

COMPUTATIONAL METHODS IN BAND THEORY

THE IBM RESEARCH SYMPOSIA SERIES

1971: Computational Methods in Band Theory

Other volumes in preparation

COMPUTATIONAL METHODS IN BAND THEORY

Proceedings of a Conference held at the
IBM Thomas J. Watson Research Center,
Yorktown Heights, New York, May 14-15, 1970, under the
joint sponsorship of IBM and the American Physical Society

Edited by

P. M. Marcus, J. F. Janak, and A. R. Williams
IBM Thomas J. Watson Research Center
Yorktown Heights, New York



PLENUM PRESS • NEW YORK-LONDON • 1971

Library of Congress Catalog Card Number 77-142039

ISBN-13: 978-1-4684-1892-7 e-ISBN-13: 978-1-4684-1890-3
DOI: 10.1007/978-1-4684-1890-3

© 1971 Plenum Press, New York
Softcover reprint of the hardcover 1st edition 1971

A Division of Plenum Publishing Corporation
227 West 17th Street, New York, N.Y. 10011

United Kingdom edition published by Plenum Press, London
A Division of Plenum Publishing Company, Ltd.
Donington House, 30 Norfolk Street, London W.C.2, England

All rights reserved

No part of this publication may be reproduced in any form
without written permission from the publisher

PREFACE

This volume contains the papers presented at the Conference on Computational Methods in Band Theory sponsored jointly by IBM and the American Physical Society and held at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York, on May 14-15, 1970.

The purpose of the conference was a sharing of information on the computational problems involved in relating models for the electron-electron and electron-ion interactions to experimentally measurable quantities. The papers comprising this volume therefore present up-to-date methodology for the calculation of single-particle energies and wave functions for periodic and near-periodic systems, the integration over these states required to describe experiment, and computationally practicable procedures for the introduction of exchange and correlation and the achievement of self-consistency.

The proceedings is actually an expansion of the conference in that, unlike the oral presentations, the papers were not limited as to length. Furthermore, time was allowed after the conference to permit the papers to be written with the conference in retrospect, and five "prepared discussion" papers written by attendees of the conference but not on the original program are included. The latter are indicated in the table of contents by asterisks.

The explicit emphasis of the conference on comparison of technique generated much lively argument, which is surely an indication of the current interest in the subject and the vigor of those working in it. It is our hope that the proceedings will make these comparisons available to the widest possible audience.

P. M. Marcus
J. F. Janak
A. R. Williams

CONTENTS

Welcoming Address	xiii
A. G. Anderson	
Comments on the Conference	xv
P. M. Marcus	
1. APW-OPW METHODS IN THE BAND PROBLEM	
A Comparison of Different Computer-Oriented Methods for the Energy Bands of Solids	3
John W. D. Connolly	
Diagonalization of Hermitian Matrices; Maximization of Speed and Accuracy	16
R. A. Faulkner	
An Alternative APW Technique: Theory and Application to Copper	25
Dale Dean Koelling	
An RAPW Expanded Basis Set	33
G. Arbman and D. Koelling	
New Version of the Modified Augmented-Plane-Wave Method	44
H. Bross, G. Bohn, G. Meister, W. Schubö, and H. Stöhr	
Gradients of $E(\vec{k})$ from the APW Determinant	59
J. H. Wood	
APW Pseudopotential Form Factors for the Alkali Metals	63
Martin J. G. Lee	

Self-Consistent Orthogonalized-Plane-Wave Calculations	82
R. N. Euwema, D. J. Stukel, and T. C. Collins	
Symmetrization Techniques in Relativistic OPW Energy Band Calculations	124
G. G. Wepfer, T. C. Collins, R. N. Euwema, and D. J. Stukel	
Some Notes on a Modified OPW Method	144
D. M. Gray and R. J. Karpien	
 2. KKR METHODS IN THE BAND PROBLEM	
Recent Developments in KKR Theory	157
A. R. Williams, S. M. Hu, and D. W. Jepsen	
*Comments on the KKR Wavefunctions; Extension of the Spherical Wave Expansion Beyond the Muffin Tins	178
O. Krogh Anderson	
Efficient Numerical Techniques for the Calculation of KKR Structure Constants	183
Harold L. Davis	
Calculations with "Non-Muffin-Tin" Potentials by the Green's Function Method	200
Benjamin Segall	
Phase Shift Parametrization: Band Structure of Silver	207
B. R. Cooper, E. L. Kreiger, and B. Segall	
Band Structure Calculations for Semiconductors and Insulators Using the KKR Method	218
H. Overhof	
Approximate KKR Band-Structure Schemes for Transition Metals	225
N. W. Dalton	
Optical Properties of the Alkalies Using the KKR-Z Method	260
A. R. Wilson, G. Dresselhaus, and C-Y. Young	

*Prepared Discussion

3. LCAO METHODS IN THE BAND PROBLEM

Discrete Variational Method for the Energy Band Problem	271
D. E. Ellis and G. S. Painter	
Discrete Variational Method for the Energy Band Problem with LCAO Basis and Non-spherical Local Potential	276
G. S. Painter and D. E. Ellis	
Recent Developments in Applying and Extending the Method of Tight Binding (LCAO) to Energy-Band Calculations	284
Earl E. Lafon, Roy C. Chaney, and Chun C. Lin	
Energy Bands by the LCAO Cellular Method	296
P. D. DeCicco	

4. INTERPOLATION AND INTEGRATION IN k-SPACE

Interpolation and k-Space Integration: A Review	305
F. M. Mueller	
Gilat-Raubenheimer Methods for k-Space Integration	323
J. F. Janak	
The Calculation of Brillouin Zone Integrals by Interpolation Techniques	340
R. L. Jacobs and D. Lipton	
Computational Method for Generalized Susceptibility	347
Joshua B. Diamond	
LCAO Interpolation Method for Nonorthogonal Orbitals	355
L. F. Mattheiss	
Interpolation in k-Space with Functions of Arbitrary Smoothness	362
Donn G. Shankland	
*Exact Solution of the Two-Band Density of States Problem	368
N. W. Ashcroft	

*Prepared Discussion

*Thermoelectric Transport Coefficients of Cubic Crystals Via k-Space Integration	373
J. M. Schoen	

5. THE BAND PROBLEM WITH BOUNDARY CONDITIONS

Bands, Bonds, and Boundaries	377
K. H. Johnson and F. C. Smith, Jr.	
$\vec{k} \cdot \vec{\pi}$ Interpolation and the Calculation of Vacancy States in PbTe	400
G. W. Pratt, Jr., E. K. Li, and F. J. Arlinghaus	
A KKR Method for Two-Dimensional Lattices and Its Application to Band Calculation	409
Kyozaburo Kambe	
The Propagation Matrix Method for the Band Problem with a Plane Boundary	416
D. W. Jepsen and P. M. Marcus	

6. SELF-CONSISTENCY, EXCHANGE, AND CORRELATION

The Self-Consistent Field Method for Crystals	447
J. C. Slater	
Approximations of the Exchange and Correlation Potentials	458
L. J. Sham	
*Some Remarks on Exchange Inhomogeneity Corrections in Many-Electron Systems	469
Irene B. Ortenburger and Frank Herman	
*Dielectric Function of Uniform Electron Gas	473
J. W. F. Woo	
Kohn-Sham Self-Consistent Scheme Applied to the Calculation of Atomic Systems and Metallic Sodium	476
B. Y. Tong	
A Potential Function for Band Structure Calculations . . .	489
David A. Liberman	
*Prepared Discussion	

CONTENTS	xi
Crystal Potentials Used in Energy Band Theory	495
A. M. Boring and E. C. Snow	
Single-Particle States in Many-Body Systems	500
Seb Doniach	
Towards Self-Consistency with the Tight Binding Approximation	512
Joseph Callaway and John L. Fry	
Toward Hartree-Fock Calculations for Simple Crystals . . .	517
Frank E. Harris and Hendrik J. Monkhorst	
The Use of the GI Method in Band Calculations on Solids	542
William A. Goddard III and Patricia M. O'Keefe	
List of Contributors	571
Subject Index	575

WELCOMING ADDRESS

CONFERENCE ON COMPUTATIONAL METHODS IN BAND THEORY

A. G. Anderson

IBM Vice President and Director of Research

Thomas J. Watson Research Center, Yorktown Heights, N. Y.

We are very pleased to be host to such an outstanding group of scientists and sophisticated users of computers.

Our earliest contacts with such scientific users extend back to times prior to World War II when IBM had early contact with W. Eckert of Columbia and others to consider what were then large problems in computation for astronomy. From that early base, contacts grew with Harvard, Princeton, and elsewhere around the world as, in 1945, the IBM Watson Scientific Computing Laboratory was established in New York City at Columbia University. That laboratory had profound impact upon computing - and computers - in the subsequent decade.

From those early beginnings, the IBM Research Division has evolved a substantial role as a user of computers. Such a role is in our own best interest and forms a vital part of the mission of the Research Division within IBM. Our user activities include large scale computation in scientific problems, laboratory automation, and time sharing.

A group in our West Coast San Jose Laboratory is engaged in work on quantum chemistry, studying problems ranging from atoms to biological molecules; much of this work is properly termed large scale computation. A recent calculation on hydrogen bonding in biological systems - guanine, cytosine - required 10^{11} integrals and took the equivalent of eight twenty-four hour days on the 360/195. This San Jose group, under Frank Herman, is now also engaged in the calculation of band structures and electronic properties of bulk matter - as is the Theoretical Physics Group,

under Paul Marcus, here in Yorktown. These interests in bulk matter are, of course, a primary source of motivation for sponsoring this conference.

Elsewhere in the Research Division there are large scale computation groups with broad interests in seismology, meteorology, astronomy, and solid state device applications. Beyond that we are avid computer users in laboratory automation and time sharing. In the latter case, to give you an idea of the present extent of this development, over 200 terminals and three different time sharing services in the Research Center here in Yorktown Heights. (I hope you will forgive the advertising; we strongly believe that science and engineering can be rapidly moved through efforts like these!)

This is the third topical conference in solid state physics organized by our Physical Sciences Department, and held at Yorktown Heights, in the past year. The previous ones were on "Current Instabilities in Semiconductors," and "Magnetic Semiconductors." We have found a two-day intensive conference of this kind on a specialized field to be a convenient and useful way to bring together workers and information in the field, and I'd like to express my best wishes for an equally successful conference in the present case.

COMMENTS ON THE CONFERENCE

INTRODUCTORY REMARKS

P. M. Marcus

IBM Thomas J. Watson Research Center

Yorktown Heights, New York

The focus of this Conference is on computational technique, which means we want to discuss matters frequently passed over lightly or swept under the rug in presentations of work in band theory, even though most of the research effort may go into developing that technique. Computational technique is, of course, for physicists only a tool for the achievement of physical understanding, but we hope by deliberately studying and improving this tool that we shall derive important benefits for the solution of the over-all research problem. It may be of interest that in my initial letter requesting Physical Society sponsorship for this Conference, I commented that increased efficiency in the vast computational activity in band theory might be a consequence of a Conference like this. In his letter to Edward Purcell, Chairman of the APS Committee on Meetings, recommending sponsorship of this Conference, Milan Fiske, Secretary of the Solid State Division of the APS, stated, "The subject appears timely, especially from the view of cost reduction." Perhaps we have an obligation to consider the terms of that recommendation, although I suspect that greater efficiency does not always lead to less computing.

