Model of Electron Correlation in Solids

Everett G. Larson and Walter R. Thorson

Articles you may be interested in

Optical probes of electron correlations in solids
AIP Conf. Proc. 918, 203 (2007); 10.1063/1.2751992

The cumulant expansion approach for strongly correlated electron models
AIP Conf. Proc. 695, 268 (2003); 10.1063/1.1639595

Modeling localized electron pair correlation energies
J. Chem. Phys. 103, 5662 (1995); 10.1063/1.470548

Electron Correlations, Magnetic Ordering, and Mott Insulation in Solids
J. Math. Phys. 9, 349 (1968); 10.1063/1.1664587

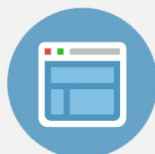Electron Correlation Models for Optical Activity
J. Chem. Phys. 48, 1127 (1968); 10.1063/1.1668773

Model of Electron Correlation in Solids

Everett G. Larson

Department of Physics, Brigham Young University, Provo, Utah

AND

Walter R. Thorson

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received 3 January 1966)

The usual Hartree–Fock model (energy-band theory) does not always give an adequate description of electronic structure in a solid, because it ignores the effects of electron correlation. It was shown first by Wigner that such a situation always develops in an electron "gas" at sufficiently low density; a solid structure described by "resonance" of Heitler–London pair bonds between electrons localized on neighboring atoms is then a good model of the system.

The transition from a Bloch-type state to such a highly correlated state as a function of electron density (lattice parameter) is a problem of considerable interest for the theory of solids, particularly those with tight binding and high electron correlation. This work is an extension of the "alternant molecular orbital" (AMO) model for molecules to a simple infinite lattice structure (one atom per lattice point), for which the lattice itself is composed of two interpenetrating and equivalent sublattices; a bcc lattice is an example. Electron correlation is treated by a variable parameter which controls alternation of electron density for a given spin between the sublattices. The ground state of this model dissociates into neutral atoms as the lattice parameter $a \to \infty$; the energy-band model ground state does not. des Cloiseaux seems to have been the first to consider a wavefunction of the AMO form for a solid; in his work, however, no real calculation is carried out, the second quantization formalism being employed and some drastic approximations made to obtain a semiquantitative description of electron behavior. In this work an explicit energy expression is obtained which is practical for exact calculations; the energy expression is also a variational form. This is important because in our opinion the model may not show all the properties ascribed to it by des Cloiseaux, and accurate calculation can establish this. This work also differs in some respects from the method of "different bands for different spins" of Calais. While both are extensions of the AMO method to infinite lattices, there are certain incorrect assumptions in Calais' treatment which lead to errors in the case of two- and three-dimensional lattices.

An approach differing from those of des Cloiseaux and Calais is used here, employing a transformed set of basis functions localized in $\mathbf{r}$ space, the localized alternant orbitals (LAO's). The LAO description explicitly shows aspects of the model which are not obvious in the $\mathbf{k}$-space basis set, particularly suggesting the relation to a "resonating" Heitler–London model.

In addition, the LAO basis set makes it easy to obtain practical energy expressions valid to higher order in N^{-1} ($2N$ electrons), for *spin eigenstates* projected from the AMO single-determinant wavefunction, for spin $s \ll N^{\frac{1}{2}}$ (and for the ferromagnetically ordered state, $s = N$). des Cloiseaux and Calais considered only the single determinant. The density of states of spin s in the vicinity of the $^1\Gamma_1$ state varies strongly with the degree of electron correlation. In this model either the $^1\Gamma_1$, or the ferromagnetic state ($s = N$) lies lowest; it is also conjectured though not definitely established that the $^1\Gamma_1$ state is always in fact the ground state, with the state $s = N$ separated from it by terms of order 1 in the energy per particle. All the energy coefficients determining these splittings are obtained as interactions of a single site with its local environment.

The spin correlation between two lattice sites is computed and it is found that for all states with spin $s \ll N^{\frac{1}{2}}$, the AMO correlation is antiferromagnetic (i.e., long-range order exists), a result agreeing with the obvious character of the single-determinant wavefunction from which spin eigenstates are projected. For $s = N$, the correlation is of course completely ferromagnetic.

Though the model is not a good one for metals, the type of electron correlation it considers closely resembles the spin-density waves (SDW) recently used by Overhauser to discuss the alkali metals. By contrast, though, the conditions of primary validity of this model are not those of the "electron gas" but the very low density limit where correlation effects are dominant.

I. INTRODUCTION

MOST theories of electronic structure in solids are based, at least in principle, on the Hartree–Fock self-consistent-field model. The energy-band method, which is a computational approximation to the Hartree–Fock scheme, is the basis for the theory of metallic systems, and can be applied with somewhat less success to semiconductors and insulators as well, at least for discussing electrical properties. In the Hartree–Fock approximation, one-electron states are determined as solutions to a self-consistent-field problem for one electron. Since the SCF potential exhibits the symmetry of the lattice, Bloch's well-known theorem states that such one-electron functions, the Bloch orbitals, are eigenfunctions of the lattice translation operators, associated with vectors $\mathbf{k}$ in the reciprocal lattice, located within the first Brillouin zone. The many-electron wavefunction for the ground state of the system is a determinantal product based on double occupancy of the

1539

Bloch orbitals $\psi(\mathbf{k})$ of lowest one-electron energies. Other states involve excitation of electrons to higher Bloch orbitals. This model is well suited to a formal treatment in one-electron terms, and the energy-band description of solids and their properties is such a treatment.

The Hartree–Fock scheme is an approximation because it ignores the correlation of electrons. For metallic systems this correlation is not an important factor for many properties involving one-electron excitations; only for properties such as the cohesive energy does it become significant. The situation is less clear cut for semiconductors and insulators, where correlation effects can dominate the wavefunction. In the diamond crystal, for example, a rather better account of the electronic structure can be given, in terms of localized pair correlation. Schmid and also other workers[1] have given such treatments for diamond, using the Heitler–London or "valence bond" (VB) method to describe the "saturated" tetrahedral valence pair bonds which each carbon atom forms with its near neighbors. Such studies are not directly relevant to the problems which concern us here, as is shown below, but they point up a serious defect of the Hartree–Fock method (and therefore of the energy-band method). As the lattice parameter a is increased, the Hartree–Fock uncorrelated wavefunction becomes less and less satisfactory as a description of a solid such as diamond, or even of elementary solids which *are* metals at their *equilibrium* lattice spacings. As $a \to \infty$ the limiting energy per atom for the HF SCF wavefunction is far above the true ground state of the neutral atom.[2] This work is of primary applicability to systems under conditions for which the Hartree–Fock approximation is not a good one, yet which on the other hand do not exhibit the Heitler–London "perfect pairing" behavior encountered, for example, in diamond.

The concept of pairing schemes plays an important role in the study of molecular systems. In such systems the treatment of electron correlation is based on the idea that pairs of electrons occupy local regions in space, and that *intra*pair correlations are most important. The Heitler–London or VB pair function is a simple form which includes such correlation. For atoms and molecules where it is clear that only one pairing scheme is reasonable, sophisticated methods analogous to the VB approach have been quite successful, from semiempirical and from purely theoretical viewpoints.[3-5] As noted above, diamond is a good example of a solid with perfect pairing. For some molecular systems, however, more than one Heitler–

London pairing scheme may play an important role, and methods emphasizing intrapair correlation must therefore consider all such pairings on an equivalent footing. This leads to the notion of *resonance* among the various *structures*, each representing a particular pairing scheme. The aromatic hydrocarbons are classic examples of such resonance.

An attempt to describe electronic structure in solids by means of Heitler–London localized pair bonds runs into difficulties for many solids, including metals, because the number of pairing schemes or structures which must be considered is infinite. In the case of metals, the Hartree–Fock method, of course, provides a far simpler and more elegant description. Pauling[6] has given a description of metallic solids and some of their properties, particularly cohesion and magnetic properties, in terms of a resonating valence bond model. While his work has been of immense qualitative value, and continues to stimulate productive thinking, doing accurate computations with such a model, as can be done with the Hartree–Fock scheme, is quite impractical. However, the attempt to construct a description with the same physical ideas but with computational simplicity led us to the work presented here. *An important conclusion, implied in this work, is that a "resonating valence bond" description or its physical equivalent is a quite different model of a solid from the Hartree–Fock scheme (or energy-band model), because of its emphasis on local spatial electron correlation.*

Long ago, Wigner[7] discussed the influence of electron correlation in an electron gas, and showed that at "*very*" low" densities a highly correlated structure, in which electrons are localized in space and have spin pairing with near neighbors, is a more stable state than is the usual Hartree–Fock wavefunction. The usual many-body theories of electron correlation[8] apply to the limit of *high* density and do not cover this limit considered by Wigner. Since correlations play a more important role in such a state than do the delocalization effects on the kinetic energy, it is not really an electron "gas" at all. The resonating valence bond model, with pair correlations between near-neighbor sites, really applies to the limit considered by Wigner, if we modify his gas model to a lattice structure with $2N$ ions and $2N$ electrons. It is not hard to see that a solid with this sort of electron structure is an insulator or a semiconductor rather than a metal. Since at higher electron densities the same system would certainly be metallic, the way in which the wavefunction changes, as a function of density (hence, of lattice parameter), is a topic of great importance to our understanding of materials with intermediate properties.

[1] L. A. Schmid, Phys. Rev. **92**, 1373 (1953); G. Dermit, *ibid.* **127**, 1110 (1962).

[2] J. C. Slater, Phys. Rev. **35**, 509 (1930).

[3] W. Moffitt, Proc. Roy. Soc. (London) **A210**, 245 (1951); T. Arai, J. Chem. Phys. **26**, 435 (1957).

[4] A. C. Hurley, J. E. Lennard-Jones, and J. A. Pople, Proc. Roy. Soc. (London) **A220**, 446 (1953).

[5] O. Sinanoğlu, Phys. Rev. **122**, 493 (1961); Proc. Roy. Soc. (London) **A260**, 379 (1961).

[6] L. Pauling, Phys. Rev. **54**, 899 (1938); J. Am. Chem. Soc. **69**, 542 (1947); Proc. Roy. Soc. (London) **A196**, 343 (1949).

[7] E. P. Wigner, Trans. Faraday Soc. **34**, 678 (1938).

[8] See, for example, M. Gell-Mann and K. Bruechner, Phys. Rev. **106**, 364 (1957); D. Bohm and D. Pines, *ibid.* **92**, 609 (1953); K. Sawada, K. Bruechner, N. Fukuda, and R. Brout, *ibid.* **108**, 257 (1957).

Mott[9] has drawn attention to phenomena to which energy-band theory seems inapplicable, and emphasized the importance of electron correlations in the low-density region. He has also made an interesting conjecture: in a regular lattice of this type, there should be a sharp transition from the metallic state to a nonmetallic state, as function of density. It was with Mott's idea in mind that this work was begun.

Attempts to make critical studies of electron correlation in lattices have often used finite systems as models; cyclic rings (one-dimensional lattices) have received considerable attention. In particular, rather extensive information exists for the hexagonal ring of hydrogen atoms, using a model basis set of the $1s$ atomic orbitals. Mattheiss[10] performed a complete configuration interaction calculation for this system, thus providing "exact" results against which the validity of various approximate schemes might be tested. Moskowitz[11] carried out the simple (one-parameter) alternant molecular orbital ("AMO") calculation for the ground singlet, and found that method to be more accurate than any other schemes then worked out. Recently we have presented results of calculations on this system using a novel method, the "MO–VBλ" approximation.[12] The results are in striking coincidence with those of the AMO method, being very slightly better at larger lattice spacings and very slightly worse at smaller ones. We mention the MO–VBλ scheme here because in our opinion it provides an instructive picture of the electron behavior, and helps to clarify the significance of the scheme which we describe here and apply to infinite lattices. The MO–VBλ approximation attempts to represent the ground state of the model system H_6 as a linear combination of (1) the Hartree–Fock ground-state (MO) wavefunction and (2) sophisticated valence bond wavefunctions, assuming resonance of all structures with pair bonds only between nearest-neighbor atoms. Neither of these components alone is satisfactory. At small lattice spacings, the MO wavefunction is the dominant component, as we may expect, but at larger lattice spacings the situation is reversed and the highly correlated valence bond structures dominant. (The transition between the two forms in the intermediate region is quite gradual.) The close agreement in the H_6 model between the MO–VBλ energies and those of the simple AMO method suggested that the latter might be a symmetrical account of the resonating pair correlations which appear in the valence bond model. As it finally turns out, these descriptions are in fact rather different for an infinite system (cf. Sec. VI); but the "localized alternant orbital" (LAO) represen-

tation of the AMO method presented here is a result of the attempt to display such symmetry explicitly.

