to be read directly into the auxiliary store.

H = the number of different values of R wanted in Chapter 2 of M 1, multiplied by -1 if the crystal was homonuclear.

V = the number of data elements associated with the wave functions only.

This data was followed by:

1) the core functions of atom A, in groups of three, preceded by the number of elements per wave function in this group. The functions were listed in the order 1s,2s,2p; 3s,3d,3p; 4s,4d,4p; so that the first member of each group was always an s-function and only the last member was a p-function.

2) the core functions of atom B, similarly arranged.

3) the energies of the core orbitalsm on atom A, as calculated by Dr. Mayers for the free atom, followed by those for atom B.

4) the nuclear charges Z_A and Z_B on atoms A and B, followed by the number of valence electrons, n_A and n_B, on each; we always chose A so that n_A was not larger than n_B.

5) the approximate average energy of a valence electron in the solid, the difference between the average energies of the valence electrons on neutral atoms A and B, and an arbitrary parameter.

6) the charge density due to the valence electrons for a neutral atom A, followed by those for a neutral atom B.

Tape M 1.

Chapter 3 of this forms $(\rho)_j$, $(\frac{1}{r}\rho)_j$, and $(\rho^{\frac{1}{3}})_j$, for the neutral atoms. As in all integrations involving the core functions, the integrations were carried out using Simpson's rule over the intervals r = 0(0.00125)0.025 (0.0025)0.05(0.005)0.1(0.01)0.2(0.02)0.4(0.04)0.8(0.08)1.6(0.16)3.2(0.32)6.4... as these were the intervals at which Dr. Mayers' wave functions were tabulated.

Chapter 1 forms the integrals $\int_{-\infty}^{\infty} (x-a)^m e^{-c(x-a)^2} e^{ikx} x^n e^{-cx^2} dx$

and $\int_{-\infty}^{\infty} x^m e^{-cx^2} \left[-\frac{1}{2} \frac{d^2}{dx^2} (x^n e^{-cx^2}) \right] dx$

which are the basic integrals needed for the calculation of integrals of Gaussianm functions in three dimensions. The general recurrence relation for an integral of this sort, which can be obtained by integration by parts or by differentiation under the integral sign, is that if

$$I_n = \int_{-\infty}^{\infty} x^n e^{-cx^2} e^{-c(x-a)^2} e^{ikx} dx \qquad (A\ 16)$$

then

$$I_{n+1} = \left(\frac{a}{2} + \frac{ik}{4c}\right) I_n + \frac{n}{4c} I_{n-1} \qquad (A\ 17)$$

while $I_0 = \sqrt{\frac{\pi}{2c}} e^{-\frac{ca^2}{2}} e^{\frac{ika}{2}} e^{-\frac{k^2}{8c}}$

We note that if $C = 3/R^2$ and $k^2 = n\left(\frac{\pi\sqrt{3}}{2R}\right)^2$, then $|I_0| = \sqrt{\frac{\pi}{2c}}\, e^{-ca^2/2}\, e^{-n\pi^2/32}$ which, as remarked previously, decreases rapidly with increasing n once n is greater than three.

Chapter 2 evaluates integrals of the form

$$\int_{r=0}^{\infty}\int_{\mu=-1}^{1} r^l \mu^m \rho_t(r)\, e^{-c(r^2+R^2-2Rr\mu)}\, d\mu\, dr$$

for a prescribed set of values of l and m, with $R^2 = (3+16s)A^2$ where s is an integer. The integration over μ was performed using a power-series expansion derived from Chebyshev polynomials in the region where μ was too small for iterative methods to be used; the radial integrals were performed using Simpson's rule. These integrals are needed in order to find the overlap between Gaussian functions centred about the midpoint of our reference bond and the core functions of the various atoms in the crystal.

Tape M 2.