Conversations with colleagues and scrutiny of material submitted to the Conference lead me to three general observations about the state of the field - perhaps to be modified in the light of what is said here:

1. I think we are now witnessing the impact on the field of the computer-trained generation; I am tempted to say of people who think like a computer, but perhaps it would be less controversial to say people who can take maximum advantage of a complete computer system - from the design of the numerical analysis to control of

auxiliary storage. In other words we can expect a professionalization of computational effort which may force some of us amateurs to reform our sloppy ways.

2. We are reaching the stage in band calculations where a global view of the problem is essential; after all we want to minimize the total time required to calculate a curve or, better, a set of curves of observable properties, not just the time for obtaining results at a single point in k space, or even the time for a set of energy bands. Hence efficient methods will consider the entire problem of filling and integrating over k space, and of economizing the calculation by making maximum use of any previously obtained information.

3. We are beginning to see use of band theory by experimentalists as a sophisticated way of fitting experimental data or of correlating data on different properties. Eventually this aspect of the field will belong to them, since they will know best what is significant about the data, and can design the best fitting procedures.

You will note from the program that the four topics suggested in the prospectus which so neatly divide up the field are by no means equally popular. Work on the classical band problem dominates by far, followed by substantial activity in the theory of self-consistency and the exchange correlation correction. Contributions to the interpolation and integration problem are rather fewer, and to the band problem with boundary conditions much fewer.

The keynote session which opens the Conference consists of two talks which address the two topics of major interest. We shall hear first from a man who occupies a unique place in our field, whose massive contributions to band theory and many related fields have greatly benefited us all, whose students and students' students and further academic generations together with his many associates in research form a substantial part of this audience. It is indeed a privilege to welcome Professor John Clarke Slater on his first visit to this laboratory. The first talk will be given by Professor Slater, to be followed by a talk by his outstanding former student, John Connolly.

* The talks by J. C. Slater and J.W.D. Connolly appear in this Proceeding in Sections 6 and 1 respectively.

1. APW-OPW METHODS IN THE BAND PROBLEM

A COMPARISON OF DIFFERENT COMPUTER-ORIENTED METHODS
FOR THE ENERGY BANDS OF SOLIDS

John W. D. Connolly

Pratt and Whitney Aircraft, Advanced Materials

Research and Development Laboratory, Middletown,
Connecticut

The computation of energy bands in solids is a two part problem. The first part consists of the generation of a suitable one-electron Hamiltonian, i.e. a periodic potential, and the difficulties associated with this aspect of the problem have been nicely described by Professor Slater in the preceding talk (1). This talk will describe the second part of the problem, namely the solution of the Schrödinger equation defined by the periodic potential (2).

The nature of the potential has a great deal to do with determining the type of solution. Near each lattice site, it is spherically symmetric and atomic-like. Whereas, in the interstitial regions it joins smoothly with the potentials around other sites, and tends to be relatively flat. This dual nature of the potential has given rise to two basic approaches to the solution of the equations, which are roughly related to the two historical ways of viewing a solid, either as an electron gas or as a collection of atoms. These two types of approaches can be termed as (1) the plane wave representation methods and (2) the partial wave representation methods. The first type sets up a suitable wave function in the interstitial region (i.e., a sum of plane waves which correctly satisfies the periodic boundary conditions) and corrects it in the inner atomic region with some kind of summation over angular momenta. The second type of method starts with a spherical wave function, i.e. a sum over angular momenta (partial waves) in the inner region and then adds subsidiary restrictions in order to satisfy the outer boundary conditions. A plane wave type method thus results in a secular equation whose size is determined by the

plane wave summation, i.e. the number of reciprocal lattice vectors necessary to represent the wave function, and whose matrix elements involve an internal angular momentum summation. This is the case for the orthogonalized plane wave (OPW)(3,4) and the augmented plane wave (APW) methods (5,6). A partial wave type method is just the reverse; the size of the secular equation is determined by the number of angular momenta necessary to represent the wave function whereas the sum over reciprocal lattice vectors is internally contained in the matrix elements. The three principal examples of the partial wave methods are the Korringa-Kohn-Rostoker (KKR) method (7,8), the cellular method (9) and the linear combination of atomic orbitals (LCAO) method (12,13). Traditionally, a plane wave representation has been considered most suitable for the free-electron type of solid, and a partial wave representation for those solids with more localized electrons. However, the results of many calculations in recent years seem to indicate that any of the above methods can be extended (or modified) to handle practically all types of solids. Up to the present, most reliable band calculations have been the result of applications of OPW, APW and KKR methods, and these are the ones I will concentrate on. The discussion of the LCAO method, which is currently being revived, will be taken up by subsequent speakers.

COMPARISON OF THE METHODS

The methods of the first type, which are derived from a plane wave expansion, all have secular equations of the form (for one atom per unit cell);

$$\det \left| \{(\vec{k} + \vec{g})^2 - E\} \delta_{\vec{g}, \vec{g}'} + F_{\vec{g}, \vec{g}'} (\vec{k}, E) \right| = 0 \quad (1)$$

where \vec{g} and \vec{g}' represent reciprocal lattice vectors, \vec{k} is the reduced vector and E is the energy. The F -function, in the particular case of the OPW method, has the form (with an internal summation over the ℓ -values of the atomic core),

$$F_{\vec{g}, \vec{g}'}^{OPW} = \tilde{V}_{\vec{g}-\vec{g}'} - \sum_{\alpha} E_{\alpha} \langle \vec{k} + \vec{g} | \alpha \rangle \langle \alpha | \vec{k} + \vec{g}' \rangle \quad (2)$$

where \tilde{V} is the Fourier transform of the potential and $\langle \vec{k} + \vec{g} | \alpha \rangle$ represents an overlap integral between a plane wave of the form $\exp[i(\vec{k} + \vec{g}) \cdot \vec{r}]$ and an (atomic) core wave function of energy E_{α} . The potential V is not restricted to any particular form, but in practice it is usually taken to be a sum of overlapping spherical potentials. This is still a more general form than the muffin-tin form often used in other methods. The method is evidently slowly convergent, requiring on the order of several hundred plane waves

to obtain accurate energy eigenvalues according to the calculations of Stukel et al (10). In the (local) pseudopotential (PP) which is a common approximation to the OPW method, $\vec{F}_{\vec{g} \vec{g}'}^{\text{OPW}}$, is assumed to be independent of \vec{k} and E ,

$$\vec{F}_{\vec{g} \vec{g}'}^{\text{OPW}} \simeq \tilde{V}_{\text{eff}} (|\vec{g} - \vec{g}'|) \quad (3)$$

This is a convenient device for empirical purposes, but there is no theoretical justification for this approximation. The more general $\vec{F}_{\vec{g} \vec{g}'}$ is often called a pseudopotential, and then referred to as non-local, k or E dependent.

The unmodified OPW method has difficulty in describing solids whose electrons cannot be described as either free-electron-like or localized core states (as in those solids which involve transition elements). However, it can be modified, as first suggested by Callaway (11), by the addition of Bloch sums of atomic-like functions which have been cut off to eliminate overlap and thus able to describe solids of this type. Recent calculations on Niobium by Deegan and Twose (12) and Copper by Butler, Bloom and Brown (13) by means of the modified OPW method appear to be convergent, since they are in agreement with the APW calculations of Mattheiss (14) and Burdick (15).

We now go on to compare with the APW and KKR methods. The "pseudopotential" matrix elements for the APW method are (in the particular case of a muffin-tin potential),

$$\begin{aligned} \vec{F}_{\vec{g} \vec{g}'}^{\text{APW}} &= \frac{4\pi R^2}{\Omega} \left\{ -[(\vec{k} + \vec{g}')^2 - E] \frac{j_1(|\vec{g} - \vec{g}'|R)}{|\vec{g} - \vec{g}'|R} + \sum_{\ell=0}^{\infty} (2\ell + 1) \right. \\ &\quad \times P_{\ell}(\vartheta_{\vec{g} \vec{g}'}) j_{\ell}(|\vec{k} + \vec{g}'|R) j_{\ell}(|\vec{k} + \vec{g}'|R) \\ &\quad \times \left. \left[\frac{u'_{\ell}(R, E)}{u_{\ell}(R, E)} - \frac{j'_{\ell}(|\vec{k} + \vec{g}'|R)}{j_{\ell}(|\vec{k} + \vec{g}'|R)} \right] \right\} \end{aligned} \quad (4)$$

where R is the radius of the muffin-tin sphere, Ω is the volume of the unit cell, and $u_{\ell}(r, E)$ is the radial wave function inside the sphere. P_{ℓ} and j_{ℓ} are Legendre polynomials and spherical Bessel functions which arise from the matching condition of a plane wave at the sphere radius.

The plane wave convergence of the APW method is faster than for the OPW method, probably because the wave function is more realistic in the inner regions. A large number of plane waves is usually necessary to describe an atomic-like function, and the partial wave expansion with which an APW basis function is "augmented" in the inner regions can describe one much more easily.

Although the KKR method is usually expressed as a partial wave expansion method, it can be transformed (16,17), as Ziman has shown, to a plane wave representation form, and the resulting pseudopotential is remarkably similar to that for the APW method.

$$\begin{aligned} F_{\vec{g}}^{\text{KKRZ}} &= \frac{4\pi R}{\Omega}^2 \sum_{l=0}^{\infty} (2l+1) P_l(\vartheta_{\vec{g}}) j_l(|\vec{k} + \vec{g}|R) j_l(|\vec{k} + \vec{g}'|R) \\ &\times \left[\frac{u'_l(R, E)}{u_l(R, E)} - \frac{j'_l(\sqrt{ER})}{j_l(\sqrt{ER})} \right] \end{aligned} \quad (5)$$

The only difference here is in the terms involving the Bessel functions j_l . In the APW method they are \vec{k} -dependent whereas in the KKR method, they are energy-dependent. This leads to better l -convergence for the latter method, since for the higher l -values, there is a cancellation of terms,

$$\frac{u'_l(R, E)}{u_l(R, E)} \approx \frac{j'_l(\sqrt{ER})}{j_l(\sqrt{ER})} \quad l > l_{\text{valence}}$$

The slow l -convergence of the APW method arises from the requirement that the interior sum of spherical functions be continuous with a plane wave at the sphere radius and for the higher plane waves this requires more and more l -values. There have recently been suggested, as a way around this problem, some modifications to the APW basis functions, by Kleinman and Shurtleff (18) and DeCicco (19) which may effectively accelerate this convergence.

The plane wave convergence, on the other hand, is not as good in the KKRZ method as in the APW method according to calculations by Wood (20). This is more serious since the number of plane waves determines the secular equation size. This slower convergence is consistent with a theorem due to Johnson (21) which shows that the APW method achieves the optimum convergence for the wave function in plane wave form.