In 1930 Slater[2] considered the possibility of introducing correlation effects in a solid containing two interpenetrating equivalent sublattices by associating electrons of a given spin with a given sublattice. The alternant molecular orbital method (AMO) first described by Löwdin and applied by him and co-workers to a number of molecular problems[13,14] achieves a variable degree of this sort of correlation, creating sets of molecular orbitals with alternating high and low densities on successive points, high density always being associated with a particular sublattice for one set of orbitals and with the other sublattice for a conjugate set. If α spins are associated with one set of alternant orbitals and β spins with the other, and the symmetrical spin singlet is projected from the determinantal product function, the resulting $^1\Gamma_1$ wavefunction is an approximation to the system ground states which includes a large amount of electron correlation. It can be proved (Refs. 14, 15; cf. also Sec. V) that as the number $2N$ of particles becomes large, the energy of this singlet eigenfunction differs from that of the initial determinant from which it is projected by terms of order $(2N)^{-1}$ in the energy per particle.

Using localized alternant orbitals, a representation of the $^1\Gamma_1$ AMO wavefunction is given here which leads directly to a simple expression for the total energy. The energy per atom appears as the (properly counted) interaction of a single site at the origin with its local environment, and the interactions can be grouped to give rapid convergence. We present first the energy per particle (for the single determinant) for the general problem, applicable to infinite lattices in one, two, or three dimensions; the case of the body-centered-cubic lattice is the one considered in those equations of the main text which differ for special cases. Appendix I cites special formulas valid for all the special cases considered. In Sec. V, general energy expressions, valid to terms of higher order in $(2N)^{-1}$, are derived for all the states with spin $s \ll N^{\frac{1}{2}}$ and for $s = N$, which can be projected from the single determinant.

After completion of this work, our attention was drawn to a paper by des Cloiseaux[16] and to a recent series of papers by Calais.[17] des Cloiseaux's paper contains a discussion of some of the relevant experimental facts which prompt development of an AMO-type model. He then develops a general type of AMO

[9] N. F. Mott, Can. J. Phys. **34**, 1356 (1956); Phil. Mag. **6**, 287 (1961).

[10] L. F. Mattheiss, Phys. Rev. **123**, 1209 (1961).

[11] J. W. Moskowitz, J. Chem. Phys. **38**, 677 (1963).

[12] E. G. Larson and W. R. Thorson, J. Chem. Phys. **43**, 3832 (1965).

[13] P.-O. Löwdin, Symp. Mol. Phys. Nikko, Japan, 1953, 13 (1954); Rev. Mod. Phys. **34**, 80 (1962).

[14] R. Pauncz, J. de Heer, and P.-O. Löwdin, J. Chem. Phys. **36**, 2247, 2257 (1962); J. de Heer, ibid. **37**, 2080 (1962); J. de Heer and R. Pauncz, ibid. **39**, 2314 (1963).

[15] P.-O. Löwdin, Phys. Rev. **97**, 1509 (1955).

[16] J. des Cloiseaux, J. Phys. Radium **20**, 606, 751 (1959).

[17] J. L. Calais, Arkiv Fysik **28**, 479 (1965) (I); **28**, 511 (1965) (II); **28**, 539 (1965) (III); **29**, 255 (1965) (IV).

single-determinant wavefunction for an infinite lattice, and, using the second quantization formalism, he makes a very approximate calculation based on the neglect of the more intractable off-diagonal matrix elements and some not unreasonable approximations for diagonal terms. The results show at least the gross features of a change from metallic to nonmetallic structure, and suggest that a statistical transition from one to the other as function of temperature may occur. In addition, des Cloiseaux also discusses the nature of conducting states and shows that a gap will emerge when correlation is strong. He seems to have been the first to apply the AMO single-determinant model to a solid. Our work differs from that of des Cloiseaux, in that (a) the calculation which we propose is exact and a variational form, and the energy expression is quite practicable for computation; (b) we obtain expressions for the energies of spin eigenfunctions, not simply those for a single determinant.

Using matrix projection operators in the ordinary Schrödinger representation, Calais[17] derives the energy expression for the energy per particle (i.e., the single-determinant energy) for the simple AMO model, in terms of the interaction of a site with its local environment. *Formally*, this expression corresponds exactly to ours. Calais does not seem to have known of des Cloiseaux's work. Calais has actually performed calculations, on a linear chain and on a bcc lattice of hydrogenic atoms; this demonstrates the practicability of the method. However, there is a conceptual error in his paper[17a] which affects the results for all but the one-dimensional lattice; this is due to an incorrect procedure in constructing alternant orbitals (see Sec. II). As noted above we have also obtained practical energy expressions for spin eigenfunctions and not merely for the single determinant.

Finally, in our opinion the localized alternant orbital (LAO) representation provides worthwhile additional insight into the physical situation for highly correlated electrons, which is not obvious in the *k*-space representations utilized by other workers.

II. ONE-ELECTRON BASIS FUNCTIONS

Consider a system with $2N$ physically equivalent sites, based on a lattice, with one site at each lattice point; the lattice consists of two equivalent interpenetrating sublattices, the nearest neighbors of any site in one sublattice belonging to the opposite sublattice. (Such a system is a special case of a more general class of systems called *alternant* structures.) Sites are *even* or *odd* depending on their relation to a chosen reference site, **0**; those on the same sublattice are *even*. The parity can always be displayed in a vector quantity **i** which uniquely specifies the actual site coordinate $\mathbf{R}_i - \mathbf{R}_0$.

[17a] *Note added in proof*: Dr. Calais has also independently drawn our attention to the need for this correction.

Assume that associated with each site **i** is a single *site orbital*, $u(\mathbf{i})$, which exhibits the full point symmetry of the site. Except for this restriction, $u(\mathbf{i})$ may otherwise be a freely variational form, at least in principle. In particular, though in the tight-binding limit we may expect the dominant contributions to $u(\mathbf{i})$ to come from "*s*-like" atomic functions centered on Site **i**, we may also form basis functions which are linear combinations of *p*, *d*, etc., type functions centered on *neighbor* sites of **i**, and which are invariant to point-group operations about **i**. (Such basis functions are like the "ligand orbitals" of crystal-field theory of complex ions.) These contributions from neighboring sites are very important for treating the "band mixing" or "broad band" effects which appear in the breakdown of the tight-binding scheme. They account for the fact that the true Wannier functions for a solid in the Hartree–Fock scheme may be rather delocalized. In this model, it is considered that $u(\mathbf{i})$ *is* more or less localized, but contributions from nearer neighbors may be included; since we argue that the model we treat is of primary validity in just the cases where tight binding tends to be valid, this practical restriction to "local" $u(\mathbf{i})$'s is not drastic.

The restriction of $u(\mathbf{i})$ to point-symmetry invariance at **i** means that in the present treatment no account is given of band symmetries other than those which can interact with a pure "*s*-type" band. The extension of the theory to include sets of site orbitals which span other representations of the space group should be easily possible but we do not present it in this paper.

We wish to calculate the energy of the ground state and certain low-lying excited states for such a system, assuming that with $2N$ electrons it is electrically neutral. For the case of finite N (appearing only in the one-dimensional case) we restrict ourselves to those alternant systems for which the molecular orbital theory ground state is a $^1\Gamma_1$ (closed shells.) For the case $N \rightarrow \infty$ we nominally consider the ground state to be $^1\Gamma_1$, though it can be shown that states of nonzero multiplicity (but $\ll N^{\frac{1}{4}}$) are degenerate with the $^1\Gamma_1$ to terms of order at most N^{-1} in the energy per particle.

For any system meeting the above restrictions the method to be developed is generally valid. The special cases for which we have derived certain special results are (1) the one-dimensional lattice, (2) the square-plane lattice, (3) the simple-cubic lattice, (4) the body-centered-cubic lattice. In the Appendix the special results for Cases (1)–(3) are cited, while the main text applies to Case (4).

Born–von Karman periodic boundary conditions are assumed for the infinite lattice cases. In the bcc system we take the periods to be $2n$ lattice spacings along each of the primitive axes, so that $2N = (2n)^3$. The dimensionless site vector **j** has components $0, \pm 1, \cdots, \pm (n-1), n$ on each of these axes. In the more convenient Cartesian coordinates for the bcc lattice the allowed site vector components j_x, j_y, j_z are either all

integers (even sublattice) or all odd half-integers (odd sublattice).

Let us now define various types of one-electron functions, all of them based on the set $\{u(\mathbf{j})\}$.

A. Bloch-Type Orbitals $\psi(\mathbf{k})$

These are given by the usual expression

$$\psi(\mathbf{k}) = (2n)^{-\frac{3}{2}} C(\mathbf{k}) \sum_j \exp[i\pi(\mathbf{k}\cdot\mathbf{j})] u(\mathbf{j}), \qquad (1)$$

where $n\mathbf{k}$ is a dimensionless vector with components $0, \pm 1, \cdots, \pm(n-1)$, n on each of the fundamental reciprocal lattice translations. The Cartesian components of $\mathbf{k}$ can be calculated readily but it must be remembered in counting $\mathbf{k}$ points for later calculation that the reciprocal lattice is face centered cubic. The allowed values of $\mathbf{k}$ may be located in the first Brillouin zone of the reciprocal lattice, which is the rhombic dodecahedron shown in Fig. 1. $C(\mathbf{k})$ is a normalization constant to be determined from the equation

$$[C(\mathbf{k})]^{-2} = \sum_j \exp(i\pi\mathbf{k}\cdot\mathbf{j}) S_0(\mathbf{j}), \qquad (2)$$

where the overlap integral $S_0(\mathbf{j})$ is defined

$$S_0(\mathbf{j}) \equiv \int d\tau_1 u(\mathbf{j}; 1) u(0; 1). \qquad (3)$$

Although Eq. (1) represents the Bloch-type orbital formally as a tight-binding expression, the fact that $u(\mathbf{i})$ may contain contributions from sites other than $\mathbf{i}$ means that we are not actually making the tight-binding approximation.

B. Wannier-Type Orbitals (Orthogonal Site Orbitals), $\omega(\mathbf{j})$

These are given by either an inverse transformation from the $\psi(\mathbf{k})$ or directly from $u(\mathbf{i})$

$$\omega(\mathbf{j}) = (2n)^{-\frac{3}{2}} \sum_k \psi(\mathbf{k}) \exp[-i\pi(\mathbf{k}\cdot\mathbf{j})] \qquad (4a)$$

$$= \sum_l d_0(\mathbf{l}-\mathbf{j}) u(\mathbf{l}), \qquad (4b)$$

where

$$d_0(\mathbf{m}) = \int d^3k\, C(\mathbf{k}) \exp(-i\pi\mathbf{k}\cdot\mathbf{m}); \qquad (5a)$$

it is also equal to the $(0, \mathbf{m})$ element of the square root of the inverse of the cyclic overlap matrix $\mathbf{S}$, whose $(\mathbf{i}, \mathbf{l})$ element is $S_0(\mathbf{l}-\mathbf{i})$ (cf. Ref. 13), i.e.,

$$d_0(\mathbf{m}) = (\mathbf{S}^{-\frac{1}{2}})_{0,\mathbf{m}} \qquad (5b)$$

(note that $\mathbf{S}$ and $\mathbf{S}^{-\frac{1}{2}}$ are symmetric). In (5a) the integration extends over the first Brillouin zone of the reciprocal lattice, but can be reduced to that over a single octant of the zone because of the cubic symmetry.
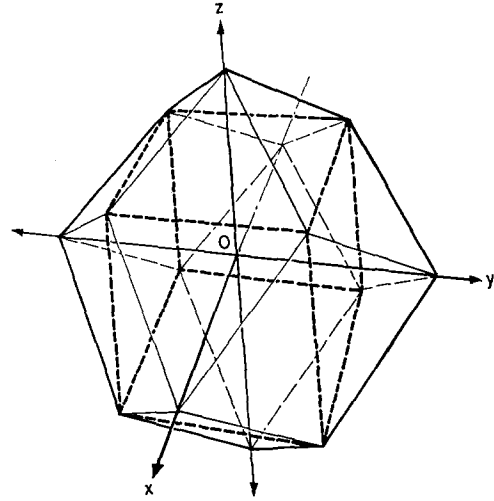

FIG. 1. The first Brillouin zone of the body-centered-cubic lattice. The inscribed cube shown is the sublattice zone. In Cartesian $\mathbf{k}$ space the edges of the cube are of *length* 2 but on such an edge only n $\mathbf{k}$ points are encountered since the reciprocal lattice is face centered cubic; the cube thus contains $4n^3$ $\mathbf{k}$ points and the full Brillouin zone contains $8n^3$ $\mathbf{k}$ points.