We write $$S^{i\tilde{n}} = \begin{pmatrix} P^{i\tilde{n}} & Q^{i\tilde{n}} \\ \tilde{Q}^{i\tilde{n}} & R^{i\tilde{n}} \end{pmatrix} \qquad (A\ 19)$$

Chapters 1 and 2 form $Q^{i\underline{n}}$, and subsequently $AQ^{i\underline{n}}\tilde{B} + BQ^{\tilde{i}\underline{n}}\tilde{A}$, using the results of M 1 Chapter 2.

Chapter 3 forms $BR^{i\underline{n}}\tilde{B}$, using the results of M 1 Chapter 1, and prints out $TS^{i\underline{n}}\tilde{T}$. The basis core functions are stored in the form n, a_0, a_1, a_2,.... which represents the function $[a_0(x^n+y^n+z^n) + a_1(x^{n-1}y + x^{n-1}z + y^{n-1}z + y^{n-1}x + z^{n-1}x + z^{n-1}y) + \ldots]e^{-cr^2}$

Chapter 4 forms the matrix T using the Schmidt orthonormalisation procedure, and also forms $AP^{i\underline{n}}\tilde{A}$ for those bonds having an atom in common with the reference bond; core orbitals on different ~~orbitals~~ atoms were assumed to be orthogonal.

Tape M 3.

Chapter 1 forms and prints out $BR_{\underline{K}}\tilde{B}$ for the various $\underline{K}$ we are using, and also $B(-\tfrac{1}{2}\nabla^2)\tilde{B}$

Tape M 4.

Chapters 1 and 2 generate and store the integrals needed to form $P_{\underline{K}}$, i.e. the matrix elements of $\sum_{i=0}^{3} e^{-i\beta_i \underline{K}\cdot\underline{r}}$ between the core functions.

Chapters 3 and 0 form ρ_n^{core} for atoms A and B, and also forms the matrices $BR_{\underline{K}}\tilde{B} - AP_{\underline{K}}\tilde{A}$ which can be used as approximations to $\beta TR_{\underline{K}}\tilde{T}$.

Tape M 5.

Chapter 1 calculates integrals of the form

$$\iiint P_t(r)\, Y_{lm}(\theta,\phi)\, e^{i\underline{K}.\underline{r}}\, v_m(\underline{r}-\underline{a})\, d\underline{r}$$

where $v_m(\underline{r}-\underline{a})$ is a Gaussian function centred about the midpoint of the reference bond and $P_t(r)\, Y_{lm}(\theta,\phi)$ is a core wave function for an atom at one end of the bond. Integration over ϕ leads to a zero-order Bessel function of r and θ, which can readily be calculated from a power-series expansion, but the integrations over θ and r have to be done numerically. The 12-point Gaussian formula was used for the integration over θ, and Simpson's rule for the integration over r. As one would expect, this double numerical integration is a comparatively slow process, and for this reason M 5 was omitted from the production runs.

Chapter 2 uses the results of Chapter 1 to calculate $AQ_{\underline{K}}\check{B} + BQ_{\underline{K}}\tilde{A}$, and combines these with the previous results to form and print out $TR_{\underline{K}}\tilde{T}$.

Tape M 6.

Chapter 3 forms the matrix of the Hamiltonian appropriate to a crystal containing atoms stripped of their valence electrons and with the zero of energy shifted by W; this matrix is stored for use in Chapter 2, and the various parameters needed in the calculation of a self-consistent H and Y are also calculated and stored.

Chapter 1 uses Powell's convergence routine[39], with the modifications mentioned in section 4 of chapter 3.

Chapter 2 contains a routine for calculating Y and its derivatives for an unnormalised approximation to a_1.

Tape M 7.

Chapter 1 calculates and prints out a number of overlap and other integrals associated with the equivalent orbitals found by M 6, and also calculates the modulus of $H|a_1\rangle - \varepsilon_0'|a_1\rangle - \sum_{\underline{c}} \varepsilon_{\underline{c}}'(\underline{c})\, S^{\underline{c}}|a_1\rangle$

Chapter 2 calculates the converged values of the parameters E_j' using an iterative procedure in conjunction with equation 4.6.