The KKR method is, of course, more suitably handled in its partial wave form. The secular equation in this case can be expressed in the form (9),

$$\det \left| \sqrt{E} \delta_{ll}, \delta_{mm}, + \tan \eta_l \sum_{LM} c_{lm; l'm'}^{LM} D_{LM}(\vec{k}, E) \right| = 0 \quad (6)$$

η_l is a scattering phase shift related to the radial function u_l inside the muffin-tin sphere. $c_{lm; l'm'}^{LM}$ are the Gaunt coefficients, so that the internal LM summation is finite. D_{LM} are the KKR structure factors, which are the coefficients in a partial wave expansion of the one-electron Green's function;

$$G_E(\vec{R}) = -\frac{1}{\Omega} \sum_{\vec{g}} \frac{e^{i(\vec{k} + \vec{g}) \cdot \vec{R}}}{(\vec{k} + \vec{g})^2 - E} \equiv -\frac{\cos \sqrt{E} R}{4\pi R} + \sum_{LM} i^L D_{LM}(\vec{k}, E) j_L(\sqrt{E} R) Y_{LM}(\hat{R}) \quad (7)$$

Since this secular equation is equivalent to the KKRZ form, the convergences in angular momentum and plane waves are the same. However, in this form, the poor plane wave convergence (which now is contained in the expression for D_{LM}) can be obviated by the use of a Ewald summation technique. This procedure divides the structure factors into two sums, one over reciprocal space, the other over direct space, in a way such that both become highly convergent.

The good convergence in l leads to a smaller secular equation for the KKR method than that of the other methods. However, one pays the price for this smaller secular equation by having more complicated matrix elements. It turns out that the limiting factor on the computation time for the KKR method is the evaluation of the many thousands of structure factors necessary for a standard energy band calculation. Each D_{LM} involves long summations over exponential functions, spherical harmonics and integrals which must be evaluated numerically.

This complication is an example of a general "conservation of effort" rule, i.e. that a computational method which has a smaller secular equation has more complicated matrix elements and vice versa. Although, it is difficult to compare computer programs, it is found empirically when one compares the APW and KKR methods, that equivalent calculations take roughly equivalent amounts of computer time.

TABLE I

A Comparison of Several Energy Band Calculations on Face-Centered Cubic Copper

$E_d = X_5 - X_1$ is a measure of the d-band width, $E_{sp} = X_4^+ - \Gamma_1$ is a measure of the sp-band width and $E_{sd} = \Gamma_{25}^- - \Gamma_1$ is the position of the d-band with respect to the sp-band. All energies are in Rydbergs.

Author	Ref.	E_d	E_{sp}	E_{sd}	Potential	Secular Equation Size	Method
Burdick	15	.249	.804	.399	Chodorow ^a	27	APW
Arlinghaus	22	.2516	.8071	.4027	Chodorow ^a	59	APW
Arlinghaus	22	.2276	.7938	.3884	Mattheiss ^b	59	APW
Snow & Waber	23	.191	.717	.216	SC/XS ^c		APW
Snow	24	.224	.794	.375	SC/X $\alpha = 5/6$		APW
Segall	25	.300	.807	.331	ℓ - dependent	9	KKR
Wakoh	26	.245		.386	SC/XS ^c	9	KKR
Davis et al	27	.2538	.7995	.4654	Mattheiss ^b	9	KKR
Butler et al	28	.254	.808	.398	Chodorow ^a	24 ^d	MOPW
Altmann et al	29	.2511	.8084	.4031	Chodorow ^a	98	CELLULAR
Mueller & Phillips	30	.242	.800	.396	Fit to Burdick	9	Emp. LCAO/PP
Dresselhaus	31			.489	Fit to FS & Optical Data	9	Emp. LCAO
Fong & Cohen	32	.242	.77	.369	Fit to Optical Data	140	Emp. PP

^aPotential generated from a Cu⁺ Hartree-Fock calculation according to the prescription of M. I. Chodorow, M. I. T. Ph.D Thesis, 1939.

^bA superimposed atom potential generated by the prescription of L. F. Mattheiss, Phys. Rev. 133, A184 (1964).

^cSelf-consistent using the Slater exchange.

^dThis is the number of symmetrized OPW's. The number of non-symmetrized OPW's required in Nb (Ref. 12) was ~ 200 for a convergence of .01 Ry. A similar number should be required for Cu.

COPPER CALCULATIONS BY MANY METHODS

Table I shows a comparison of a representative sampling of the many calculations for face-centered cubic copper that have been published in the last decade. In the table, each calculation is characterized by three variables, the d-band width E_d , the sp-band (conduction band) width E_{sp} and the relative positions of the two bands E_{sd} . Calculations which have these three variables equal tend to be identical in all other respects. We note that all the calculations have roughly equal conduction band widths since these levels are close to the free electron values. Through comparison with the optical and photoemission experimental data, it appears that the calculation due to Snow (24), which is a self-consistent calculation using an exchange factor intermediate between the Slater and Kohn-Sham-Gaspar values (34), has the best agreement with experiment. With a few exceptions (23,27,31), most of the other calculations agree with this one to within a few hundredths of a Rydberg. The main point of this table is that all methods, including the OPW and empirical pseudopotential methods, are capable of adequately describing the energy band structure of a material like Cu which contains both localized and non-localized electrons. Also included here is the recent calculation by Altmann (29) using the supposedly obsolete cellular method, which has now been shown to be at least as accurate as the other methods. About the only major method not represented here is the ab initio LCAO method. However, a recent paper by Callaway and Zhang (33) has shown that this method can be applied to Nickel, whose energy band structure is very similar to that of copper.

MODIFICATIONS OF THE METHODS

Up to this point, what I have described are the bare skeletons of each of the major methods of band theory. Modifications to the methods are, of course, numerous. One of the most important of these is the extension to more general potentials than the muffin-tin or spherical forms, and how easily this can be done is one criterion for the usefulness of a method. This is of concern mainly to those who use the APW and KKR methods. The OPW method already uses a potential which is a more general form than a muffin-tin, (i.e., a sum of overlapping spherical potentials). As far as I know, there has been no attempt to generalize the OPW method to a sum of overlapping non-spherical potentials, and for certain types of solids this may eventually become necessary. Before listing the various modifications, it should be mentioned that it is sometimes possible to do quite well without going beyond the muffin-tin potential, even on covalent materials which would be expected to be quite non-spherical. This has been shown to be true in calculations on diamond by Keown (35), on boron nitride by Wiff and Keown (36) and on some II-VI compounds by Treusch, Eckelt and Madelung (37).

For the APW method, the simplest method of generalizing the potential is by assuming a non-constant potential outside the spheres about each lattice site. This is sometimes referred to as the "warped muffin-tin" (38,39). Since an APW basis function is simply a plane wave outside the spheres, the additional matrix elements consist merely of the Fourier components of this non-constant potential, which complicates the method hardly at all. A more severe modification arises from the non-spherical corrections inside the spheres. These can be calculated by evaluating the many extra matrix elements of a harmonic expansion of the potential between the inner spherical functions, as shown in the recent Lithium calculation by Rudge (40). An alternative to this procedure is to modify the APW basis functions to be non-spherical. This leads to the complication of m -dependent radial functions satisfying coupled differential equations, but has the effect of somewhat simplifying the extra matrix elements (18,41).

There are two other possibilities suggested by workers in the KKR method. One of these is the ordinary perturbation theory, which was first shown to be a practical procedure by Ham and Segall (42). The other is by assuming, as in the OPW method, a potential which is the sum of overlapping spheres. Williams (43) has shown that to first order of approximation, the formalism of the KKR method does not change when the potential is generalized to this form, and, therefore, can be applied to this more general form of potential without major changes in the computer programs.

Another criterion for the usefulness of a particular method of band theory is whether it lends itself easily to empirical modifications, as a convenient device for the interpretation of experimental data. Examples of this type of procedure would be the core shifts in the OPW method (44) or the change in the intersphere potential in the APW method (35), both of which allow one to adjust a first principles band structure through the use of only one parameter. For a more complete parameterization of an energy band structure, the empirical LCAO method (45) has turned out to be useful for a wide range of solids, including covalent materials like Germanium and Silicon (46) and transition metals like Iron and Chromium (47).

One of the most popular empirical schemes is, of course, the pseudopotential, either alone or in combination with the tight binding approximation (48,49) and these schemes have been extremely useful when used to fit experimental data or ab initio calculations. However, there is a dangerous temptation to try to use it as an ab initio method, and I would like to present here an example which shows the possible error involved in using a pseudopotential which has not been empirically derived. This figure shows the local pseudopotential calculated by Animalu and Heine (50) for Aluminum, which one would expect to be well characterized by the pseudopotential.

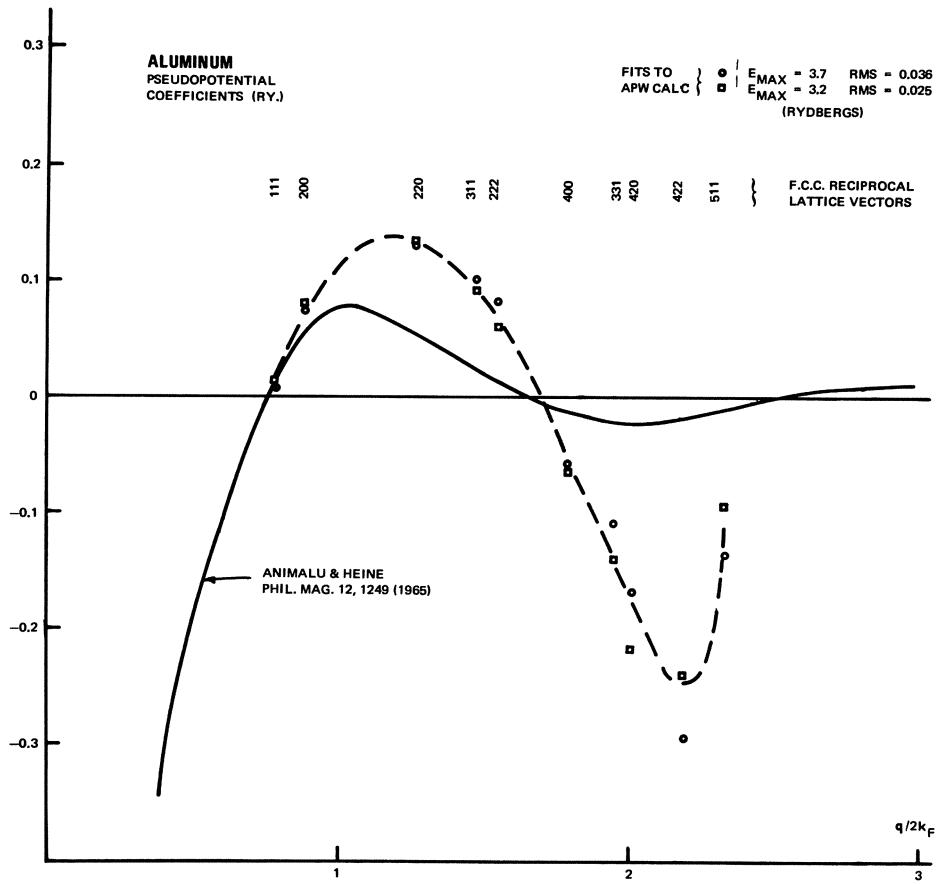


Figure 1. The pseudopotential coefficients obtained from a least squares fit to an APW calculation for Al. RMS is the root-mean-square deviation of the fit up to the energy E_{MAX}.

approximation since it is a "free-electron" like metal. As a comparison, the figure also shows the local pseudopotential components determined from a least squares fit to an APW calculation (51). As can be seen, the deviation for the first two components (which determine the Fermi surface) is not too large, but for the higher components, the Animanu and Heine pseudopotential is much too weak. As a result, Aluminum is much less free-electron like than one might think, and this is confirmed experimentally by the soft x-ray absorption and electron reflection data (51).