The degree of localization of the $\omega(\mathbf{j})$ depends upon that of the $u(\mathbf{j})$. The set $\{\omega(\mathbf{j})\}$ is an orthonormal set.

C. Sublattice Bloch-Type Orbitals, $\xi(\mathbf{k})$ and $\bar{\xi}(\mathbf{k})$

So far no use has been made of the two-sublattice structure. The first Brillouin zone of the *sublattice* contains just N $\mathbf{k}$ vectors, half as many as that of the full lattice. Let the sublattice zone be placed symmetrically around zero inside the full lattice zone; the remaining portion of the full lattice zone is also the *second* Brillouin zone of the sublattice; we call it the *residual* zone. For the bcc lattice, the sublattice zone is the inscribed cube, shown in Fig. 1. For every $\mathbf{k}$ inside the sublattice zone, there corresponds an unique conjugate $\bar{\mathbf{k}}$ in the residual zone, such that the functions

$$\xi(\mathbf{k}) = 1/\sqrt{2}[\psi(\mathbf{k}) + \psi(\bar{\mathbf{k}})],$$

$$\bar{\xi}(\mathbf{k}) = 1/\sqrt{2}[\psi(\mathbf{k}) - \psi(\bar{\mathbf{k}})], \qquad (6)$$

are sublattice Bloch-type orbitals (SBO's) having density [in terms of the $\omega(\mathbf{j})$'s] only on a single sublattice, $\xi(\mathbf{k})$ being associated with the *even* sublattice and $\bar{\xi}(\mathbf{k})$ with the *odd* sublattice. In order to construct SBO's which truly have this property, it is *necessary* to choose the $\mathbf{k}$'s in the sublattice zone and the $\bar{\mathbf{k}}$'s in the *residual* zone. The relation between $\mathbf{k}$ and $\bar{\mathbf{k}}$ for the bcc system is

$$\bar{\mathbf{k}} = \mathbf{k} + \mathbf{K}, \qquad (7)$$

where $\mathbf{K}$ can be any one of the vectors $(2, 0, 0)$, $(0, 2, 0)$, $(0, 0, 2)$ (*Cartesian system*); (7) determines $\bar{\mathbf{k}}$ to within a reciprocal lattice translation, and uniquely within the residual zone. The division of the $\mathbf{k}$ vectors in the first Brillouin zone of the full lattice into the set falling in the sublattice zone and the set in the

residual zone is an important difference between this work and that of Calais.[17] Calais assumed that independent and orthogonal sets of SBO's can be constructed regardless of the choice of the sets of **k**'s and **k̄**'s; in particular, he took the set of vectors **k** to be those lying within a sphere of half the volume of the first Brillouin zone, as an attempt to achieve a **k**-space occupancy in the low-correlation limit similar to that for an electron gas. Calais' procedure is formally rather different from that used here. He defines [Paper I, Eqs. (9)] certain operators O_{kI}, O_{kII}, which would generate, respectively (except for a factor $1/4n^3$), our $\bar{\xi}(\mathbf{k})$ and $\xi(\mathbf{k})$, from $\omega(0)$, *provided that the set of* **k**'s *used lies in the sublattice zone.* From the operators O_{kI}, O_{kII} Calais constructs "splitting operators" Q_{kI}, Q_{kII}, which analogously would generate our alternant Bloch-type orbitals $\bar{\Phi}(\mathbf{k})$ and $\Phi(\mathbf{k})$ (cf. Sec. II.D below). However, his choice of **k**'s makes the corresponding Φ's and $\bar{\Phi}$'s neither complete nor orthogonal. It is easy to show that his Eq. (12a), which asserts their orthogonality,

$$Q_{kI,II}Q_{k'I,II}=0, \qquad \mathbf{k}'\neq\mathbf{k},$$

is not satisfied. There exist **k**'s, inside a Fermi sphere with $4n^3$ **k** points, which lie in the *residual* zone, i.e., in the *second* Brillouin zone of the sublattice. As a result, each such **k**' is equivalent, to within a reciprocal vector *of the sublattice*, to a vector **k**'' lying *inside the sublattice zone.* Since either $O_{k''I}O_{k'I}$ or $O_{k''II}O_{k'II}$ contains a full sum over the *sublattice translations* only, they do not give zero for such a pair, nor does the left side of Calais' equation (12a). However, with the one modification that the **k**'s be assigned as in this work to the sublattice zone, all the equations of Calais' paper become formally valid. In the development we use we note points of equivalence between his results and ours.

D. Alternant Bloch-Type Orbitals, $\Phi(\mathbf{k})$ and $\bar{\Phi}(\mathbf{k})$

These are given by

$$\Phi(\mathbf{k};\lambda)=(1+\lambda^2)^{-\frac{1}{2}}[\xi(\mathbf{k})+\lambda\bar{\xi}(\mathbf{k})]$$
$$=\cos\theta\psi(\mathbf{k})+\sin\theta\psi(\bar{\mathbf{k}}),$$
$$\bar{\Phi}(\mathbf{k};\lambda)=(1+\lambda^2)^{-\frac{1}{2}}[\bar{\xi}(\mathbf{k})+\lambda\xi(\mathbf{k})]$$
$$=\cos\theta\psi(\mathbf{k})-\sin\theta\psi(\bar{\mathbf{k}}), \qquad (8)$$

with $0\leq\lambda\leq1$, $0\leq\theta\leq\pi/4$, and

$$\cos\theta=(1+\lambda)/[2(1+\lambda^2)]^{\frac{1}{2}}. \qquad (9)$$

The variable mixing parameter λ (or θ) describes the degree of alternation in density encountered in the Φ or $\bar{\Phi}$, on passing from one sublattice to the other.

E. Conjugate Wannier-Type Orbitals, $\bar{\omega}(\mathbf{j})$

Since $\bar{\mathbf{k}}$ is given by Eq. (7) and $\exp(-i\pi\mathbf{K}\cdot\mathbf{j})=+1$ if **j** is even, -1 if **j** is odd, the function $\omega(\mathbf{j})$, for **j**

even, can be written

$$\omega(\mathbf{j})=(4n^3)^{-\frac{1}{2}}\sum_{\mathbf{k}\text{ in s.z.}}\exp(-i\pi\mathbf{k}\cdot\mathbf{j})\xi(\mathbf{k}); \qquad (10a)$$

the notation "**k** in s.z." means the sum is performed over the *sublattice zone.* The analogous transform (again with **j** *even*) of $\bar{\xi}(\mathbf{k})$ is a new function, to which we have given the name conjugate Wannier-type orbital (CWO):

$$\bar{\omega}(\mathbf{j})=(4n^3)^{-\frac{1}{2}}\sum_{\mathbf{k}\text{ in s.z.}}\exp(-i\pi\mathbf{k}\cdot\mathbf{j})\bar{\xi}(\mathbf{k}) \qquad (\mathbf{j}\text{ even}).$$
$$(10b)$$

The roles of ξ and $\bar{\xi}$ are reversed if **j** is an odd site vector, $\bar{\xi}$ then generating $\omega(\mathbf{j})$ and ξ, $\bar{\omega}(\mathbf{j})$; $\bar{\omega}(\mathbf{i})$ may be expanded in $\omega(\mathbf{j}')$'s,

$$\bar{\omega}(\mathbf{i})=\sum_{\mathbf{j}'}b_0(\mathbf{j}'-\mathbf{i})\omega(\mathbf{j}'); \qquad (11)$$

the coefficients $b_0(\mathbf{m}')$ are nonzero only for $\mathbf{m}'$ on the odd sublattice. They are closely related to certain matrices **U** defined by Calais [Eqs. (20), Ref. 17, first paper]; in particular, formally,

$$[U_IU_{II}^++U_{II}U_I^+]_{1,j'}=b_0(1-\mathbf{j}'). \qquad (12)$$

They can easily be computed from the formula

$$b_0(\mathbf{m})=\int_{\mathbf{k}\text{ in subzone}}d\mathbf{k}\exp(-i\pi\mathbf{k}\cdot\mathbf{m})-\delta(\mathbf{m}), \qquad (13a)$$

which gives for the bcc case

$$b_0(\mathbf{m})=\left(\frac{\sin\pi m_x}{\pi m_x}\right)\left(\frac{\sin\pi m_y}{\pi m_y}\right)\left(\frac{\sin\pi m_z}{\pi m_z}\right)-\delta(\mathbf{m}); \qquad (13b)$$

$\delta(\mathbf{m})$ is 1 for $\mathbf{m}=0$ and zero otherwise. The integration in (13a) must take into account the fcc density of points in **k** space, if Cartesian coordinates are used. The matrix **B** whose 1, **j**' element is $b_0(\mathbf{j}'-1)$ is cyclic, orthogonal, and symmetric, and $b_0(\mathbf{m})$ also exhibits cubic point symmetry. We therefore have

$$\omega(\mathbf{j})=\sum_{1'}b_0(\mathbf{j}-1')\bar{\omega}(1'), \qquad (14)$$

and, if **j** and 1 are of common parity

$$\langle\bar{\omega}(\mathbf{j})|\bar{\omega}(1)\rangle=\delta(1-\mathbf{j}), \qquad (15a)$$
$$\langle\omega(j)|\bar{\omega}(1)\rangle=0. \qquad (15b)$$

Since the coefficients $b_0(\mathbf{m})$ fall off rather slowly with increasing $\mathbf{m}$ in certain directions, $\bar{\omega}(0)$ is *not* as localized as is $\omega(0)$. This makes no difficulties for the convergence of the energy expression which we derive, but it is a point of some relevance for physical interpretations of the model.

F. Localized Alternant Orbitals, $\phi(\mathbf{j};\lambda)$ and $\bar{\phi}(\mathbf{j};\lambda)$

These are given by

$$\phi(\mathbf{j};\lambda)=(1+\lambda^2)^{-\frac{1}{2}}[\omega(\mathbf{j})+\lambda\bar{\omega}(\mathbf{j})],$$
$$\bar{\phi}(\mathbf{j};\lambda)=(1+\lambda^2)^{-\frac{1}{2}}[\bar{\omega}(\mathbf{j})+\lambda\omega(\mathbf{j})], \qquad (\mathbf{j}\text{ even}) \qquad (16)$$

or equivalently by transforms of $\Phi(\mathbf{k};\lambda)$ and $\bar{\Phi}(\mathbf{k};\lambda)$ analogous to Eqs. (10). One might just as well, of course, have constructed LAO's for the *odd* sublattice points, and the question arises whether the apparent asymmetric bias in favor of one sublattice is useful. We show that the singlet state constructed from *either* set of $2N$ LAO's is the same ${}^1\Gamma_1$ state and that it is identical to the AMO ${}^1\Gamma_1$ state constructed in the usual way from the $\Phi(\mathbf{k})$ and $\bar{\Phi}(\mathbf{k})$.