Chapter 0 calculates the energies for the valence bands in the $\langle 100\rangle$ and $\Sigma\langle 111\rangle$ directions, using equations 4.2 and 4.3.

Rather than present the full computer programmes in Mercury Autocode, which would occupy some eighty pages and in any case would not be very efficient for the larger computers on which any future calculations ought to be carried out, we prefer to give here just a flow analysis of these programmes, which may perhaps be of more general use.

M100Z, Chapter 0.

This just puts the basic data into the auxiliary store.

M 1, Chapter 1.

a) Evaluate and store $e^{ikL/2}\int_{-\infty}^{\infty} x^n e^{-2cx^2} e^{ikx}\,dx$, $0 \le n \le 10$, $k = (0,1,2,3,4)\pi/L$

b) Form and store $\int_{-\infty}^{\infty} x^n e^{-cx^2}\left[-\tfrac{1}{2}\frac{d^2}{dx^2}(x^m e^{-cx^2})\right]dx$, $0 \le m,n \le 5$

c) Form $\int_{-\infty}^{\infty} x^m e^{-cx^2} e^{-c(x-a)^2}\,dx$, $0 \le m \le 10$, $a = (0,1,2,3,4)L/2$

d) Using c), form and store $\int_{-\infty}^{\infty} (x-a)^m e^{-c(x-a)^2} x^n e^{-cx^2}\,dx$, $0 \le m,n \le 5$

M 1 Chapter 2.

a) Set $R = (3+16n)A^2$, $n = 0, 1, 2, \ldots |H|$, in turn.

b) Choose sets of A and B core functions in turn; each set contains three functions.

c) Form $\int_{-1}^{1} \mu^n e^{2CvR\mu}\,d\mu$

d) Multiply by $2\pi r^s e^{-c(r^2+R^2)}$, by the core functions, and by r if the core functions are p-type.

e) Integrate using Simpson's rule, and store the results temporarily. Return to a) until all the values of n are exhausted, then return to Chapter 0.

M 1, Chapter 3.

a) Set up in the fast store $4\pi r^2\rho_A$, and then $4\pi r^2\rho_B$

b) Form the integrands $4\pi r^2\rho \sin kr / kr$, $(4\pi)^{-\frac{2}{3}}\rho^{\frac{1}{3}} \sin kr$, and $(4\pi)^{-1} \sin kr / \rho$.

c) Integrate using Simpson's rule.

M 1, Chapter 0.

a) Add $4\pi r^2\rho_{\text{valence}}$ to $4\pi r^2\rho_{\text{core}}$.

b) Do chapter 3.

c) Read in the coefficients for the power series used in c) of chapter 2.

d) Do chapter 1.

e) Do chapter 2.

M 2, Chapter 1.

a) Find the position $\underline{R}$ of the A atom at the end of a given bond with respect to the mid-point of the reference bond.

b) Rotate the axes to put the new z-axis in the same direction as $\underline{R}$, and express the expansion functions and core orbitals in terms of the new coordinates.

c) Go to chapter 2.

d) Do the above procedure for the B atom.

e) Allow for a factor of -1 in the p-orbitals on the B atom.

f) Form the matrix $Q^{i,\underline{n}}$.

g) If $(i,\underline{n}) = (0,\underline{0})$, return to chapter 0; otherwise, form $AQ\tilde{B} + B\tilde{Q}\tilde{A}$, and return to a) for the next bond.

M 2, Chapter 2

a) For each core orbital in turn, use the results of M 1, chapter 2, to form

$$\int \rho_l(r)\, Y_{lm}(\theta,\phi)\, e^{-c(R^2+r^2-2Rr\cos\theta)}\, v_n(\underline{r}-\underline{R})\, d\underline{r}.$$

with slightly different routines for $l = 0$, 1, and 2.