CONCLUSIONS

In conclusion, the main point I wish to make is that any of the methods of band theory mentioned here can be used successfully on all types of solids, and there is no reason why any of them cannot be extended to arbitrarily large periodic systems. However, if one wishes to determine the electronic structure of more general (non-periodic) systems, such as surfaces, impurities or other localized excitations, there is probably an advantage to using one of the partial wave type of methods. Plane wave expansions are peculiarly suitable to periodic systems, whereas the partial wave expansions are more flexible and, therefore, more easily generalized. Of these, the KKR method appears to be the method which can be extended most conveniently to non-periodic systems, and it has been shown recently by Johnson and others to be applicable to isolated defects, clusters of impurities and surfaces (52,53). These extensions will be the subject of a paper by Johnson at this conference (54).

REFERENCES

1. J. C. Slater, preceding paper, (1970). (See Section 6)
2. J. C. Slater, in Methods in Computational Physics, Vol. 8, Energy Bands of Solids, (edited by B. Alder, S. Fernbach and M. Rotenberg, Academic Press, New York, 1968), p. 1.
3. L. M. Falicov in Metallurgical Society Conferences, Vol. 45, Energy Bands in Metals and Alloys, Gordon and Breach, 1968, p. 73.
4. F. Herman, R. L. Kortum, C. D. Kuglin and J. P. Van Dyke, Methods in Computational Physics, Vol. 8, (1968), p. 193.
5. J. C. Slater, Phys. Rev. 51, 846 (1937).
6. L. F. Mattheiss, J. H. Wood and A. C. Switendick, Methods in Computational Physics, Vol. 8, (1968), p. 63.

7. J. Korringa, *Physica* 13, 392 (1946); W. Kohn and N. Rostoker, *Phys. Rev.* 94, 1111 (1954).
8. B. Segall and F. S. Ham, *Methods in Computational Physics*, Vol. 8, (1968), p. 251.
9. S. L. Altmann, *Proc. Roy. Soc. A* 244, 141, 153 (1958); S. L. Altmann and C. J. Bradley, *Proc. Phys. Soc.* 86, 915 (1965).
10. D. J. Stukel, R. N. Euwema, T. C. Collins, F. Herman and R. L. Kortum, *Phys. Rev.* 179, 740 (1969).
11. J. Callaway, *Phys. Rev.* 99, 500 (1955).
12. R. A. Deegan and W. D. Twose, *Phys. Rev.* 164, 993 (1967).
13. F. A. Butler, F. K. Bloom and E. Brown, *Phys. Rev.* 180, 744 (1969).
14. L. F. Mattheiss, *Phys. Rev.* B1, 373 (1970).
15. G. A. Burdick, *Phys. Rev.* 129, 138 (1963).
16. J. M. Ziman, *Proc. Phys. Soc.* 86, 337 (1965).
17. J. C. Slater, *Phys. Rev.* 145, 599 (1966).
18. L. Kleinman and R. Shurtleff, *Phys. Rev.* 188, 1111 (1969).
19. P. D. DeCicco, M. I. T. Solid State and Molecular Theory Group Quarterly Progress Report No. 65, p. 53, (1967), unpublished.
20. See Appendix C of Reference 6.
21. K. H. Johnson, *Phys. Rev.* 150, 429 (1966).
22. F. J. Arlinghaus, Ph.D. Thesis, Department of Chemistry, M. I. T., 1965, unpublished.
23. E. C. Snow and J. T. Waber, *Phys. Rev.* 157, 570 (1967).
24. E. C. Snow, *Phys. Rev.* 171, 785 (1968).
25. B. Segall, *Phys. Rev.* 125, 109 (1962).
26. S. Wakoh, *J. Phys. Soc. Japan* 20, 1894 (1965).
27. H. L. Davis, J. S. Faulkner and H. W. Joy, *Phys. Rev.* 167, 601 (1968).

28. F. A. Butler, F. K. Bloom and E. Brown, Phys. Rev. 180, 744 (1969).
29. S. L. Altmann, B. L. Davies and A. R. Hartford, J. Phys. C. 1, 1633 (1968).
30. F. M. Mueller and J. C. Phillips, Phys. Rev. 157, 600 (1969).
31. G. Dresselhaus, Solid State Commun. 7, 419 (1969).
32. C. Y. Fong and M. L. Cohen, Phys. Rev. Letters 24, 306 (1970).
33. J. Callaway and H. M. Zhang, Phys. Rev. B1, 305 (1970).
34. See Ref. 1 for an explanation of these terms.
35. R. Keown, Phys. Rev. 150, 568 (1966).
36. D. R. Wiff and R. Keown, J. Chem. Phys. 47, 3113 (1967).
37. J. Treusch, P. Eckelt and O. Madelung in Semiconducting II-VI Compounds, edited by D. G. Thomas, (Benjamin 1967), p. 588.
38. P. D. DeCicco, Phys. Rev. 153, 931 (1967).
39. D. D. Koelling, Phys. Rev. 188, 1049 (1969).
40. W. E. Rudge, Phys. Rev. 181, 1024 (1969).
41. J. W. D. Connolly, M. I. T. Solid State and Molecular Theory Group Semi-Annual Progress Report No. 72 (1970), p. 26, unpublished.
42. F. S. Ham and B. Segall, Phys. Rev. 124, 1786 (1961).
43. A. R. Williams, Phys. Rev. B1 (1970, in press).
44. F. Herman, R. L. Kortum, C. D. Kuglin and J. P. Van Dyke in Metallurgical Society Conferences, Vol. 45, Energy Bands in Solids, Gordon and Breach, 1968, p. 19.
45. J. C. Slater and G. F. Koster, Phys. Rev. 94, 1498 (1954).
46. G. Dresselhaus and M. S. Dresselhaus, Phys. Rev. 160, 649 (1967).
47. J. W. D. Connolly, Proceedings of the N. B. S. Conference on the Electronic Density of States (1969, in press).

48. L. Hodges, H. Ehrenreich and N. D. Lang, Phys. Rev. 152, 505 (1966).
49. F. M. Mueller, Phys. Rev. 153, 659 (1967).
50. A. O. E. Animalu and V. Heine, Phil. Mag. 12, 1249 (1965); W. A. Harrison, Pseudopotentials in the Theory of Metals (W. A. Benjamin, Inc., New York, 1966) Table 8-4, p. 309.
51. J. W. D. Connolly, Int. J. Quantum Chem. 3S, 807 (1970).
52. K. H. Johnson, Int. J. Quantum Chem. 2S, 233 (1968).
53. K. H. Johnson and J. W. D. Connolly, Physics Letters 28A, 291 (1968).
54. K. H. Johnson and F. C. Smith, Jr., "Bands, Bonds and Boundaries," (See Section 5)

DIAGONALIZATION OF HERMITIAN MATRICES;
MAXIMIZATION OF SPEED AND ACCURACY

R. A. Faulkner

Bell Telephone Laboratories, Incorporated

Murray Hill, New Jersey 07974

In this paper I describe an algorithm for finding all of the eigenvalues and eigenvectors of an Hermitian matrix (either complex or real, symmetric). The algorithm is a combination of Householder reduction to tri-diagonal form and a modified QR method for obtaining both eigenvalues and eigenvectors of the reduced matrix. The central new practical technique communicated here lies in the modification of the QR method when generating eigenvectors along with the eigenvalues.

As a practical procedure for diagonalization of general Hermitian matrices on a digital computer, this method is unsurpassed in the accuracy of the computed results and is as fast as the current state of the art allows. Its speed can be surpassed only by sacrificing guaranteed accuracy, but the fastest known algorithm cannot so much as double the speed.

The principal work to which I shall refer is the classic book by Wilkinson(1). I urge all who would learn of the eigenvalue problem to study it carefully. Every numerical algorithm mentioned in this paper is described there in full detail.

THE EIGENVALUE PROBLEM

Given an Hermitian matrix, A , the eigenvalue problem seeks a diagonal matrix, Λ , and a unitary

matrix, B , such that

$$A B = B \Lambda. \quad (1)$$

The diagonal elements of Λ are the eigenvalues of A and the columns of B are the eigenvectors of A . Equation (1) may be rewritten to emphasize that B "diagonalizes" A ;

$$B^* A B = \Lambda. \quad (2)$$

The asterisk is used to denote the complex conjugate transpose of a matrix.

BACKGROUND NOTES

Historically, the Jacobi method was the first truly successful method used to diagonalize Hermitian matrices on digital computers. It proceeds by applying a sequence of plane rotations to the matrix until the off-diagonal elements have been reduced below an assigned error limit.

$$A_0 = A \quad (\text{initial matrix}) \quad (3a)$$

$$B_0 = I \quad (\text{identity matrix}) \quad (3b)$$

$$A_{i+1} = Q_i^* A_i Q_i \quad (3c)$$

$$B_{i+1} = B_i Q_i \quad (3d)$$

$$\Lambda = \lim_i A_i \quad (3e)$$

$$B = \lim_i B_i \quad (3f)$$

Each plane rotation reduces one off-diagonal element to zero. The time required for the diagonalization is proportional to N cubed, where N is the number of rows or columns in the matrix.

Jacobi, though very slow, enjoys great numerical stability because each elementary transformation is a unitary transformation. The eigenvectors produced by Jacobi form a complete orthonormal set regardless of the distribution of the eigenvalues.

In 1954, Givens developed a method for reducing a general Hermitian matrix to tri-diagonal form using a sequence of unitary transformations. In 1958,

Householder showed how to accomplish the same reduction in half the time and with increased accuracy. Eigenvalues of tri-diagonal matrices can be found in a time proportional to N^2 rather than to N^3 . Consequently, Householder's method has become widely used as a very powerful first step in the solution of the Hermitian matrix eigenvalue problem.

Most methods which find eigenvalues of tri-diagonal matrices do not proceed by applying unitary transformations. Instead, they obtain the eigenvalues directly from the elements of the tri-diagonal matrix. The method of Bisection and Laguerre's method are among these. Using the computed eigenvalues, the corresponding eigenvectors may be found. The method of Inverse Iteration is a very fast and generally good method for obtaining the eigenvectors.

Difficulties arise with Inverse Iteration when eigenvalues become quasi-degenerate. For two degenerate eigenvalues, Inverse Iteration returns only a single eigenvector, not a pair of orthogonal eigenvectors such as Jacobi's method would have produced.

No completely satisfactory method of finding both eigenvalues and eigenvectors of Hermitian matrices has been found which does not proceed by applying elementary unitary transformations to the matrix at every stage.

The QR method described in the next section does apply elementary unitary transformations at every stage and derives its extraordinary stability from this fact. The QR method was developed by Francis(2) in 1961 for general, nonsymmetric matrices, but it is very powerful when applied to symmetric matrices.

MODIFIED QR METHOD

Figure 1 indicates the stages in a computer program designed to find all of the eigenvalues and eigenvectors of an Hermitian matrix from setup to wrapup. HECEVV stands for High Efficiency Complex EigenValues and Vectors(3).

Let the matrix have N rows and columns. Then the Householder step requires $2/3 N^3$ multiplications and the Eigenvector Householder step requires N^3

Subroutine HECEVV

SETUP

Dimensions, constants, addresses, testing for logical input errors, etc.

SCALE

Subtract a constant to render the matrix traceless and scale the matrix, establishing the error limit and the extended error limit.