The LAO's of a given site parity have the following properties:

$$\langle\phi(1)\mid\phi(\mathbf{j})\rangle=\delta(\mathbf{j}-1)=\langle\bar{\phi}(1)\mid\bar{\phi}(\mathbf{j})\rangle,$$

$$\langle\phi(1)\mid\bar{\phi}(\mathbf{j})\rangle=\delta(\mathbf{j}-1)[2\lambda/(1+\lambda^2)]$$

$$(1,\mathbf{j}\ \text{same parity}),\quad(17)$$

and

$$\bar{\phi}(\mathbf{j})=\sum_{\mathbf{l}'}b_0(\mathbf{l}'-\mathbf{j})\phi(\mathbf{l}'),$$

$$\phi(\mathbf{j})=\sum_{\mathbf{l}'}b_0(\mathbf{l}'-\mathbf{j})\bar{\phi}(\mathbf{l}').\quad(18)$$

III. MANY-ELECTRON STATES

This section concerns the construction and properties of wavefunctions for certain spin states (mainly those with total spin $s\ll N^{\frac{1}{2}}$).

Consider three generating products defined as follows:

$$\chi_{\mathrm{I}}=\prod_{\mathbf{j}'}[\phi(\mathbf{j}';\lambda;\mu_{j'})\alpha(\mu_{j'})\bar{\phi}(\mathbf{j}';\lambda;\mu_{j'}+N)\beta(\mu_{j'}+N)],$$

$$(19\mathrm{a})$$

where the site vector $\mathbf{j}'$ runs over only the *odd* sublattice and $\mu_{j'}$ is an ordered indexing of electrons from 1 to N;

$$\chi_{\mathrm{II}}=\prod_{\mathbf{j}}[\phi(\mathbf{j};\lambda;\mu_j)\alpha(\mu_j)\bar{\phi}(\mathbf{j};\lambda;\mu_j+N)\beta(\mu_j+N)],$$

$$(19\mathrm{b})$$

where $\mathbf{j}$ runs over only the *even* sublattice; and

$$\chi_{\mathrm{AMO}}$$
$$=\prod_{\mathbf{k}\ \text{in s. z.}}[\Phi(\mathbf{k};\lambda;\mu_k)\alpha(\mu_k)\bar{\Phi}(\mathbf{k};\lambda;\mu_k+N)\beta(\mu_k+N)],$$

$$(19\mathrm{c})$$

$\mathbf{k}$ running over the sublattice Brillouin zone.

In order to satisfy the Pauli principle, construct the determinantal products $A(2N)\chi_{\mathrm{I}}$, $A(2N)\chi_{\mathrm{II}}$, and $A(2N)\chi_{\mathrm{AMO}}$, the antisymmetrizer $A(2N)$ being given by

$$A(2N)=(2N\,!)^{-\frac{1}{2}}\sum_P(-1)^PP,\quad(20)$$

the sum being over the group of permutations P of $2N$ electrons.

Spin eigenfunctions are formed by the projection operator technique of Löwdin.[15] Following his notation, we define

$$T_0{}^{\mathrm{I}}\equiv A\chi_{\mathrm{I}};\quad T_0{}^{\mathrm{II}}\equiv A\chi_{\mathrm{II}};\quad T_0{}^{\mathrm{AMO}}\equiv A\chi_{\mathrm{AMO}};\quad(21\mathrm{a})$$

any one of these determinants T_0 can be written symbolically

$$T_0\equiv\{\alpha\alpha\alpha\cdots\alpha\mid\beta\beta\beta\cdots\beta\},\quad(21\mathrm{b})$$

a form which expresses the ordered assignments of α spins to ϕ's or Φ's and of β spins to $\bar{\phi}$'s or $\bar{\Phi}$'s, in the products χ. Now, consider a sequence of symmetrical sums of determinants T_μ, obtained by adding all possible distinct determinants formed by switching any μ α spins in the left-hand part of the curly brackets of T_0 with any μ β spins in the right-hand part. T_1 for example can be symbolically written

$$T_1=(\beta\alpha\alpha\cdots+\alpha\beta\alpha\cdots+\alpha\alpha\beta\cdots+\cdots+\alpha\alpha\alpha\cdots\alpha\beta)\mid(\alpha\beta\beta\beta\cdots+\beta\alpha\beta\beta\cdots+\beta\beta\alpha\beta\cdots+\cdots+\beta\beta\beta\cdots\beta\alpha),\quad(22)$$

and so forth. There are $\binom{N}{\mu}^2$ determinants in T_μ. T_N is just given by

$$T_N=\{\beta\beta\beta\cdots\mid\alpha\alpha\alpha\cdots\}.\quad(23)$$

The projection operator ${}^{2s+1}O$ for a state with multiplicity $2s+1$ is

$${}^{2s+1}O\equiv\prod_{\substack{\nu=1\\\nu\neq s}}^{N}\left\{\frac{S^2-\nu(\nu+1)\hbar^2}{[s(s+1)-\nu(\nu+1)]\hbar^2}\right\},\quad(24)$$

and it can be shown[15] that

$${}^{2s+1}\Psi\equiv{}^{2s+1}OT_0=\sum_{\mu=0}^{N}C(s;\mu)T_\mu;\quad(25)$$

the specific form of the coefficients $C(s;\mu)$ is treated later (Sec. V), but for the present we require only the property

$$C(s;N-\mu)=(-1)^{N+s}C(s;\mu);\quad(26)$$

this can easily be deduced from results cited in Sec. V or Ref. 15.

The translational symmetry properties of ${}^{2s+1}\Psi_{\mathrm{I}}$, ${}^{2s+1}\Psi_{\mathrm{II}}$, and ${}^{2s+1}\Psi_{\mathrm{AMO}}$, and their equivalence, can be show as follows: $T_N{}^{\mathrm{II}}$ is the Slater determinant based on the ordered product

$$\prod_{\mathbf{j}\ \text{even}}[\phi(\mathbf{j};\lambda;\mu_j)\beta(\mu_j)\bar{\phi}(\mathbf{j};\lambda;\mu_j+N)\alpha(\mu_j+N)];$$

let $\mathbf{M}_N{}^{\mathrm{II}}$ be the function matrix whose determinant is $T_N{}^{\mathrm{II}}$. With the transformations (18) we can easily prove that

$$\mathbf{M}_N{}^{\mathrm{II}}=\mathbf{B}'\mathbf{M}_0{}^{\mathrm{I}\prime},\quad(27)$$

where $\mathbf{B}'$ is a matrix which factors into diagonal block form

$$B' = \begin{bmatrix} \mathbf{B} & 0 \\ 0 & \mathbf{B} \end{bmatrix}, \qquad (28)$$

each block being $N \times N$, and $\mathbf{B}$ is the orthogonal symmetric matrix with $b_0(\mathbf{1}'-\mathbf{j})$ as its $\mathbf{1}'$, $\mathbf{j}$ element; $M_0^{I'}$, as its name suggests, is a matrix differing from M_0^{I}, the matrix whose determinant is T_0^{I}, by permutation of N rows or N columns. Hence we have, since $\mathrm{Det} \, | \, B \, | = \pm 1$, that

$$T_N^{II} = \mathrm{Det} \, | \, M_0^{I'} \, | = (-1)^N T_0^{I}. \qquad (29)$$

Similarly, we can prove $T_N^{I} = (-1)^N T_0^{II}$, and, in general,

$$T_\mu^{II} = (-1)^N T_{N-\mu}^{I}. \qquad (30)$$

It follows from Eqs. (26) and (25), therefore, that

$$^{2s+1}\Psi_I = (-1)^s \, {}^{2s+1}\Psi_{II}. \qquad (31)$$

The translational properties of $^{2s+1}\Psi_I$ and $^{2s+1}\Psi_{II}$ may now be demonstrated by observing that an elementary translation which moves sites on one sublattice to the opposite sublattice converts χ_{II} into χ_I, except for a permutation of even parity; thus $^{2s+1}\Psi_{II}$ goes into $^{2s+1}\Psi_I$ under such a translation. It follows from Eq. (31) that $^{2s+1}\Psi_{II}$ (and therefore also $^{2s+1}\Psi_I$) is a $^{2s+1}\Gamma_1$ state, associated with a wave vector $\mathbf{k}=0$ in the reciprocal lattice if s is *even*, while if s is *odd*, $^{2s+1}\Psi_{II}$ is associated with $\mathbf{k}=\mathbf{K}$, and changes sign under a primitive lattice

translation of one unit in any direction. We call such states $^{2s+1}\Gamma(\mathbf{K})$ states.

The equivalence of $^{2s+1}\Psi_{II}$ to $^{2s+1}\Psi_{AMO}$ can be proved by noting that the matrix of the transformation generating χ_{II} from χ_{AMO} is unitary, with a determinant $+1$ [consider the analog of Eqs. (10) for ϕ, $\bar{\phi}$ in terms of Φ, $\bar{\Phi}$]; as a consequence, $T_0^{AMO} = T_0^{II}$, and $^{2s+1}\Psi_{AMO} = {}^{2s+1}\Psi_{II}$. Though we work in the even-site LAO basis, we now drop the subscripts I, II, AMO on the many-particle wavefunctions $^{2s+1}\Psi$.

It can be proved (Refs. 14 and 15; cf. Sec. V) that to terms of order (N^{-1}), the energy per atom for states with $s \ll N^{\frac{1}{2}}$ is just that of the single determinant T_0^{II}. In Sec. IV we give an expression for this energy per atom. Section V is concerned with the spin-dependent splittings which are of higher order.

IV. SINGLE-DETERMINANT ENERGY

In symbolic form the Hamiltonian is

$$H = \sum_{i=1}^{2N} h(i) + \frac{1}{2} \sum_i \sum_{i \neq i} g(i,j) + W_N, \qquad (32)$$

where $h(1)$ is the kinetic energy of Electron 1 plus its electrostatic interaction with each lattice site (nucleus), $g(1,2)$ is the Coulomb repulsion of 1 and 2, and W_N is the (nuclear) repulsion energy of the lattice site cores. The cores may be taken to be other than Coulomb potential sources but for convergence the system must be neutral with one electron per atom. Utilizing orthogonalities and the physical equivalence of lattice sites one obtains the energy per atom, $\bar{\epsilon}$

$$\bar{\epsilon} = \frac{1}{2}\{ [\langle \varphi_0 | h | \varphi_0 \rangle + \langle \bar{\varphi}_0 | h | \bar{\varphi}_0 \rangle + \langle \varphi_0 \varphi_0 | g | \bar{\varphi}_0 \bar{\varphi}_0 \rangle]$$

$$+ \frac{1}{2} \sum_{j \neq 0,\, \text{even}} [\langle \varphi_0 \varphi_0 | g | \varphi_j \varphi_j \rangle - \langle \varphi_0 \varphi_j | g | \varphi_0 \varphi_j \rangle + \langle \bar{\varphi}_0 \bar{\varphi}_0 | g | \bar{\varphi}_j \bar{\varphi}_j \rangle - \langle \bar{\varphi}_0 \bar{\varphi}_j | g | \bar{\varphi}_0 \bar{\varphi}_j \rangle + 2 \langle \varphi_0 \varphi_0 | g | \bar{\varphi}_j \bar{\varphi}_j \rangle] \} + (1/2N) W_N;$$

$$(33)$$

in the symbolic notation for two-electron integrals, functions to the lhs of g have argument 1, those to the rhs have argument 2. The electronic energy in (33) is clearly the pair self-energy plus half their energy of interaction with all other pairs (all divided by 2). Using Eqs. (16) we can express $\bar{\epsilon}$ in terms of the variable parameter $z = 2\lambda/(1+\lambda^2)$ and integrals involving ω_j and $\bar{\omega}_j$ (j here *always* refers to sites of *even* parity only). The result is (we drop the symbols ω and g)

$$\bar{\epsilon} = \frac{1}{2}\{ \langle 0 | h | 0 \rangle + \langle \bar{0} | h | \bar{0} \rangle + 2z \langle 0 | h | \bar{0} \rangle \} + (1/2N) W_N$$

$$+ \sum_{j\, \text{even}} \{ \frac{1}{2} \langle 00 | j\bar{j} \rangle + \frac{1}{4} [\langle 00 | jj \rangle + \langle \bar{0}\bar{0} | j\bar{j} \rangle] + z^2 [\langle 0\bar{0} | j\bar{j} \rangle] + z [\langle 00 | j\bar{j} \rangle + \langle 0\bar{0} | j\bar{j} \rangle] - \frac{1}{4} [\langle 0j | 0j \rangle + \langle \bar{0}j | \bar{0}j \rangle]$$

$$- \frac{1}{8} z^2 [2 \langle 0\bar{j} | \bar{0}j \rangle + 2 \langle 0\bar{j} | 0j \rangle + 2 \langle 0j | \bar{0}j \rangle - \langle 0j | 0j \rangle - \langle \bar{0}j | \bar{0}j \rangle] - \frac{1}{2} z [\langle 0j | 0\bar{j} \rangle + \langle \bar{0}j | \bar{0}\bar{j} \rangle] \}. \qquad (34)$$