M 2, Chapter 3

a) Select a bond, and choose from the results of M 1, chapter 1 the sets of integrals appropriate to the overlap between this bond and the reference bond, and make allowance for whether the bond has $i = 0$ or 1.

b) Form $\int v_l v_m d\underline{r}$ for each l and m with $m \leq l$, and hence form the upper half of $R^{i,\underline{n}}$.

c) If $S^{i,\underline{n}}$ is not symmetric, repeat for the inverse bond and form half the sum of the two sets of integrals for use as a symmetric $R^{i,\underline{n}}$.

d) If $(i,\underline{n}) = (0,\underline{0})$, return to chapter 0; otherwise, form $BR\tilde{B}$, add to the results of chapters 1 and 4 to form $TR\tilde{T}$, and store and print this matrix in triangular form.

e) Go back to a), and consider the next bond.

M 2, Chapter 4

a) Carry out the Schmidt orthonormalisation procedure described in Appendix I, and then return to chapter 0.

b) For first-neighbour bonds, form $AP\tilde{A}$ and add it to $AQ\tilde{B} + B\tilde{Q}\tilde{A}$, and then return to chapter 0.

M 2, Chapter 0.

a) Read in the specifications of the bonds and of the expansion functions.

b) Go to chapter 3 for the reference bond, i.e. the bond with $(i,\underline{n}) \neq (0,\underline{0})$.

c) Go to chapter 1 for the reference bond.

d) Go to chapter 4, part a), to form the matrix T, and print it out.

e) Go to chapter 1 for the remaining bonds.

f) Go to chapter 4, part b).

g) Go to chapter 3 for the remaining bonds.

M 3, Chapter 1

a) Use the routine of M 2, chapter 3 b, to form $R_{\underline{K}}$ for each $\underline{K}$ in turn, and also to form the matrix of $-\frac{1}{2}\nabla^2$.

b) Form $BR_{\underline{K}}\tilde{B}$ and $B(-\frac{1}{2}\nabla^2)\tilde{B}$, store them, and print them out.

M 3, Chapter 0

Carry out chapter 1.

M 4, Chapter 1

a) For the core functions of atom A first, and then for those of atom B, find the number and type of the non-zero elements of $Z_{\underline{K}} = \sum_{i=0}^{3} e^{-i\beta_i \underline{K}\cdot\underline{r}}$ between the core functions. $Z_{\underline{K}} = 0$ automatically between an s and a p function, and between an s and a d function. The routine is complicated by the fact that different sets of core orbitals may contain different numbers of elements.

b) Go to chapter 2 for the s-s and p-p integrals, and store the results temporarily.

c) If there are any d core functions, go to chapter 2 for the d-d and p-d integrals, and store the results temporarily.

d) After completing the B orbitals, return to chapter 0.

M 4, Chapter 2

a) Form the integrand as a function of r; for those integrals involving d-functions we use the methods of M 1, chapter 2 c, writing ikr in place of -2CrR.

b) Integrate using Simpson's rule, and then return to chapter 1.

M 4, Chapter 3

a) For each $\underline{K}$ in turn, use the results of chapter 1 to form $P_{\underline{K}}$, making the correct allowances for phase factors.

b) Form $AP_{\underline{K}}\tilde{A}$ and $BR_{\underline{K}}\tilde{B} - AP_{\underline{K}}\tilde{A}$.

M 4, Chapter 0

a) Do chapter 1.

b) Using the methods of M 1, chapter 3, calculate the Fourier components of the core charge density.

c) Do chapter 3.

M 5, Chapter 1

a) To form the integrand for radial integration, form the Bessel function J_0 of the appropriate function of r and μ, where $\mu = \cos\theta$, and integrate over μ by the 12-point Gaussian method.

b) Repeat this for ten values of r, and store the results temporarily.

c) Take the core functions, in sets of 3, and using Simpson's rule plus the stored results integrate over r in sections, taking all intermediate results from and returning them to the auxiliary store.

d) Carry out the above for all the necessary sets of values of r, then return to chapter 2.