HOUSEHOLDER

Reduce the matrix to tri-diagonal form by Householder's method. Save the elements of the transformation in the matrix itself.

PHASE

Apply phase factors to make the matrix real, symmetric tri-diagonal.

QR

Find the eigenvalues and eigenvectors of the real, symmetric tri-diagonal matrix by the QR method.

EIGENVECTOR PHASES

Apply phase factors to the eigenvector matrix.

EIGENVECTOR HOUSEHOLDER

Complete the eigenvector matrix by applying the sequence of Householder transformations used to transform the input matrix initially, but in reverse order.

WRAPUP

Re-order the eigenvalues and eigenvectors into descending order of the eigenvalues. De-scale and shift the eigenvalues to recover the solutions to the original problem.

Figure 1

Major stages in a computer program designed to find all of the eigenvalues and eigenvectors of a complex, Hermitian matrix. (Reference 3)

multiplications. These are complex multiplications if the matrix really is complex instead of simply real, symmetric.

We will now concern ourselves with the QR stage. Any complex, Hermitian tri-diagonal matrix can be made real, symmetric by a simple sequence of unitary phase transformations. We need concern ourselves therefore only with the real, symmetric problem.

The QR method proceeds by forming the sequence:

$$A_i - \lambda_i I = Q_i R_i \quad (4a)$$

$$A_{i+1} = R_i Q_i + \lambda_i I \quad (4b)$$

$$\text{or: } A_{i+1} = Q_i^* A_i Q_i \quad (4c)$$

where:

Q_i = unitary matrix

R_i = upper triangular matrix

λ_i = shift

See Wilkinson(1) for the intimate details.

The tri-diagonal nature of A is preserved in these operations. The shift at each step is used to speed the convergence. I use the shift strategy of Kahan and Varah(4), namely, λ = that eigenvalue of the lower right hand 2 by 2 sub-matrix of A which is closer to the last diagonal entry of A . In general, it takes approximately two iterations to isolate the last diagonal entry, that is, to make the off-diagonal element on the lower right hand end effectively zero. The matrix is then said to be deflated. Further operations need be performed only on the remaining, smaller, tri-diagonal matrix.

The QR iteration scheme may be continued until all off-diagonal elements have been reduced essentially to zero. The time required to find the eigenvalues is proportional to N squared.

The eigenvector matrix could be generated along with the eigenvalues by forming the sequence:

$$Q = Q_1 Q_2 Q_3 \dots Q_m \quad (5)$$

where m is the number of QR iterations required to diagonalize A . Because all of the elementary transformations in the sequence are unitary transformations, this procedure is guaranteed to produce a complete orthonormal set of eigenvectors.

The procedure requires a time proportional to N cubed which, though much less than the time required for Jacobi's method, surpasses the time required for the application of the Householder transformations to the original matrix and to the eigenvector matrix. To reduce the time required in this critical step, let us refer back to Equation (4) for the standard QR transformation.

At that point, if λ were an exact eigenvalue and we were using exact arithmetic, only one QR step would be required to isolate the last diagonal element, thereby deflating the matrix. Because constructing the eigenvector matrix is the expensive step, we can save time by finding an eigenvalue in two or three QR iterations and then backing up to take a single QR step to deflate the matrix and to make a single transformation on the eigenvector matrix. With this trick, we have saved a factor of two in the running time of this stage.

Great care must be taken at this step, however. Even if we knew the exact eigenvalue, deflation cannot always be performed in a single step because of round-off error in the finite-precision arithmetic we must use. A test of the last off-diagonal element must be made after the deflation step to determine whether another QR step with the same shift is necessary to insure accuracy.

When I first wrote the program which incorporated this deflation procedure, I found that it was taking an average of two deflation steps for each eigenvalue. Some careful thought explained why: I was computing the eigenvalue to the error limit and was then deflating and testing to see if the last off-diagonal element was less than this same error limit. Due to round-off error, the last off-diagonal element, though very small, still was larger than the error limit. The trouble was resolved by going to double precision computation and finding the eigenvalue to a tighter limit than required overall. Having an eigenvalue

accurate to the extended error limit, computing the parameters of the QR transformation to double precision, though still transforming the eigenvector matrix in single precision, with a final test of the last off-diagonal element against the old error limit, the desired result was produced: an average of one QR deflation for each eigenvalue.

At this point, it becomes possible to take advantage of the properties of the QR transformation and the fact that the eigenvector matrix begins as the identity matrix.

After one QR deflation, the upper triangle of the eigenvector matrix has been filled in with non-zero elements, but the lower triangle is still all zero except for the first sub-diagonal. After the second QR deflation, the first two sub-diagonals are non-zero. After m QR deflations, m sub-diagonals are non-zero. The computing loops which update the eigenvector matrix can be coded to omit the transformations on the zero elements, thereby reducing the running time of this stage by an additional factor of $2/3$, for an overall savings of a factor of three in the QR stage for the eigenvectors.

The number of multiplications required to generate the eigenvector matrix is $4/3 N$ cubed, slightly less than the number, $5/3 N$ cubed, required for the Householder transformations applied to the input matrix and to the eigenvector matrix. If the original matrix is complex instead of real, the Householder transformations will require four real multiplications for each complex multiplication. This makes the QR time pale into insignificance in comparison. A summary of execution times is presented in Figure 2.

There are two distinct steps at which QR transformations are employed. One is while the search for the next eigenvalue is proceeding. The other is the deflation step which is performed to transform the eigenvector matrix once the eigenvalue has been isolated. The latter QR step requires a square root evaluation to determine the sine and cosine of each plane rotation used. However, the former, which is a search for the eigenvalue only, can be performed without having to evaluate any square roots. The method given by Wilkinson(1) for avoiding square roots is not numerically stable and should be avoided. The method given by Kahan and Varah(4), which is a variation of

Summary of Execution Times

1. Householder reduction to tri-diagonal form:
 $\frac{2}{3} N^3$ multiplications (x 4 for complex)
2. Householder transformations of the eigenvectors:
 N^3 multiplications (x 4 for complex)
3. Eigenvectors of tri-diagonal matrix by QR method:
 $\frac{4}{3} N^3$ multiplications (x 1 for complex)

Comparison of Total Times for Three MethodsReal, symmetric matrices:

Householder / QR : $3 N^3$ multiplications
 House. / Inv. Iter. : $1.67 N^3$ multiplications
 Jacobi : $24 N^3$ multiplications

Complex, Hermitian matrices:

Householder / QR : $8 N^3$ multiplications
 House. / Inv. Iter. : $6.67 N^3$ multiplications
 Jacobi : $72 N^3$ multiplications

Figure 2

Summary of execution times for major stages in the diagonalization of an Hermitian matrix and a comparison of total times required for three procedures.

Wilkinson's, is numerically stable. I recommend the use of this latter algorithm. Care must be taken to code the algorithm exactly as Kahan and Varah have given it, for although it appears deceptively similar to Wilkinson's, the subtle differences are of vital importance. The differences are of the nature of the difference between producing a small number by subtraction of two large numbers, with the large round-off error accompanying this procedure, and producing a small number by the multiplication of a large number and a small number, a procedure which yields practically no round-off error.

SUMMARY

The use of the QR algorithm, properly coded, has been shown to be a very fast and very accurate practical procedure for finding all eigenvalues and eigenvectors of Hermitian matrices on a digital computer. Its use should relieve many programmer fears concerning the accuracy of computed eigenvectors at a modest cost in computer time.

REFERENCES

- (1) J. H. Wilkinson, The Algebraic Eigenvalue Problem, Clarendon Press, Oxford (1965).
- (2) J. G. F. Francis, Computer J. 4, 265 and 332 (1961).
- (3) HECEVV is part of a larger package available from the GESHUA Library: HEMP - High Efficiency Matrix Package, GES1030, GE-600 Series Users' Library. HEMP is written in assembly language for General Electric 600-Series computers.
- (4) W. Kahan and J. Varah, Computer Science Department Technical Report No. CS43, Stanford University (1966).

AN ALTERNATIVE APW TECHNIQUE: THEORY AND APPLICATION TO COPPER^{*}

Dale Dean Koelling

Magnetic Theory Group, Northwestern University, Evanston,
Illinois 60201

The variational formulation for the solution of the periodic potential problem has been known for some time¹. Thus, the development of a new band structure method usually consists of creating a basis function set which has some convenient features. What is to be described here is a variant of the APW basis function set.

The APW method², which predates - but can be derived from³ - the variational formulation of the problem, has been used with considerable success for many years⁴. The basis set used is constructed by joining continuously - but with a discontinuous derivative - the solutions of the spherically symmetric potential inside the muffin-tin spheres onto plane waves outside the spheres. This yields a secular equation whose matrix elements depend on the logarithmic derivatives of the radial functions at the sphere boundary. And it is here that one can see a difficulty that can plague the method: As the node of a radial function is brought through the muffin-tin radius, the corresponding logarithmic derivative goes through a singularity (known as an asymptote). Obviously, numerical accuracy is difficult to achieve near this asymptote. In practice, this has not posed any great problem since few bands have been found which cross through the asymptote regions. However, with increasing atomic number, the density of asymptotes in the energy region of interest increases. Further, as we are more and more interested in compounds, it becomes more likely that the bands due to one constituent will cross the asymptote region of another. Thus we are motivated to investigate alternative APW formulations where this difficulty will not appear.

* Supported by the Advanced Research Projects Agency through the Northwestern University Materials Research Center and by the Air Force Office of Scientific Research.

The approach to be used in setting up the basis function is the reverse of the "standard APW" approach. In the standard approach, the differential (i.e., Schrödinger) equation is solved exactly inside the muffin-tin spheres and the boundary conditions are satisfied as a result of the application of the variational formalism. In the alternate approach, the boundary conditions are satisfied exactly and the solution of the differential equation is obtained by the application of the variational formalism. This alternate approach is also considered (in a different fashion) by another paper of this session⁵.

The alternative APW (AAPW) method will be described by presenting a full description of the construction of the basis function set and then sketching the details of the variational formalism. A more complete description of the formalism can be found elsewhere⁶. Then, consistent with the aims of this conference, a number of calculational details and testings will be presented. The entire discussion will be limited to a muffin-tin potential and kept non-relativistic although neither restriction is mandatory, but greatly simplifies the discussion.

The AAPW basis set is to be constructed by expanding the radial function inside the muffin-tin spheres in two sets of functions for each ℓ which are solutions of the eigenvalue problems:

$$h_\ell u_\ell^n(r) = \epsilon_\ell^n u_\ell^n(r); \quad u_\ell^n(R) = 0 \text{ and } \frac{d}{dr} u_\ell^n(R) = 1$$

$$h_\ell v_\ell^\mu(r) = e_\ell^\mu v_\ell^\mu(r); \quad v_\ell^\mu(R) = 1 \text{ and } \frac{d}{dr} v_\ell^\mu(R) = 0$$

In these expressions, h_ℓ is the radial hamiltonian and R is the muffin-tin sphere radius. The value of the non-zero boundary condition merely affects the normalization and can be chosen arbitrarily. The value one is a very convenient choice as it makes the overlap integral of a u with a v function merely R^2 times the reciprocal of the energy difference. Several features of these functions should be noted. It is obvious that, with only one u and one v , the boundary conditions can be easily satisfied although this would not give adequate variational freedom in most cases. The u set and the v set are both orthogonal, complete sets. The resulting overcompleteness will never cause any difficulty as we will be drastically truncating the expansion. Should linear dependence occur, it would show up at an early stage of the calculation and could be easily eliminated. Further, in the full infinite expansion limit with either the u 's or the v 's, one is back to the standard APW basis function⁷. But, as we are truncating the expansion, this will not be relevant.