Were further simplification not possible, the slow convergence properties of the coefficients $b_0(\mathbf{m}')$ would make computation of this expression impractical. Fortunately, the orthogonal and symmetric properties of the $\mathbf{B}$ matrix,

$$\sum_{j'} b_0(\mathbf{j}'-\mathbf{i}) b_0(\mathbf{j}'-\mathbf{l}) = \delta(\mathbf{i}-\mathbf{l}), \qquad (35)$$

permit us to derive the following formulas:

$$\langle \bar{\omega}_0 \mid h \mid \bar{\omega}_0 \rangle = \langle \omega_0 \mid h \mid \omega_0 \rangle; \tag{36a}$$

$$\sum_{j \text{ even}} \langle \omega_0 \omega_0 \mid g \mid \bar{\omega}_j \bar{\omega}_j \rangle = \sum_{j' \text{ odd}} \langle \omega_0 \omega_0 \mid g \mid \omega_{j'} \omega_{j'} \rangle; \tag{36b}$$

$$\sum_{j \text{ even}} \langle \omega_0 \bar{\omega}_j \mid g \mid \omega_0 \bar{\omega}_j \rangle = \sum_{j' \text{ odd}} \langle \omega_0 \omega_{j'} \mid g \mid \omega_0 \omega_{j'} \rangle; \tag{36c}$$

$$\sum_{j \text{ even}} \langle \bar{\omega}_0 \bar{\omega}_0 \mid g \mid \bar{\omega}_j \bar{\omega}_j \rangle = \sum_{j' \text{ odd}} \langle \bar{\omega}_0 \bar{\omega}_0 \mid g \mid \omega_{j'} \omega_{j'} \rangle$$

$$= \sum_{j' \text{ odd}} \langle \omega_0 \omega_0 \mid g \mid \bar{\omega}_{j'} \bar{\omega}_{j'} \rangle$$

$$= \sum_{j \text{ even}} \langle \omega_0 \omega_0 \mid g \mid \omega_j \omega_j \rangle; \tag{36d}$$

similarly,

$$\sum_{j \text{ even}} \langle \bar{\omega}_0 \bar{\omega}_j \mid g \mid \bar{\omega}_0 \bar{\omega}_j \rangle = \sum_{j \text{ even}} \langle \omega_0 \omega_j \mid g \mid \omega_0 \omega_j \rangle; \tag{36e}$$

$$\sum_{j \text{ even}} \langle \omega_0 \bar{\omega}_0 \mid g \mid \bar{\omega}_j \bar{\omega}_j \rangle = \sum_{j' \text{ odd}} \langle \omega_0 \omega_0 \mid g \mid \omega_{j'} \bar{\omega}_{j'} \rangle. \tag{36f}$$

Using these formulas to reduce the matrix elements as far as possible to those involving only ω's, and utilizing Eq. (11) for the rest of the reduction, we find $\bar{\epsilon}$ in terms of matrix elements of Wannier-type orbitals only (ω and g are again dropped)

$$\bar{\epsilon} = \{ \langle 0 \mid h \mid 0 \rangle + \tfrac{1}{2} \sum_{j \text{ even}} [\langle 00 \mid jj \rangle - (1 - \tfrac{1}{2} z^2) \langle 0j \mid 0j \rangle] + \tfrac{1}{2} \sum_{j' \text{ odd}} [\langle 00 \mid j'j' \rangle - \tfrac{1}{2} z^2 \langle 0j' \mid 0j' \rangle] \}$$

$$+ z \{ \sum_{j' \text{ odd}} b_0(j') [\langle 0 \mid h \mid j' \rangle + \sum_{i, \text{ even}} (2 \langle 0j' \mid ii \rangle - \langle 0i \mid j'i \rangle)] \}$$

$$+ \tfrac{1}{4} z^2 \{ \sum_{i \text{ even}} \sum_{j' \text{ odd}} \sum_{l' \text{ odd}} b_0(j') b_0(l' - i) [4 \langle 0j' \mid il' \rangle - \langle 0l' \mid ij' \rangle - \langle 0i \mid l'j' \rangle] \} + (1/2N) W_N. \tag{37}$$

This expression gives the correct ϵ for $z \to 0$, which can be deduced by writing ω_j for φ_j, $\bar{\omega}_j$ for $\bar{\varphi}_j$ in (33).

We now show that there are no serious convergence difficulties when this expression is properly arranged. For convenience let us take the special case that the lattice sites are unit positive charges. Let the distance between two sites i, j be R_{ij}

$$\frac{W_N}{2N} = (2N)^{-1} \left(\tfrac{1}{2} \sum_i \sum_{k \neq i, \text{ even}} \frac{e^2}{R_{ik}} + \tfrac{1}{2} \sum_{j'} \sum_{l' \neq j', \text{ odd}} \frac{e^2}{R_{j'l'}} + \sum_{i \text{ even}} \sum_{j' \text{ odd}} \frac{e^2}{R_{ij'}} \right), \tag{38}$$

which can be simplified to

$$\frac{W_N}{2N} = \frac{1}{2} \left(\sum_{l \neq 0, \text{ even}} \frac{e^2}{R_{0l}} + \sum_{j' \text{ odd}} \frac{e^2}{R_{0j'}} \right). \tag{39}$$

The one-electron Hamiltonian can be written

$$h(1) = T(1) + \sum_{i, \text{ even}} V_i(1) + \sum_{j' \text{ odd}} V_{j'}(1). \tag{40}$$

For both odd and even sites let us separate the long-range component from the higher multipoles in the potential of a site charge cloud $\omega_i^2(2)$

$$\int d\tau_2 \omega_i^2(2) \frac{e^2}{r_{12}} = J_i(1) - V_i(1). \tag{41}$$

It is convenient to define the following quantities (for both odd and even sites):

$$U_i \equiv \langle \omega_0(1) \mid \tfrac{1}{2} [(e^2/R_{0i}) + V_i(1) + J_i(1)] \mid \omega_0(1) \rangle, \tag{42a}$$

$$X_i \equiv \langle \omega_0(1) \omega_i(1) \mid g(1, 2) \mid \omega_0(2) \omega_i(2) \rangle, \tag{42b}$$

$$\epsilon_0 \equiv \langle \omega_0(1) \mid T(1) + V_0(1) \mid \omega_0(1) \rangle. \tag{42c}$$

These groupings make it possible to collect the nuclear repulsion energy and the first curly brackets of Eq. (37) in a way which gives rapid convergence of the energy expression. The terms linear in z, in the second curly brackets of (37), can be conveniently computed as follows:

$$z \sum_{j', \text{odd}} b_0(j') [\langle 0 \mid h \mid j' \rangle + \sum_i (2\langle 0j' \mid ii \rangle - \langle 0i \mid j'i \rangle)]$$

$$= z \sum_{j' \text{ odd}} b_0(j') [\langle 0 \mid T \mid j' \rangle + \langle 0 \mid \sum_{l' \text{ odd}} V_{l'} - \sum_{i \text{ even}} V_i \mid j' \rangle + \sum_i (2\langle 0 \mid J_i \mid j' \rangle - K_{ij'}{}^0)], \quad (43a)$$

where

$$K_{ij'}{}^0 \equiv \langle 0i \mid g \mid j'i \rangle.$$

Since $\langle 0 \mid \sum_{l'} V_{l'} - \sum_i V_i \mid j' \rangle$ vanishes by symmetry, we have simply

$$z \sum_{j'} b_0(j') [\langle 0 \mid T \mid j' \rangle + \sum_{i \text{ even}} (2\langle 0 \mid J_i \mid j' \rangle - K_{ij'}{}^0)]. \quad (43b)$$

The terms in the third curly brackets of (37) cannot be simplified further. A convenient way of grouping them, however, is to treat them as interactions on the overlap density of $0, j'$ due to overlap densities clustered at sites i and summed over i; they then can be grouped with (43b)

$$\tfrac{1}{4}z^2 \sum_{j'} b_0(j') \{ \sum_{i \text{ even}} [\sum_{l'} b_0(l'-i)(4\langle 0j' \mid il' \rangle - \langle 0l' \mid ij' \rangle - \langle 0i \mid j'l' \rangle)]\}. \quad (43c)$$

The convergence of the expressions (43b) and (43c) is more rapid than might appear at first glance. In (43b) the terms of longest range come from the multipole interactions J_i and act on the overlap density $(0, j')$. Because of the orthogonality of these orbitals the interactions fall off strongly for i far from the center of 0 and j'. Similarly, none of the terms in (45c) have long-range character, for the same reason, except for the case $l'=j'$ and $i=0$; these terms can be grouped together,

$$-\tfrac{1}{4}z^2 \sum_{j'} [b_0(j')]^2 \langle 00 \mid j'j' \rangle, \quad (44)$$

and even for them convergence will be sufficiently rapid to make practical computation feasible, since all $b_0{}^2(j')$ fall off at least as strongly as $R_{0j}{}^{-2}$ and in most cases much more rapidly. The final expression for the energy per particle is

$$\bar{\epsilon} = (\epsilon^0 + \tfrac{1}{4}z^2 \langle 00 \mid 00 \rangle) + \sum_{i \neq 0, \text{ even}} [U_i - \tfrac{1}{2}(1 - \tfrac{1}{2}z^2)X_i]$$

$$+ \sum_{j' \text{ odd}} (U_{j'} - \tfrac{1}{4}z^2 X_{j'}) + z \sum_{j'} b_0(j') [\langle 0 \mid T \mid j' \rangle + \sum_{i \text{ even}} (2\langle 0 \mid J_i \mid j' \rangle - K_{ij'}{}^0)]$$

$$+ \tfrac{1}{4}z^2 \sum_{j'} b_0(j') \sum_{i \text{ even}} [\sum_{l'} b_0(l'-i)(4\langle 0j' \mid il' \rangle - \langle 0l' \mid ij' \rangle - \langle 0i \mid l'j' \rangle)]. \quad (45)$$

If the integrals in terms of Wannier orbitals, or certain symmetrical sums of them, were available, direct evaluation of this expression could be made. However, it seems more practical to use the transformation of Eq. (4b) and its symmetry properties to reduce (45) to an expression involving integrals with the fundamental site orbitals $u(i)$. Even without doing that reduction in detail, however, it is a routine if tedious procedure to show that Eq. (45) is equivalent *formally* to the energy expression obtained by Calais [Ref. 17, Paper I, Eqs. (50), (51)]. Calais' expression also gives the energy as the interactions of a single site with its environment, but his expression contains the elements of matrices R_i which combine the effects of non-orthogonality of the site orbitals with the effects of alternation which appear, for example, in the matrix elements $b_0(m')$. The terms in (45) which contain sums over $b_0(m')$ correspond exactly to terms arising from the elements q_2 (elements of the matrix R_2) in

Calais' equations (51). The equivalence of the two energy expressions is, of course, a formal one only because of Calais' incorrect choice of "occupied" k's for generating AMO's.