M 3, Chapter 2

a) For a given $\underline{K}$, set up the parameters needed in calculating the matrix elements of $e^{-i\underline{K}.\underline{r}}$ between core orbitals and our expansion functions.

b) Go to chapter 1.

c) Allow for the odd ~~xxxxxxxxxxxxxxxxxxxxxxxxxx~~ expansion functions in the integrals involving the core orbitals on the B atom, and also multiply these integrals by the correct phase factor, since the origin is at an A atom.

d) Add the results to that part of $Q_{\underline{K}}$, if any, which has already been formed.

e) Return to chapter 0.

M 3, Chapter 0

a) Set up all necessary parameters in terms of the number and type of core functions.

b) For each set of β_ν $\underline{K}$ in turn, choose those $\underline{K}$ for which $Q_{\underline{K}}$ has to be found, and then go to chapter 2.

e) When Q_K is complete, form $AQ_K\tilde{B} + BQ_K\tilde{A}$, add it to $2BR_K\tilde{B}$, and subtract $BR_K\tilde{B} - AP_K\tilde{A}$ to form $TR_K\tilde{T}$, which is then printed out and stored.

M 6, Chapter 1

This chapter contains the convergence routine described in chapter 3 of the thesis. It uses chapter 2 to form Y and its derivatives.

M 6, Chapter 2

a) Given a vector $\underline{x}$, such that in the notation of Appendix I $\sum_{i=m+1}^{m+l} x_i W_i$ is an unnormalised approximation to the equivalent orbital, form the valence charge densities and their derivatives with respect to $\underline{x}$, using the stored matrices $TR_K\tilde{T}$ and a fast routine for multiplying vectors by matrices stored in triangular form.

b) Form the matrix of the self-consistent Hamiltonian H.

c) Form Y and its derivatives with respect to $\underline{x}$, using the stored matrices $TS^{i,n}\tilde{T}$.

M 6, Chapter 3

a) Using the results of M 1, chapter 3, and M 4, chapter 0, form the Fourier components of the neutral-atom potential, and also calculate the factors by which $\underline{x}'TR_K\tilde{T}\underline{x} / \underline{x}'\underline{x}$ has to be multiplied to give the Fourier components of the nominal valence electron potential.

b) Form and store the matrix of the bare core Hamiltonian, which is the Hamiltonian appropriate to a crystal built up of neutral atoms minus the nominally self-consistent potential which would arise from the valence charge distribution on these atoms.

M 6, Chapter 0

a) Make a suitable adjustment to the zero of energy.

b) Do chapter 3.

c) Set up the parameters for the convergence routine and the initial approximation viz. a_1 is proportional to e^{-Cr^2}.

d) Do chapter 1, i.e. obtain a value of $\underline{x}$ which minimises Y.

e) Print out $\underline{x}$ and some related quantities.

M 7, Chapter 1

a) Uses a routine similar to that of M 6, Chapter 2, to form the parameters $\varepsilon_i(\underline{n})$ and $s_i(\underline{n})$, and also the valence charge density, for the equivalent orbital obtained in M 6, and prints them out.

b) Do chapter 2.

c) Form and print out the square of the modulus of

$$H|a_1\rangle - \varepsilon_0' |a_1\rangle - \sum_{i,\underline{\alpha}} \varepsilon_i'(\underline{\alpha})\, s^{i\underline{\alpha}} |a_1\rangle .$$

M 7, Chapter 2

Calculate the parameters $\varepsilon_i'(\underline{n})$ using equation 4.6 and an iterative procedure; we never needed more than 3 or 4 iterations to obtain convergence.

M 7, Chapter 0

a) Normalise $\underline{k}$.

b) Do chapter 1.

c) Calculate the energies for all the valence bands at 11 points, at intervals of $\pi/20$, along the directions (21, 0, 0)/L and (1, 1, 1)/L in reciprocal space, using equations 4.2 and 4.3, and print them out.