The radial function is expanded in these functions as,

$$w_\ell(r, k, E) = \sum_{v,v} c(k, E) \mu_\ell^v(r) + \sum_\mu d_\mu^\ell(k, E) v_\ell^\mu(r)$$

The coefficients are determined by the variational requirement that

$$\delta \langle w_\ell | H - E_b | w_\ell \rangle = 0$$

with the additional requirement that the basis function and its derivative be continuous at R. Using a number N of u functions and a number M of v functions, this leads to an $(N + M + 2)$ dimensional matrix equation. The solution of this equation yields

$$c_n^\ell(k, E_b) = q_{n\ell}^1(E_b) j_\ell'(kR) + q_{n\ell}^2(E_b) j_\ell(kR)$$

$$d_\mu^\ell(k, E_b) = q_{\mu\ell}^1(E_b) j_\ell'(kR) + q_{\mu\ell}^2(E_b) j_\ell(kR)$$

The important point to note is that the k-dependence of the expansion coefficients appear in the spherical Bessel functions and their derivatives. Thus, the matrix equation need only be solved once per energy for each ℓ . This represents a considerable savings in calculational effort. In addition, one can consider storing these k-independent coefficients on an energy mesh and interpolating - much as in the treatment of the logarithmic derivatives for the standard APW method. However, as these quantities are more slowly varying, one could even hope to simply use a polynomial fit over a wide range of energies. (This, however, has not yet been tested.) Such a tabulation - interpolation or a polynomial fit scheme can result in a considerable saving in computer time as it eliminates the need to repeatedly set up and solve $(\ell_{\max} + 1)$ matrix equations; the cost in storage space is very modest.

With the radial functions w_ℓ constructed, the basis functions are

$$(N_c \Omega)^{1/2} f(r, k, E_b) \equiv \exp ik \cdot r$$

$$+ 4\pi \sum_n \theta(r_n, R) \exp ik \cdot R \sum_\ell^{\ell_{\max}} i \ell Y_L^*(k) Y_L(r_n)$$

$$\otimes \Delta_\ell(r_n, k, E_b)$$

$$\Delta_\ell(r_n, k, E_b) \equiv w_\ell(r_n, k, E_b) - j_\ell(k r_n)$$

$$\tilde{r}_n \equiv r - R_n$$

$$L \equiv \{\ell, m\}$$

where the R_i are the positions of the lattice sites, the carat is used to indicate a unit vector, and N_c is the number of unit cells of volume Ω . $\theta(r, R)$ is the step function with value one for $r \leq R$ and zero for $r > R$. The value of ℓ_{\max} is certainly less than four and must be at least two for a transition metal.

With this basis function set, one gets a secular equation where the Hamiltonian matrix is

$$H_{ij}(E_b) = k_j^2 \delta_{ij} + V(\hat{k}_j - \hat{k}_i) + \sum_\ell P_\ell(\hat{k}_i \cdot \hat{k}_j) G_{ij}^\ell(E_b)$$

$$G_{ij}^\ell(E_b) \equiv \frac{4\pi}{\Omega} (2\ell+1) \{ [w_\ell(k_i, E_b) | h_\ell | w_\ell(k_j, E_b)] - [j_\ell(k_i r) | h_\ell | j_\ell(k_j r)] \}$$

and the overlap matrix is

$$S_{ij}(E_b) = \delta_{ij} + \sum_\ell P_\ell(\hat{k}_i \cdot \hat{k}_j) F_{ij}(E_b)$$

$$F_{ij}^\ell(E_b) = \frac{4\pi}{\Omega} (2\ell+1) \{ [w_\ell(k_i, E_b) | w_\ell(k_j, E_b)] - [j_\ell(k_i r) | j_\ell(k_j r)] \}$$

In these expressions the square bracket is used to indicate a radial integral - the angular integrals having already been done. Thus,

$$[j_\ell(kr) | j_\ell(kr)] = \frac{1}{2} R^3 \{ j_\ell'(x)^2 + (\frac{1}{x}) j_\ell'(x) j_\ell(x) + (1 - \ell(\ell+1)/x^2) j_\ell(x)^2 \}$$

$$[j_\ell(k_1 r) | j_\ell(k_2 r)] = \frac{R^2}{(k_2^2 - k_1^2)} \{ j_\ell'(x_1) j_\ell(x_2) - j_\ell'(x_2) j_\ell(x_1) \}$$

$$x_i \equiv k_i R$$

Obviously, since the w_ℓ functions are expanded in eigenvalue function sets whose overlaps are known, their integrals are easily and quickly done. The Bessel function overlap integrals have been displayed above and

$$[j_\ell(k_1 r) | h_\ell | j_\ell(k_2 r)] = k_2^2 [j_\ell(k_1 r) | j_\ell(k_2 r)] + [j_\ell(k_1 r) | V | j_\ell(k_2 r)]$$

Thus, the only radial integrals to be performed numerically are the Bessel function integrals of the potential. (And these need be done only once for a given point in reciprocal space.) Of course, since the $V(\tilde{k}_1 - \tilde{k}_2)$ are just the Fourier components of the potential, these can be tabulated and stored for a given potential. Thus, one can see that it is easy to set up the matrix elements rapidly. (Note that the presence of $V(\tilde{k})$ in the calculation for the muffin-tin potential makes the extension to the warped muffin-tin potential⁸ especially easy.)

To test the applicability of this AAPW method, it has been applied to a Chodorow copper potential. This potential was obtained by copying it from the paper by Burdick⁹ and then interpolating it onto the radial mesh used by Liberman, Waber, and Cromer¹⁰ for their atomic calculations (a standard mesh in our work). Unfortunately, this procedure restricts the accuracy of the test since the potential is modified by both the inaccuracies of the interpolation and the change in muffin-tin sphere radius. There is the further (untested) effect that the radial equations inside the muffin-tin sphere were solved by applying the Milne method to the coupled (non-relativistic) first order differential equations¹¹ rather than the applying the Numerov method to the Herman-Skillman mesh as was done in the APW⁹ and KKR¹² calculations. Even with these restrictions, it is still possible to learn much from the applying the method to this potential. One finds that convergence in both reciprocal space and angular momentum are quite acceptable and that program efficiency is good¹³. And this is with a simple test code which has had almost no optimization.

To compare the eigenvalue results, the first six bands at the points Γ , X , and L are to be compared to the Burdick results. To do this, the following four quantities are calculated¹⁴:

$$\Delta \equiv \frac{1}{18} \sum_{n, \tilde{k}} e(n, \tilde{k})$$

$$\sigma^2 \equiv \frac{1}{18} \sum_{n, \tilde{k}} e^2(n, \tilde{k}) - \Delta^2$$

$$\Delta_{sd} \equiv \frac{1}{3} \sum_{\tilde{k}} \delta_{sd} (\tilde{k})$$

$$\sigma_{sd}^2 \equiv \frac{1}{3} \sum_{\tilde{k}} \delta_{sd}^2 (\tilde{k}) - \Delta_{sd}^2$$

where

$$e(n, \tilde{k}) \equiv E_n(\tilde{k}) - E_n^{BURDICK}(\tilde{k})$$

$$\delta_{sd}(\tilde{k}) \equiv \frac{1}{5} \sum_d e_d(n, \tilde{k}) - e_s(\tilde{k})$$

In these expressions, an s or d subscript is used on the e's to indicate that only the states of that character are to be included. The values of these quantities are shown in Table I together with the identical KKR values¹² (for comparison).

Table I. Values of Quantities Defined in Text.
(Units are Milli Rydbergs.)

	AAPW	KKR
Δ	3.3	-1.0
σ	5.5	6.4
Δ_{sd}	1.4	-13.7
σ_{sd}	12.8	12.8

As can be seen, the results are quite reasonable. Although the results thus far have been obtained on an admittedly crude pilot code, they are sufficiently encouraging to warrant further development of the method. In fact, current plans are to go immediately to a relativistic calculation on a compound.

REFERENCES

1. W. Kohn, "Variational Methods for Periodic Lattices", Phys. Rev. 87, 472 (1952).
2. J.C. Slater, "Wave Functions in a Periodic Potential", Phys. Rev. 51, 846 (1937).
3. R.S. Leigh, "The Augmented Plane Wave and Related Methods for Crystal Eigenvalue Problems", Proc. Phys. Soc. (London) 69, 388 (1956); H. Schlosser and P.M. Marcus, "Composite Wave Variational Method for Solution of the Energy Band Problem in Solids", Phys. Rev. 131, 2529 (1963); P.M. Marcus, "Variational Methods in the Computation of Energy Bands", Int. J. of Quantum Chem. 1, 567 (1967).
4. L.F. Mattheis, J.H. Wood, and A.C. Switendick, "A Procedure for Calculating Electronic Energy Bands Using Symmetrized Augmented Plane Waves", Meth. in Comp. Physics (vol. 8), Academic Press, New York (1968).
5. H. Bross, "Ein Neues Verfahren zur Berechnung von Einelektronenzuständen in Kristallen", Phys. Kondens. Materie 3, 119 (1964); H. Bross and H. Stohr, "Modified Augmented Plane Wave Method for the Calculation of Energy Bands and Electronic Wave Functions in Metals", Phys. Lett. 8, 240 (1964); H. Bross, "On a Formalism to Calculate Single Particle Wave Functions in Crystals Taking into Account Many-Body Aspects", Z. für Physik 215, 485 (1968).
6. D.D. Koelling, "Alternative Augmented-Plane-Wave Technique: Theory and Application to Copper", Phys. Rev. (to be published) (July, 1970).
7. J.C. Slater, "An Augmented Plane Wave Method for the Periodic Potential Problem", Phys. Rev. 92, 603 (1953); M.M. Saffren and J.C. Slater, "An Augmented Plane Wave Method for the Periodic Potential Problem II", Phys. Rev. 92, 1126 (1953).
8. D.D. Koelling, A.J. Freeman, and F.M. Mueller, "Shifts in the Electronic Band Structure of Metals Due to Non-Muffin-Tin Potentials", Phys. Rev. 1, 1318 (1970).
9. G.A. Burdick, "Energy Band Structure of Cu", Phys. Rev. 129, 138 (1963).
10. D. Liberman, J.T. Waber, and D.T. Cromer, "Self-Consistent-

Field Dirac-Slater Wave Functions for Atoms and Ions I.
Comparison with Previous Calculations", Phys. Rev. 137, A27
(1965).

11. The procedure used is a modified version of the FOVRG routine of Loucks with $c^{-1} = 0$. Sec. App. 7 of T. Loucks, Augmented Plane Wave Method, W.A. Benjamin, New York (1967).
12. B. Segall, "Fermi Surface and Energy Bands of Copper", Phys. Rev. 125, 109 (1962).
13. A detailed discussion of this convergence and speed questions is given in Appendix B of ref. 6.
14. These are the same quantities defined to study the effects of the warping of the muffin-tin potential. See ref. 8.