In this paper we have not considered the problem of actually calculating the energy from Eq. (45) or its transformation in terms of site-orbital integrals. It was originally our intent simply to demonstrate that this calculation is a feasible one. Since this work was completed, Calais' papers[17] appeared, using his energy expression and the method of "combined atomic orbitals" for computations (Papers II, III, IV). The CAO method utilizes the cubic symmetry of the interaction sums to reduce computation to the minimum required. Calais' calculation of the energy for the linear chain is numerically correct because in one dimension the choice of k's for generating AMO's and LAO's is obviously the lower half of the Brillouin zone whether we use symmetry or the energy-band

model Fermi surface to divide it. However, the calculation on bcc atomic hydrogen is formally incorrect and should be repeated with a correct set of $\mathbf{k}$'s.

There is no point in discussing practical computations further at present, in view of the extensive and efficient treatment given the problem by Calais (Paper III, Ref. 17). We are undertaking a similar approach in this laboratory. In Sec. V, however, we have gone further and have also obtained computationally feasible forms for the higher-order coefficients determining the splitting of different spin states from each other and from the single-determinant energy. The LAO basis set is very convenient for this purpose.

V. ENERGIES OF SPIN EIGENSTATES

In Sec. IV the energy expression associated with the single determinant T_0 was obtained, the argument being made that for the states of low multiplicity the energy of T_0 differs from the true eigenstate energies by terms of higher order in N^{-1}. We now obtain ex-

pressions for these higher order terms for states of spin s, for $s \ll N^{\frac{1}{2}}$. The case $s = N$ can also be solved but we do not examine it in detail. At the outset our procedure differs from that of Pauncz[18] only in minor details.

Consider the expectation value $\langle ^{2s+1}\Psi \mid \mathcal{L} \mid ^{2s+1}\Psi \rangle$, where $^{2s+1}\Psi$ is given by Eq. (25) and $\mathcal{L}$ is either the identity or the Hamiltonian. Since ^{2s+1}O defined by Eq. (24) is Hermitian, idempotent, and commutes with the symmetric operator $\mathcal{L}$,

$$\langle ^{2s+1}\Psi \mid \mathcal{L} \mid ^{2s+1}\Psi \rangle = \sum_k C(s;k)\langle T_0 \mid \mathcal{L} \mid T_k \rangle. \quad (46)$$

It is convenient to define one-electron and two-electron operators in H

$$F = \sum_{i=1}^{2N} h(i), \quad (47a)$$

$$G = \tfrac{1}{2}\sum_i \sum_{i \neq i} g(i,j). \quad (47b)$$

A straightforward calculation yields the results

$$\langle T_0 \mid T_k \rangle = (-1)^k z^{2k} \binom{N}{k}, \quad (48a)$$

$$\langle T_0 \mid F \mid T_k \rangle = (-1)^k z^{2k} \binom{N}{k} (\langle T_0 \mid F \mid T_0 \rangle + kf), \quad (48b)$$

$$\langle T_0 \mid G \mid T_k \rangle = (-1)^k z^{2k} \binom{N}{k} \left[\langle T_0 \mid G \mid T_0 \rangle + k(g_1 + g_2) + \frac{k(k-1)}{(N-1)}(g_3 - g_2) \right], \quad (48c)$$

where

$$f = (2\langle \varphi_0 \mid h \mid \bar{\varphi}_0 \rangle / z - \langle \varphi_0 \mid h \mid \varphi_0 \rangle - \langle \bar{\varphi}_0 \mid h \mid \bar{\varphi}_0 \rangle); \quad (49a)$$

$$g_1 = \sum_{j \text{ even}} \{ (1/z)[2(\langle \varphi_0\varphi_0 \mid g \mid \varphi_j\varphi_j \rangle + \langle \bar{\varphi}_0\bar{\varphi}_0 \mid g \mid \bar{\varphi}_j\bar{\varphi}_j \rangle) - (\langle \varphi_0\varphi_j \mid g \mid \bar{\varphi}_0\varphi_j \rangle + \langle \varphi_0\bar{\varphi}_j \mid g \mid \bar{\varphi}_0\bar{\varphi}_j \rangle)]$$

$$- 1[(\langle \varphi_0\varphi_0 \mid g \mid \varphi_j\varphi_j \rangle + \langle \bar{\varphi}_0\bar{\varphi}_0 \mid g \mid \bar{\varphi}_j\bar{\varphi}_j \rangle + 2\langle \varphi_0\varphi_0 \mid g \mid \bar{\varphi}_j\bar{\varphi}_j \rangle) - (\langle \varphi_0\varphi_j \mid g \mid \varphi_0\varphi_j \rangle + \langle \bar{\varphi}_0\bar{\varphi}_j \mid g \mid \bar{\varphi}_0\bar{\varphi}_j \rangle)]\}; \quad (49b)$$

$$g_2 = \sum_{j \text{ even}} [(1/z^2)\langle \varphi_0\bar{\varphi}_j \mid g \mid \varphi_0\bar{\varphi}_j \rangle - (1/z)(\langle \varphi_0\varphi_j \mid g \mid \varphi_0\bar{\varphi}_j \rangle + \langle \varphi_0\bar{\varphi}_j \mid g \mid \bar{\varphi}_0\bar{\varphi}_j \rangle) + (1)\langle \varphi_0\varphi_j \mid g \mid \bar{\varphi}_0\bar{\varphi}_j \rangle]; \quad (49c)$$

$$g_3 = \sum_{j \text{ even}} \{ (1/z^2)(2\langle \varphi_0\bar{\varphi}_0 \mid g \mid \varphi_j\bar{\varphi}_j \rangle - \langle \varphi_0\bar{\varphi}_j \mid g \mid \bar{\varphi}_0\varphi_j \rangle)$$

$$- (1/z)[2(\langle \varphi_0\varphi_0 \mid g \mid \varphi_j\varphi_j \rangle + \langle \varphi_0\bar{\varphi}_0 \mid g \mid \bar{\varphi}_j\bar{\varphi}_j \rangle) - (\langle \varphi_0\varphi_j \mid g \mid \bar{\varphi}_0\varphi_j \rangle + \langle \varphi_0\bar{\varphi}_j \mid g \mid \bar{\varphi}_0\bar{\varphi}_j \rangle)]$$

$$+ \tfrac{1}{2}[(\langle \varphi_0\varphi_0 \mid g \mid \varphi_j\varphi_j \rangle + \langle \bar{\varphi}_0\bar{\varphi}_0 \mid g \mid \bar{\varphi}_j\bar{\varphi}_j \rangle + 2\langle \varphi_0\varphi_0 \mid g \mid \bar{\varphi}_j\bar{\varphi}_j \rangle) - (\langle \varphi_0\varphi_j \mid g \mid \varphi_0\varphi_j \rangle + \langle \bar{\varphi}_0\bar{\varphi}_j \mid g \mid \bar{\varphi}_0\bar{\varphi}_j \rangle)]\}. \quad (49d)$$

Expanding in terms of ω and $\bar{\omega}$, we may group the terms as follows [the redundant symbols g and ω have been dropped, so that j means $\omega(\mathbf{j})$, $\bar{j}$ means $\bar{\omega}(\mathbf{j})$, etc.]

$$\eta_1 = f + g_1 + g_2$$

$$= (1 - z^2)\{ (2/z)\langle 0 \mid h \mid \bar{0} \rangle + \sum_{j \text{ even}} [4\langle 0\bar{0} \mid j\bar{j} \rangle + (2/z)(\langle 0\bar{0} \mid jj \rangle + \langle 0\bar{0} \mid \bar{j}\bar{j} \rangle)$$

$$+ \tfrac{1}{4}(\langle 0j \mid 0j \rangle + \langle \bar{0}\bar{j} \mid \bar{0}\bar{j} \rangle) - \tfrac{3}{2}\langle 0j \mid \bar{0}j \rangle - \tfrac{1}{2}\langle 0j \mid \bar{0}j \rangle + [(2 - 3z^2)/2z^2]\langle 0j \mid 0\bar{j} \rangle - z^{-1}(\langle 0j \mid \bar{0}j \rangle + \langle 0\bar{j} \mid \bar{0}\bar{j} \rangle)]\}; \quad (50a)$$

$$\eta_2 = g_3 - g_2$$

$$= [(1 - z^2)^2/z^2] \sum_{j \text{ even}} [2\langle 0\bar{0} \mid j\bar{j} \rangle - (\langle 0\bar{j} \mid 0j \rangle + \langle 0j \mid \bar{0}j \rangle)]. \quad (50b)$$

---

[18] R. Pauncz, J. Chem. Phys. **37**, 2739 (1962).

The appearance of $(1-z^2)$ reflects the fact that all these terms vanish when $\bar{\varphi}(\mathbf{j}) \rightarrow \varphi(\mathbf{j})$, as required by the Pauli principle. Equations (50) can be still further reduced by the use of Eqs. (36), (11), and (41) to remove the $\bar{\omega}(\mathbf{j})$ and obtain more convenient and rapidly convergent forms. It should be noted that η_1 and η_2 are of the order of magnitude of the energy per particle, at *most* $\sim 1/N$ of $\langle T_0 \mid H \mid T_0 \rangle$; also, that at $z \rightarrow 0$ they behave at worst as z^{-2}.

The energy $E^{(2s+1)}$, the expectation value $E^{(2s+1)} \equiv \langle ^{2s+1}\Psi \mid H \mid ^{2s+1}\Psi \rangle / \langle ^{2s+1}\Psi \mid ^{2s+1}\Psi \rangle$ can be written

$$E^{(2s+1)} = 2N\epsilon^{(2s+1)} = \langle T_0 \mid H \mid T_0 \rangle + \eta_1(z^2) y_1(s; z^2) + \eta_2(z^2) y_2(s; z^2), \tag{51}$$

where $y_1(s; z^2)$ and $y_2(s; z^2)$ can be derived from a fundamental quantity, $y_0(s; z^2)$ (the normalization integral) given by

$$y_0(s; t) = \sum_{k=0}^{N} (-1)^k \binom{N}{k} C(s; k) t^k; \tag{52a}$$

$$y_1(s; t) = t\{d[\ln y_0(s; t)]/dt\}: \tag{52b}$$

$$y_2(s; t) = [t^2/(N-1)][y_0^{-1} d^2 y_0(s; t)/dt^2],$$

$$y_2(s; t) = (N-1)^{-1}[t(dy_1/dt) + y_1(y_1 - 1)]. \tag{52c}$$

Thus all the energies are determined by $y_0(s; t)$, for which we can obtain certain expressions.

Though $C(s: k)$ is related to the Clebsch–Gordan coefficient[14] $(\frac{1}{2}N, m_1, \frac{1}{2}N, -m_1 \mid \frac{1}{2}N, \frac{1}{2}N, s, 0)$ with m_1 given by $(\frac{1}{2}N - k)$, it is most convenient to use the recursion formula of Löwdin.[15] If we define

$$A(s; k) = (-1)^k \binom{N}{k} C(s; k), \tag{53}$$

this recursion formula is

$$-(N-k)(k+1)(A_{k+1} - A_k) - s(s+1) A_k + (N-k+1) k(A_k - A_{k-1}) = 0. \tag{54}$$

If $M = N - 2k$ be used to express this formula, it is easy to show that $A(s; k)$ is a polynomial in M of order and parity s. Pauncz[18] gives explicit expressions for these polynomials up to $s = 5$. For our purposes, with $N \rightarrow \infty$, it is more useful to introduce the quasicontinuous variable $\mu = M/N$. By converting the difference equation to a differential equation and computing higher order corrections in N^{-1} we obtain

$$A(s; k) \equiv a(s; \mu) = P_s{}^0(\mu) - (N^{-1}) \sum_{\substack{l=2, \text{ even}}}^{(s)} (2s - 2l + 1) P_{s-l}{}^0(\mu)$$

$$+ \frac{1}{6N^2} \sum_{l=2, \text{ even}}^{(s)} (2s - 2l + 1)[2(l-2)(2s-l) + (s-l)(s-l+3)] P_{s-l}{}^0(\mu) + O(N^{-3}) \cdots, \tag{55}$$

where $P_j{}^0(\mu)$ is the jth Legendre polynomial. The sum over l (even) extends to s or $s-1$ depending on the parity of s. $O(N^{-3})$ vanishes for $s < 5$. The formula is valid when $s \ll N$ and is useful when $s \ll N^{\frac{1}{2}}$.