AN RAPW EXPANDED BASIS SET[†]

G. Arbman^{*} and D. Koelling

Magnetic Theory Group, Physics Department

Northwestern University, Evanston, Illinois 60201

I. INTRODUCTION

The relativistic augmented plane wave (RAPW) method¹ had been used extensively to perform energy band calculations for the heavier (higher atomic number) materials. Except for the use of the relativistic formalism, it is precisely the same method as the non-relativistic APW method² which has done so well in the region of lower atomic numbers. Both are energy-variational based techniques^{3,4} using basis functions constructed by joining atomic-like functions near the nuclei onto plane waves in the outer regions. However, a new feature of an additional spin variable arises in the RAPW method because the spin-orbit coupling term has been included. This makes the RAPW method considerably more time consuming to apply because it doubles the size of the basis set in addition to making the matrix elements slightly more complicated.

Adding to the difficulty is the fact that in order to properly describe the often densely spaced bands of the heavier materials normally studied with this method, one must demand high numerical accuracy (i.e. good convergence) of the method. Consequently, the computational times of this method--although apparently not much longer than those of other relativistic methods--are long enough to make the cost of evaluating the energies for a large number of points in reciprocal (k) space undesirably high.

For an optimally programmed RAPW calculation, most of the calculational time is spent evaluating determinants (roughly 90%). Since the time required to evaluate a determinant varies as the matrix size to the n 'th power (in practice, $2 < n < 3$), an obvious way to reduce the computational time is to decrease the size of the matrix. To do so, a more efficient basis set than the individual

[†]Supported by the Air Force Office of Scientific Research.

^{*}On leave from FOA, Stockholm.

RAFW's is needed. We are currently investigating the possibilities of using linear combinations of the RAPW basis functions to create more efficient basis function sets. Although the main objective is to cut down the size of the matrix needed at a general k -space point and thus reduce the computational time, there are many other advantages arising from the flexibility of such a procedure. For example, by the proper choice of prescription for one's linear combinations, one can perform a symmetrized calculation^{5,6}, "telescope" a partially converged solution to convergence, or obtain local (in k -space) eigenvalues with small matrices to avoid interpolation schemes.

Because we are interested in the general approach of using linear combinations of RAPW's, we first discuss the practical general computational problems of such an approach in Sec. II with no prescription for the linear combinations to use. The computer codes that we are working with are constructed along these lines so that one can explore the full spectrum of possibilities of the approach. In Sec. III, a number of possible prescriptions for forming linear combinations are discussed. These, of course, are to be tailored to one's prior knowledge of the system for which the calculations are performed. Finally, in Sec. IV, we discuss the particular prescription designed from the concepts of the $\vec{k} \cdot \vec{p}$ method and some very preliminary results of our experiments with it.

II. GENERAL CONSIDERATIONS FOR THE USE OF LINEAR COMBINATIONS

The discussion to follow assumes that the reader is familiar with the RAPW formalism^{4,7}. An RAPW basis function will be denoted as $|\vec{k}_n s\rangle$ where $\vec{k}_n \equiv \vec{k} + \vec{K}_n$ with \vec{K}_n a reciprocal lattice vector and s the spin coordinate. In order to facilitate the discussion which follows, the matrix elements can be written in the following somewhat unorthodox form:

$$\begin{aligned} \langle \vec{k}_n s | H - E | \vec{k}_m s' \rangle &= (s | s') \{ A_{nm} - EB_{nm} \\ &+ \sum_{\vec{t}}^{\ell_1} \sum_L c_L^{\vec{t}}(\vec{k}_n)^* \xi_{\ell}(E) c_L^{\vec{t}}(\vec{k}_m) \} \\ &+ \sum_{\vec{t}}^{\ell_1} \sum_L (s | c_L^{\vec{t}}(\vec{k}_n)^* \vec{L} \cdot \vec{\sigma} c_L^{\vec{t}}(\vec{k}_m) | s') \\ A_{nm} &= \vec{k}_n \cdot \vec{k}_m U(\vec{k}_m - \vec{k}_n) + S(\vec{k}_m - \vec{k}_n) \sum_{\ell=\ell_1+1}^{\ell_{\max}} \alpha_{\ell} P_{\ell}(\hat{k}_n \cdot \hat{k}_m) j_{\ell}(k_n R) j_{\ell}(k_m R) \end{aligned}$$

$$B_{nm} = U(\vec{k}_m - \vec{k}_n) + S(\vec{k}_m - \vec{k}_n) \sum_{\ell=\ell_1+1}^{\ell_{max}} \beta_\ell P_\ell(k_n \cdot k_m) j_\ell(k_n R) j_\ell(k_m R)$$

$$c_L^t(\vec{k}) = \left(\frac{12\pi V_o}{(2\ell+1)R} \right)^{1/2} \exp[i\vec{k} \cdot \vec{t}] j_\ell(kR) v_L(k)$$

$$L \equiv \{\ell, m\}$$

$$S(\vec{k}) = \left(\frac{3V_o}{R} \right) \sum_t \exp[i\vec{k} \cdot \vec{t}]$$

$$U(\vec{k}) = \delta_{K0} - S(\vec{k}) j_1(kR)/k$$

This expression can be shown to be algebraically equivalent to those used previously in RAPW calculations.⁸ The symbols used above are defined as follows: $|s\rangle$ is a two component spinor; E is the trial energy; V_o is the ratio of the volume of one APW sphere to that of the unit cell; R is the radius of the APW sphere; ξ_ℓ and η_ℓ are combinations of the large component logarithmic derivatives for $j = \ell \pm s$, i.e., with

$$g'_\kappa/g_\kappa \approx cf_\kappa/g_\kappa - (\kappa+1)/R$$

$$j = |\kappa| - 1/2$$

$$\ell = \begin{cases} |\kappa| - 1 & \kappa < 0 \\ \kappa & \kappa > 0 \end{cases}$$

one defines

$$\xi_\ell = \ell g'_\ell/g_\ell + (\ell+1) g'_{-\ell-1}/g_{-\ell-1}$$

$$\eta_\ell = g'_\ell/g_\ell - g'_{-\ell-1}/g_{-\ell-1}$$

so that in the non-relativistic limit ξ_ℓ becomes $(2\ell+1)$ times the logarithmic derivative and η_ℓ becomes zero; α_ℓ and β_ℓ are obtained from the linear energy dependence of the ξ_ℓ for $\ell > \ell_1$

$$\xi_\ell(E) = \alpha_\ell - E\beta_\ell$$

(It is interesting to note that β_ℓ is almost independent of ℓ .); the vectors t are the positions of the atoms within the unit cell. L is the angular momentum operator; and $\vec{\sigma}$ is the (2×2) Pauli spin operator.

The linear combination basis function is to be of the form

$$|\#i\rangle \equiv \sum_{ns} c_i^{ns} |\vec{k}_n s\rangle$$

so that one can write the matrix elements straightforwardly as

$$\langle \#i | H-E | \#j \rangle = \sum_{\substack{n,s \\ n',s'}} c_i^{ns} {}^* \langle \vec{k}_n s | H-E | \vec{k}_{n'} s' \rangle c_j^{n's'}$$

Of course, if one were to merely program this expression as it stands, one would expend a prohibitive amount of time evaluating matrix elements. In symmetrization, one is saved by the idempotent property of the projection operator--which reduces an N_{lc}^2 process down to an N_{lc} process. Here, however, it is necessary to be a bit more clever. However, from the form in which the RAPW matrix elements were written, it is easily seen that one can express the matrix element as

$$\begin{aligned} \langle \#i | H-E | \#j \rangle &= (A)_{ij} - E(B)_{ij} + \sum_t \sum_L (\vec{c}_L^t)^\dagger \xi_\ell (\vec{c}_L^t)_j \\ &\quad + \sum_t \sum_L (\vec{c}_L^t)_i \eta_\ell \vec{L} \cdot \vec{\sigma} (\vec{c}_L^t)_j \end{aligned}$$

$$(A)_{ij} = \sum_{\substack{n,s \\ n',s'}} c_i^{ns} {}^* A_{nn'} c_j^{n's'}$$

$$(B)_{ij} = \sum_{\substack{n,s \\ n',s'}} c_i^{ns} {}^* B_{nn'} c_j^{n's'}$$

$$(\vec{c}_L^t)_i = \sum_{ns} c_i^{ns} c_L^t(\vec{k}_n) |s\rangle$$

Obviously the RAPW matrix elements were written in a form to make the various calculational expediencies transparent; they need not be described in detail.

The ability to use the linear fit for the ξ_ℓ 's of higher ℓ is extremely important since it means that relatively few C_L 's are needed, so that storage requirements are reasonable. Usually ℓ_1 can be set to 2 or 3 and is certainly safe at 4. The matrices

(A)_{ij} and (B)_{ij} are energy independent and need to be set up only once. Thus it is not advantageous to attempt a factorization. However, because of the large number of matrices multiplying the logarithmic derivatives, it is very desirable to factorize into the (\vec{C}_L^T)_i vectors--thereby reducing from a N_{lc}^2 to an N_{lc} process and requiring much less storage. As an added bonus, it is no longer necessary to store separate quantities for the spin-orbit terms of the matrix elements.

Our general-case experimental computer program has been written in this form so that any linear combinations of RAPW's can be used as basis functions. As a result, we are able to explore the use of linear combinations of RAPW's in depth, with the expectation that a great deal will be discovered about wave functions in the process.

III. SOME PRESCRIPTIONS FOR FORMING LINEAR COMBINATIONS

It is when one demands the prescription for forming the linear combinations that one arrives at the heart of the linear combination approach. And it is to this point that all one's prior knowledge of the system must be applied because the success of the approach depends precisely on this point. There is, of course, no single "best" prescription. We list here several interesting prescriptions--some of which have already proven successful, some of which we are currently testing, and some of which are purely speculative. The order of presentation will be roughly from most sure to most speculative.

A. SRAPW

The one proven successful application of the linear combination approach is that of the symmetrized RAPW (SRAPW) method⁶ (and also SAPW for the non-relativistic APW method⁵) based on the use of projection operator techniques. Of course, this is only applicable to k-space points which have additional symmetry beyond the translational symmetry. A very useful (and crucial) test of the computer codes is to input the linear combinations that are formed by the application of the projection operator and thus perform an SRAPW calculation--which can then be compared with the results of other SRAPW calculations.

B. Telescoping

Another useful technique which is suggestively called telescoping works as follows: After one has performed a calculation

which has not fully converged to the accuracy required, one uses the partially converged solution together with a set of RAPW's (some of which may already be included in the partially converged solution) as a new basis set. This procedure is continued until the desired accuracy is obtained. Thus one has a vehicle for pushing to extreme convergence limits without the requirement of extremely large matrices.

C. a la $\vec{k} \cdot \vec{p}$

The prescription which we are currently exploring is based on the success of the $\vec{k} \cdot \vec{p}$ methods--i.e., on the use of the solutions at one \vec{k} -point to make up the basis functions for another \vec{k} -point. It should be pointed out, however, that this prescription is only in the spirit of the $\vec{k} \cdot \vec{p}$ method. Other workers have already applied the APW momentum matrix elements to the $\vec{k} \cdot \vec{p}$ formalism--which is not a linear combination approach. The details of the actual prescription used are discussed in the next section together with an analysis of its limitations.