We can use it to compute $y_0(s; t)$ in these cases. Using the Euler–MacLaurin formula to evaluate Eq. (52a) as an integral, we obtain after some manipulation the *exact* result

$$y_0(s; t) = \frac{1}{2}[1 + (-1)^s t^N] a(s; 1) + \sum_{i=0}^{\infty} \left(\frac{2}{N}\right)^i (i!)^{-1}[1 + (-1)^{s-i+1} t^N] a^{(i)}(s; 1) \gamma^{(i)}(\xi), \tag{56}$$

where $a^{(i)}(s; 1) \equiv [d^{(i)} a(s; \mu)/d\mu^{(i)}]_{\mu=1}$, $\gamma^{(i)}(\xi)$ is the ith derivative of $\gamma(\xi) \equiv \frac{1}{2} \coth(\frac{1}{2}\xi)$, and $\xi = -\ln t$. This formula is useful if $(1-t)$ is $\gg N^{-1}$ (i.e., is finite) and if $s \ll N^{\frac{1}{2}}$. It can be simplified then to

$$y_0(s; t) = \frac{1}{1-t} \left\{ 1 - \frac{s}{2N} \left[(s-1) + \frac{t(s+1)}{1-t} \right] \right.$$

$$\left. + \frac{s(s-1)}{4N^2} \left[\frac{1}{6}(s-2)(3s-1) + (s-2)(s+1)\frac{t}{1-t} + (s+2)(s+1)\frac{t(1+t)}{(1-t)^2} \right] + O(N^{-3}) \cdots \right\}. \tag{57}$$

The resulting expressions for y_1 and y_2 (to order N^{-1}) are

$$y_1(s; t) = [t/(1-t)] - (1/N)[s(s+1)t/2(1-t)^2] + O(N^{-2}) \cdots, \tag{58a}$$

$$y_2(s; t) = (1/N)[2t^2/(1-t)] + O(N^{-2}) \cdots. \tag{58b}$$

Therefore, if $(1-t)$ is *finite* as $N \rightarrow \infty$, the separation of the states $s \ll N^{\frac{1}{2}}$ but not equal to 0 from the singlet is of

order N^{-2} in the energy per particle, and from the energy of T_0 it is of order N^{-1} in the energy per particle. This abnormally high density of states does not persist, of course, over any appreciable number of states compared to N. The limit $t = 1 - (\delta/N)$ is of some interest. For this regime we can prove that y_1 and y_2 behave differently and lead to a splitting of the states with $s \ll N^{\frac{1}{4}}$ from one another which is of order N^{-1} in the energy per particle. Again using the Euler–MacLaurin formula to evaluate (52a), we obtain in this case

$$y_0(s; t) = N\{\tfrac{1}{2}\exp(-\tfrac{1}{2}\delta)F_s(\tfrac{1}{2}\delta)\} + 1\{\tfrac{1}{2}[1+(-1)^s e^{-\delta}] - \tfrac{1}{8}\delta^2 \exp(-\tfrac{1}{2}\delta)F_s(\tfrac{1}{2}\delta)$$

$$+ \tfrac{1}{2}\exp(-\tfrac{1}{2}\delta)[\tfrac{1}{4}\delta^2 F'_s(\tfrac{1}{2}\delta) - \sum_{l=2,\text{ even}}^{(s)}(2s-2l+1)F_{s-l}(\tfrac{1}{2}\delta)]\} + O(N^{-1})\cdots, \quad (59)$$

where

$$F_s(x) = (2\pi/x)^{\frac{1}{2}}I_{s+\frac{1}{2}}(x) \tag{60}$$

($I_{s+\frac{1}{2}}$ is the Bessel function usually so denoted). Evaluating y_1 and y_2 in lowest order only, we find

$$y_1(s; t) = \tfrac{1}{2}N[1 - F'_s(\tfrac{1}{2}\delta)/F_s(\tfrac{1}{2}\delta)] + \cdots, \tag{61a}$$

$$y_2(s; t) = \tfrac{1}{4}N[1 - 2(F'_s/F_s) + (F''_s/F_s)] + \cdots. \tag{61b}$$

For the special case $s = 0$ the results are

$$y_1 = N\{[1-(1+\delta)e^{-\delta}]/\delta(1-e^{-\delta})\} - 1\{[1-e^{-\delta}(1+\delta+\tfrac{1}{2}\delta^2)]/(1-e^{-\delta})^2\} + \cdots, \tag{62a}$$

$$y_2 = 2N\{[1-e^{-\delta}(1+\delta+\tfrac{1}{2}\delta^2)]/\delta^2(1-e^{-\delta})\} + \cdots. \tag{62b}$$

Now, when $(1-t) = \delta/N$, η_1 behaves as (δ/N) and η_2 as $(\delta/N)^2$, so the net splitting of states with $s \ll N^{\frac{1}{4}}$ from one another is of order N^{-1} in the energy per particle; this is quite different from the correlated region, $(1-t)$ finite, corresponding to $(1-\lambda)$ finite.

The above results are all valid for $s \ll N^{\frac{1}{4}}$. Solutions can also be obtained for the special case $s = N$. $C(s, k) = 1$ in this case, so that for $(1-t)$ finite

$$y_0(N, t) = (1-t)^N, \tag{63a}$$

$$y_1(N, t) = -Nt/(1-t), \tag{63b}$$

$$y_2(N, t) = +N[t/(1-t)]^2. \tag{63c}$$

We see that the state $s = N$ is well separated from the single determinant in energy, by terms of order 1 in the energy per particle. Another useful application of Eqs. (63) is to check the accuracy of the expressions for η_1 and η_2 [Eqs. (50)], since the energy of the ferromagnetic state can easily be obtained from a single-determinant wavefunction with all spins parallel, and it is obviously independent of t, according to the Pauli principle.

As a last exercise, we have calculated the spin correlation between two lattice sites, **0** and **m**, where **m** can be even or odd. This is defined as the expectation value $\langle \mathbf{s}(\mathbf{0}) \cdot \mathbf{s}(\mathbf{m}) \rangle$, where the local spin operators $\mathbf{s}(\mathbf{j})$ are defined by projection with the Wannier-type orbital $\omega(\mathbf{j})$ on the density matrix (in this case for two particles); we obtain results as follows:

(a) **m** even

$$\langle \mathbf{s}(\mathbf{0}) \cdot \mathbf{s}(\mathbf{m}) \rangle = \tfrac{1}{4}\hbar^2(1-t)[1-(1/N)y_1(s; t)]; \tag{64a}$$

(b) **m** odd

$$\langle \mathbf{s}(\mathbf{0}) \cdot \mathbf{s}(\mathbf{m}) \rangle = \tfrac{1}{4}\hbar^2 \left\{ -[(1-t)+\tfrac{3}{2}tb_0^2(\mathbf{m})] \right.$$

$$+ N^{-1}(1-t)y_1(s; t)\left[1 - \frac{2(1-t)}{t} - \tfrac{7}{2}b_0^2(\mathbf{m})\right]$$

$$\left. + N^{-1}(1-t)^2 y_2(s; t)\left[2\frac{[1-b_0^2(\mathbf{m})]}{t}\right]\right\}. \tag{64b}$$

Substitution of correct values for y_1 and y_2 shows that for $s \ll N^{\frac{1}{4}}$ the correlation is $+\tfrac{1}{4}\hbar^2(1-t)$ for **m** even and $-\tfrac{1}{4}\hbar^2[(1-t)+\tfrac{3}{2}tb_0^2(\mathbf{m})]$ for **m** odd, i.e., there is long-range antiferromagnetic order nearly independent of spin and distance. For $s = N$ the correct ferromagnetic result of $+\tfrac{1}{4}\hbar^2$ is obtained for both even and odd sites **m**.

We may make certain suggestions based on these results about the energies of the various spin eigenstates in this model. Though in principle the optimum value of t can be determined separately for each s, for the states $s \ll N^{\frac{1}{4}}$ it is determined entirely by the single-determinant energy $\bar{\epsilon}$, as is the form of the optimum site orbital, $u(\mathbf{j})$. On the other hand, the energy of the ferromagnetic state is, as noted above, independent of t and depends only on the site-orbital variation. If we could assume that the optimum choice of $u(\mathbf{j})$ is the same for all the values of s, we can state that the signs of the coefficients η_1 and η_2 determine whether in this model the ferromagnetic state or the $^1\Gamma_1$ state is the ground state. We can be reasonably sure that η_2 is positive in any case, but the sign of η_1 is more difficult to establish. It is certainly always negative in the H_6 ring system, and it is not too un-

reasonable to suppose that it will remain so for all lattice spacings as $N \to \infty$, at least for hydrogenic "1s"-type orbitals. So, if we can conjecture that η_1 is negative, and if we assume a common site orbital for all the states, we find that the $^1\Gamma_1$ state and never the ferromagnetic state is the ground state. The ferromagnetic states lies above the ground state in this case by terms of order 1 in the binding energy per particle.

VI. DISCUSSION

The LAO representation suggests that the simple AMO (or "different bands for different spin") method for high correlation is a description of a solid in which *local* electron correlation plays some role. The relationship of this description to that of a Heitler–London or valence bond theory is an important question. In the valence bond method[6,12] an electron associated with a given site forms a bond pair with an electron on a near-neighbor site, and these bond pairings or bonds then "resonate" among the possible forms of equivalent energy. Any particular valence bond "structure" has an unique pairing of particular sites which are "bonded" in the structure; each such pair appears as a *pair singlet* in the wavefunction for the structure. In order to construct a $^1\Gamma_1$ state it is then necessary to consider a linear combination of all such possible structures; the symmetry of the interaction of a site with its environment is achieved by the interaction of the superposed wavefunctions of the structures. By contrast, the LAO wavefunction reflects from the beginning the total symmetry of the system, since the conjugate Wannier orbitals $\bar{\omega}(\mathbf{j})$ have full point symmetry about $\mathbf{j}$. The other significant difference is that there is *not* an unique pairing of the electron on $\varphi(\mathbf{j})$ with that on $\varphi(\mathbf{j})$, to form a pair singlet, although the components of spin *are* opposed. Indeed, the character of the singlet projection operator of Eq. (24) is that it weights all the possible spin pairings on an equal footing. An ingredient of the physical concept in the valence bond description is retained in the spatial correlation of $\omega(\mathbf{j})$ and $\bar{\omega}(\mathbf{j})$, but a pair singlet bond is not formed between them. In addition, $\bar{\omega}(\mathbf{j})$ is not *fully* localized at the sites close to $\mathbf{j}$, so that some "delocalization" is introduced.

It was our original hypothesis, based on results obtained for the H_6 ring system,[12] that the simple AMO method offers a similar description, even in an infinite system, to that given by a resonating valence bond model. However, there are two facts which suggest that this may not be the case: (1) The results of Calais' calculation on the linear chain (Ref. 17, Paper II) show that the system does not bind, in the simple AMO model. By contrast, we can be reasonably certain that a simple valence bond wavefunction representing bonds between pairs of near neighbors is binding, and the resonant combination of the two possible such structures will have an even lower energy. May we ex-

pect that in two or three dimensions this advantage of the valence bond model may persist? (2) In Sec. V we have shown that in the simple AMO ground state (i.e., $^1\Gamma_1$) there is a long-range antiferromagnetic order, while in a $^1\Gamma_1$ state based on valence bonds between near neighbors such pair correlations will persist only for short distances, a few neighbors. Thus we may suppose that increasing divergence between the two descriptions as $N \to \infty$ is due to the difference in spin pairing, the VB model selectively pairing only those electrons which are near each other in space. It is our belief that a valence bond description will turn out to be better than the simple AMO description, for the ground state of the sort of nonmetal system we consider.

The terms "localized" or "delocalized" in connection with electron behavior are, it should be noted, only meaningful in relation to the two-particle distribution, and our discussion using these terms must be taken in that sense. For the one-particle distribution it is irrelevant whether a $\mathbf{k}$ space or an $\mathbf{r}$ space description is emphasized.

The "AMO" model treated in this paper also differs in some important respects from the energy-band description. Although one-electron basis functions are employed, the eigenstates being considered are many-particle states. In the present work we have considered only the ground state and low-lying spin states derivable from the same "configuration."

The model is biased in favor of systems which are closer to "tight-binding" conditions than to "free-electron" conditions. In the energy-band theory, the most general Bloch function $\psi^{\mathrm{band}}(\mathbf{k})$ can be written

$$\psi^{\mathrm{band}}(\mathbf{k}) = [C(\mathbf{k})/(2n)^{\frac{1}{2}}] \sum_{\mathbf{j}} \exp(i\pi\mathbf{k}\cdot\mathbf{j}) u(\mathbf{k}; \mathbf{j}); \quad (65)$$

this form is similar to our Eq. (1) in appearance, except that u is now also a function of $\mathbf{k}$. However, this *implied* extra degree of freedom in Eq. (64) is *not real*, because it is implicitly understood that $u(\mathbf{k}; \mathbf{j})$ which appears in (64) is a $\mathbf{k}$-dependent linear combination of atomic $s, p, d, \cdots$, etc., type functions *associated with site* $\mathbf{j}$. Fourier's theorem, however, permits us to assert that this $\mathbf{k}$ dependence is equally well expressed if we *absorb* it in the *phase lags* associated with contributions to our $\mathbf{k}$-*independent* site orbitals $u(\mathbf{j})$ which come from atomic orbitals of type p, d, etc., $\cdots$, centered not on $\mathbf{j}$ but on various neighbors of $\mathbf{j}$. Thus the apparent $\mathbf{k}$ dependence in (64) is merely another way of describing the delocalized character in the site orbital $u(\mathbf{j})$ of Eq. (1), using in its place a site orbital with atomic components associated only with $\mathbf{j}$. However, our description is *a priori* a more natural one for tight binding, since for practical reasons we wish to consider site orbitals $u(\mathbf{j})$ which are reasonably well localized.

The major difference between the AMO model and the energy-band theory, however, is in the $\mathbf{k}$-space

population of the electrons. In the energy-band theory, k-space population is determined by the one-electron energy levels $\epsilon(\mathbf{k})$, these being doubly occupied up to the Fermi surface $\epsilon_F(\mathbf{k})$. In the simple AMO model, on the other hand, the set of Bloch-type orbitals $\psi(\mathbf{k})$ which are doubly occupied in the limit of low correlation $(t \to 1)$ is fixed by symmetry requirements rather than by minimization of the Hartree–Fock energy. It follows that the simple AMO theory is not a useful description in the low correlation region; the energy-band model has a lower energy. Only in the high-correlation domain where $1-t$ is finite can this model give an account of electronic structure. In this general perspective, as well as in the formal treatment, the present work differs from that of Calais,[17] which considers electron correlation in, say, an alkali metal as its goal.

However, alternant correlation of the sort treated probably plays some role even in the ground state of an alkali metal. Deep in the Fermi sea there are Bloch-type orbitals $\psi(\mathbf{k})$ which are doubly occupied both in the energy-band model and in the limit $t \to 1$ of the simple AMO model, and for which the $\psi(\bar{\mathbf{k}})$ for the appropriate conjugate vectors $\bar{\mathbf{k}}$ are completely empty. Alternant correlation can then occur for these $\mathbf{k}$'s, while for $\mathbf{k}$'s nearer the Fermi surface and the sublattice zone boundary a more complicated situation arises. Confirmation that effects of this type do occur in metallic solids is provided by the work of Overhauser[19] on spin-density wave (SDW) states. Overhauser proved that the ordinary (paramagnetic) Hartree–Fock state of a free-electron gas is unstable with respect to the formation of a state with a net static spin polarization wave, the wavelength being $\sim (2k_F)^{-1}$, where k_F is the radius of the Fermi sphere. In our model the analogous alternation is associated with the wave vector $\mathbf{K}$ of Eq. (7). In Overhauser's theory there is a definite static spin polarization at each point, while in our model there is only a correlation of polarizations; but this is simply the result of not projecting a spin eigenstate. The single determinant T_0 has a static spin polarization wave like that of an SDW state. Overhauser's calculations give a fairly small net spin polarization amplitude for the alkali metals. The formation of an SDW state is accompanied by a very slight deformation of the Fermi surface. The hypothesis of an SDW ground state has been used by Overhauser to explain the optical absorption spectrum of potassium. Overhauser's SDW approach to the effects of electron correlation is undoubtedly a better way to treat metallic systems than is the AMO model, which describes the limit of high correlation.

There is some relevance to the conjecture of Mott[9]

[19] A. W. Overhauser, Phys. Rev. **128**, 1437 (1962); Phys. Rev. Letters **13**, 190 (1964).

that there is a sharp change in the character of the electronic wavefunction at some critical density, and the work of des Cloiseaux on this question. While we are as yet unable to assess the validity of des Cloiseuax's rather drastic approximations, we feel that an accurate calculation would be valuable for a critical study of the problem. Since our model gives only information on the ground (i.e., $^1\Gamma_1$) state, we cannot yet describe the conducting states. des Cloiseaux's arguments that a gap will develop for the case of high correlation may well be qualitatively correct. However, study of the ground state may provide information about this possibility, in the behavior of the optimum value of t as a function of the lattice spacing, a. Although the simple AMO model is not a valid description of the metallic state, it does provide a model by which to study the onset of correlation effects. If $t(a)$ shows a rapid change at some critical distance a_c, we may perhaps conclude that Mott's conjecture is correct. For this and other reasons, actual computation of the energy coefficients is of great interest. Work on such computation is now in progress.

ACKNOWLEDGMENT

APPENDIX: FORMULAS FOR SPECIAL CASES

Case (1): One-Dimensional Lattice

Allowed values of j and nk are $0, \pm 1, \pm 2, \cdots,$ $\pm(n-1), n; \ (n=N)$;

$$b_0(m) = \frac{\sin(\frac{1}{2}\pi m)}{N \sin(\pi m/2N)} - \delta(m), \qquad (A1)$$

which yields as $N \to \infty$

$$b_0(m) = [2 \sin(\tfrac{1}{2}\pi m)/\pi m] - \delta(m). \qquad (A2)$$

Case (2): Square-Plane Lattice

$2N = 4n^2$, $\mathbf{j}$, and $n\mathbf{k}$ have integer Cartesian components as in Case (1). The sublattice zone is a square rotated at $45°$ from the Brillouin zone of the full lattice, bounded by the constraints $|k_x| + |k_y| = 1$:

$$b_0(\mathbf{m}) = 2(\cos\pi m_y - \cos\pi m_x)/\pi^2(m_x^2 - m_y^2) - \delta(\mathbf{m}).$$

$$(A3)$$

Case (3): Simple-Cubic Lattice

$2N = (2n)^3$; $\mathbf{j}$ and $n\mathbf{k}$ have integer Cartesian components as above. The Brillouin zone is a cube, the sublattice zone the "cuboctahedron" (first zone of an fcc lattice), bounded by the planes $|k_i| = 1$, $|k_x| + |k_y| + |k_z| = \frac{3}{2}$.

If the absolute values of all components are distinct,

$$b_0(\mathbf{m}) = -(2/\pi^3)\{1/(m_x^2-m_y^2)(m_y^2-m_z^2)(m_z^2-m_y^2)\}(b_x+b_y+b_z),\tag{A4a}$$

where

$$b_z = (m_x^2-m_y^2)[m_z \sin\tfrac{1}{2}(\pi m_z)(\cos\pi m_x + \cos\pi m_y) + \cos\tfrac{1}{2}(\pi m_z)(m_x \sin\pi m_x + m_y \sin\pi m_y)];\tag{A4b}$$

b_x and b_y can be obtained by cyclic permutations. Although the terms in $\sin\pi m_i$ always vanish they are needed for the derivation of Eqs. (A5) and (A6).

If $|m_x| = |m_y| \neq |m_z|$,

$$b_0(\mathbf{m}) = [2/\pi^3(m_z^2-m_x^2)^2]\{[2m_z \sin\tfrac{1}{2}(\pi m_z)\cos\pi m_x - m_x \sin\tfrac{1}{2}(\pi m_x)\cos\pi m_z$$
$$-m_x \sin\tfrac{1}{2}(3\pi m_x)]+[2m_x \cos\tfrac{1}{2}(\pi m_z)\sin\pi m_x - m_z \cos\tfrac{1}{2}(\pi m_x)\sin\pi m_z]\}$$
$$+[1/2\pi^2(m_z^2-m_x^2)]\{[\cos\tfrac{1}{2}(3\pi m_x)+(2/\pi m_x)\sin\tfrac{1}{2}(\pi m_z)(1-\cos\pi m_z)$$
$$-\cos\pi m_z \cos\tfrac{1}{2}(\pi m_x)]+[(m_z/m_x)\sin\tfrac{1}{2}(\pi m_x)\sin\pi m_z]\}.\tag{A5}$$

If $|m_x| = |m_y| = |m_z|$,

$$b_0(\mathbf{m}) = (3/8\pi m_z)\sin\tfrac{1}{2}(3\pi m_z)+(3/4\pi^2 m_z^2)\sin\pi m_z \sin\tfrac{1}{2}(\pi m_z)+(1/4\pi^3 m_z^3)\sin\tfrac{1}{2}(\pi m_z)(1-\cos\pi m_z)-\delta(\mathbf{m}).\tag{A6}$$

Electronic Structure of the First Excited State of CO. I. SCF Wavefunction Calculated in the Restricted Hartree–Fock Formalism*

WINIFRED M. HUO†

Department of Chemistry, Harvard University, Cambridge, Massachusetts

(Received 11 April 1966)

An SCF wavefunction calculated in the restricted Hartree–Fock formalism is reported for the lowest excited state ($a\ ^3\Pi$) of CO. The expansion method is used and exponents are optimized. Even though the calculated energy and Hellmann–Feynman forces are reasonable, there is a large discrepancy between the calculated and experimental dipole moment. The discrepancy is caused by the fact that the restricted Hartree–Fock function for open-shell systems does not satisfy Brillouin's theorem and first-order corrections come in. In particular, it is found that the charge-transfer states have a large first-order effect on the dipole moment. Some preliminary results on the isoconfigurational $A\ ^1\Pi$ state are reported. They indicate that the charge distributions of the two isoconfigurational states are quite different.

I. INTRODUCTION

THE calculation of SCF wavefunctions for the ground states of first-row diatomic molecules have been quite successful.[1,2] It has been found that if the basis set is sufficiently large and carefully optimized, the SCF wavefunction is a good approximation to the Hartree–Fock function. Both the calculated energy and one-electron molecular properties are in good agreement

with the experimental values. This good agreement is a consequence of Brillouin's theorem,[3] which states that the Hartree–Fock charge distribution is a representation of the exact charge distribution to first order.

The SCF wavefunction is, in general, a sum of anti-symmetrized products of one-electron orbitals $\phi_{i\lambda\alpha}$, where λ refers to the symmetry species to which the orbital belongs, α refers to the subspecies, and i labels orbitals not distinguishable by symmetry. The index pair $i\lambda$ represents a shell, which is composed of all orbitals belonging to the same i and λ. The orbitals in a closed shell form a complete degenerate set, but in an open shell the degenerate orbitals are only partially occupied. In molecular calculations, the one-electron orbitals generally used are symmetry orbitals. They transform under symmetry operations according to the

---

[1] W. M. Huo, J. Chem. Phys. **43**, 624 (1965).

[2] For example, see, (a) C. C. J. Roothaan, *et al.* Bull. Am. Phys. Soc. **8**, 535 (1963); (b) A. C. Wahl, J. Chem. Phys. **41**, 2600 (1964); (c) P. E. Cade, K. D. Sales, and A. C. Wahl, J. Chem. Phys. **44**, 1973 (1966); (d) P. E. Cade and W. M. Huo, "Hartree–Fock Wavefunctions and Energy Quantities for the Ground States of the First-Row Hydrides, AH," and "Certain Molecular Properties for the Ground States of the First-Row Hydrides, AH," *ibid.* (to be published).

[3] L. Brillouin, Actualités Sci. et Ind., No. 71 (1933); No. 159 (1934); C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).