D. Perturbation Modified

To understand the motivation for this next prescription, it is necessary to recall that one needs very few RAPW's to converge to an accuracy of .01 Ryd. The remaining basis functions are included to reduce the convergence accuracy below this level. Thus it is reasonable to attempt to include these vectors by a perturbation-based technique. In particular, one chooses a maximum \vec{k} -vector length k_2 and a minimum perturbation length k_1 . Then one uses the basis set

$$|\#_{\vec{k}_n} s\rangle = |\vec{k}_n s\rangle - \sum_{n', s'}^{k_1 \leq |k_n'| \leq k_2} \left(\frac{\langle \vec{k}_n', s' | H-E | \vec{k}_n s \rangle}{\langle \vec{k}_n, s' | H-E | \vec{k}_n, s' \rangle} \right) |\vec{k}_n, s'\rangle$$

with $|k_n| < k_1$. One can be sure that this will work because a simplification of this procedure has already worked successfully. The simplified procedure consists of formulating this same problem but ignoring the off-diagonal terms in the second order interactions of the "perturbation terms". While this is consistent with the prescription used, the simplified procedure then loses its variational properties--which can lead to difficulties. The correct use of this perturbatively modified basis set is completely variational--with a restricted basis set.

E. Known Properties

Finally, there is the possibility of forming linear combinations which satisfy known properties of the solution. For example, one might try to analyze the overlap of atomic d states in the outside regions of the unit cell to obtain good basis functions for d-band calculations. Or one might form linear combinations which have continuous derivatives across the muffin-tin sphere boundary. Or one can use any of many other possibilities--limited only by one's inventiveness. Clearly, this open-ended area of investigation has not been exhausted by the short list presented here.

IV. RAPW A LA $\vec{k} \cdot \vec{p}$

In this section, we discuss in some detail the prescription for the linear combinations based on $\vec{k} \cdot \vec{p}$ considerations. But first let us comment on the objection that $\vec{k} \cdot \vec{p}$ approaches are normally not applied to metals--our prime (although not only) interest. The reason the $\vec{k} \cdot \vec{p}$ is not applied to metals is merely that one would get large matrices, and the $\vec{k} \cdot \vec{p}$ method is primarily an effort to get away from large matrices (and often as a basis for a parametrization scheme). We, however, are already dealing with much larger matrices so that the $\vec{k} \cdot \vec{p}$ matrix is small by comparison.

To formulate the prescription, assume that one has solved for a set of eigenvalues and eigenvectors at the point g so that one has the solutions:

$$|\#n, \vec{g}\rangle = \sum_{m s} c_{ms}(n, \vec{g}) |\vec{g}_m s\rangle$$

Then one can use as a basis function set at the point \vec{k} those functions formed by associating the same coefficient with the addition of the same reciprocal lattice vector. Then the new basis set for the point \vec{k} is chosen as

$$|\# n(\vec{g}), \vec{k}\rangle \equiv \sum_{m s} c_{ms}(n, \vec{g}) |\vec{k}_m s\rangle$$

Clearly, for the inclusion of a specified number of bands (i.e. basis functions), this basis function set will have a finite radius in \vec{k} -space for convergence to a given accuracy. The crucial question for the utility of this prescription is the convergence question--How does this radius vary with the number of basis functions used?

To graphically illustrate the conditions which restrict the radius of convergence, the first few fcc reciprocal lattice vector

TABLE I. List of Stars for the fcc Reciprocal Lattice

Star #	$\Gamma (000)$	$X (100)$	$L (\frac{1}{2}\frac{1}{2}\frac{1}{2})$
1	(000)	(000), (-200)	(000), (-1-1-1)
2	(±1±1±1)	(-1±1±1)	(-1-11), (-11-1), (1-1-1), (-200), (0-20), (00-2)
3	(±200), (0±20), (00±2)	(-2±20), (-20±2), (0±20), (00±2)	(11-1), (1-11), (-111), (-2-20), (-20-2), (0-2-2)
4	(±2±20), (±20±2), (0±2±2)	(1±1±1), (-3±1±1)	(111), (-2-2-2)
5	(±3±1±1), (±1±3±1), (±1±1±3)	(200), (-400)	(-3-1-1), (-1-3-1), (-1-1-3), (200), (020), (002)
6	(±2±2±2)	(-2±2±2), (0±2±2)	(2-20), (-220), (-202), (20-2), (02-2), (0-22), (-31-1), (-3-1-1), (1-3-1), (-1-31), (1-1-3), (-11-3)
7	(±400), (0±40), (00±4)	(-1±3±1), (-1±1±3)	(-311), (1-31), (11-3), (2-2-2), (-22-2), (-2-22)
8	(±3±3±1), (±3±1±3), (±1±3±3)	(20±2), (2±20), (-40±2), (-40±2)	(220), (202), (022), (-3-3-1), (-3-1-3), (-1-3-3)
9	(±4±20), (±40±2), (0±4±2), (±2±40), (±20±4), (0±2±4)	(1±3±1), (1±1±3), (-3±3±1), (-3±1±3)	(-400), (0-40), (00-4), (3-1-1), (-13-1), (-1-13)
10	(±4±2±2), (±2±4±2), (±2±2±4)	(2±2±2), (-4±2±2)	(-4-20), (-40-2), (0-4-2), (-2-40), (-20-4), (0-2-4)

stars at the points Γ , X, and L (in units of $2\pi/a$) are shown in Table I in order of increasing resultant vector length. The significant point to be seen from this table is that very different vector sets appear for each point--with the same reciprocal lattice vector being included at very different levels of convergence for each point. Yet these are the vector (or RAPW) groupings whose coefficients will be related by symmetry considerations. Thus we can see that to go from X or L to Γ would require the use of a large number of basis functions. And more important, just to have the required vectors for convergence at Γ , one would need to have converged the X (or L) calculations extremely well (far beyond current practice).

If, instead, one were to try going only half the distance from X (or L) to Γ , these differences in the choice of which reciprocal lattice vectors to include would be much less drastic. This is the prime feature in the limitation of the radius of convergence. The fact that different symmetry points require different vector combinations becomes less important after reaching 6 or 8 basis functions because vector sets of the same form will be mixed in the same way.

Unfortunately, we are not yet able to answer these convergence questions from actual calculations as our computer codes have not been completely tested out.¹⁰ Thus it is premature to discuss the results of the few test runs (done primarily for timing) in detail. We are, however, encouraged to find that we are able to get an increase in speed of one to two orders of magnitude for those points which we treat using this method. And the radius of convergence appears to be sufficiently large with the inclusion of six basis functions to permit the use of a mesh with a size $\frac{1}{2}$ to $\frac{1}{4}$ the Γ to X distance in a transition metal.

We can also make two other observations which are not at all surprising but very clearly evident. The first is that one need only extend the basis function sets by including the higher energy solutions. In fact, even the basis functions formed from solutions at the bottom of the band are found to contribute almost nothing to the variational freedom for those eigenvalues at the top of the band (orthogonalization effects). The second is that one should not mix basis functions produced from solutions at different \vec{k} -points. The resulting approximate linear dependence yields spurious zeroes of the determinant throughout the energy range which are extremely difficult to eliminate.

V. CONCLUSION

One can find the prime example of the "conservation of effort

theorem" described earlier by J.W.D. Connolly in the comparison of the Augmented Plane Wave (APW) and Greens function (KKR) methods. The APW matrix elements are extremely quick and easy to evaluate, but the resulting matrix can become undesirably large in size. On the other hand, much more effort must be expended to evaluate the KKR matrix elements, but then one deals with a matrix whose dimensions are usually less than or equal to nine. In fact, one can find even more support for the "conservation of effort theorem" by just looking at the KKR method in two different representations: the standard spherical harmonic and the plane-wave representation. In the plane-wave representation, one is back to the APW end of the work distribution scale with easy matrix elements and large matrices. (In fact, somewhat larger than the APW matrices). Obviously, this conservation theorem has much empirical evidence supporting it. NONETHELESS, the subject of this paper is an attempt to violate this theorem.

The basic idea of the approach is a simple one: instead of using the basis functions individually, one can form linear combinations and use these as the new basis functions. If the calculations were carried out with complete sets, this would be merely a rotation of one's basis vectors. However, since one must truncate the sum, this will be a different restriction on one's variational freedom. The basis for improvement is then dependent on being able to choose that set of linear combinations which provides adequate variational freedom with a smaller number of basis functions.

One sees that it might be possible to violate the conservation of effort theorem, because it is incomplete as currently stated. It should be modified to require the additional condition that one starts from the same state of ignorance. If one has other knowledge that can be exploited, it must be possible to conserve some effort. We have seen another successful example of the linear combination approach for the LCAO method in the paper presented in the previous session by Lafon and Liu where they performed a calculation for diamond using only bonding orbitals. Hopefully, we will see even more in the future.

REFERENCES

1. T. Loucks, Augmented Plane Wave Method, W. A. Benjamin, New York (1967).
2. J. C. Slater, Phys. Rev. 51, 846 (1937).
3. R. S. Leigh, Proc. Phys. Soc. (London) 69, 388 (1956); H. Schlosser and P. M. Marcus, Int. J. of Quantum Chem. 1, 567 (1968).

4. T. Loucks, Phys. Rev. 139, A1333 (1965).
5. L. F. Mattheiss, J. H. Wood, and A. C. Switendick, Meth. in Comp. Phys. (vol. 8), Academic Press, New York (1968).
6. D. Koelling, Phys. Rev. 188, 1049 (1969).
7. D. Koelling, "The relativistic APW Method and its Convergence in ℓ ", Quarterly Progress Report of the Solid State and Molecular Theory Group, MIT (unpublished) 68, 36 (1968). Also see ref. 6.
8. Some of the details are given in the refs. 6-7 and the use of the spherical harmonic sum rule is obvious. The rewriting of the spin-orbit term to actually look like a spin-orbit term (by G.A.) will be presented in the forthcoming paper.
9. N. Parada, "Localized Defects in PbTe via a $\vec{K} \cdot \vec{\Pi}$ - APW Energy Band Calculation", Tech. Rep. No. 8 of the Materials Theory Group, MIT (1968).
10. Our SRAPW tests indicate that there is still a bug in the construction of the RAPW matrix elements leading to small (.02 Ryd.) errors. Thus we can only discuss those features which are unaffected by these errors.

NEW VERSION OF THE MODIFIED AUGMENTED-PLANE-WAVE METHOD *

H. Bross, G. Bohn, G. Meister, W. Schubö, and H. Stöhr
Sektion Physik, University of Munich, Germany

I. Introduction

The Augmented-Plane-Wave (APW) method proposed by Slater (/1/,/2/) has proved very successful for the calculation of the band structure of solids. However, aiming especially at improved wave functions, some refinements seem desirable.

In this paper a new version of the Modified Augmented-Plane-Wave (MAPW) method is reported. The characteristics of this approach are as follows.

- (i) Exact continuity of both the wave functions and their first derivatives on the surfaces of the APW spheres. This is a necessary demand if the wave functions are to be used to calculate matrix elements, in particular those of the momentum operator and functions thereof.
- (ii) General potential. The assumption of a muffin-tin potential is not appropriate in a number of solids, particularly in self-consistent investigations it is expected to be less successful.
- (iii) Orthogonality of the wave functions of the valence electrons to the wave functions of the core electrons.
- (iv) Consideration of exchange and correlation effects beyond the Hartree-Fock approximation using a non-local potential. In addition one would like to have a formalism which can be handled on a computer using standard techniques of numerical analysis.

Different modifications of the APW method (/3/-/10/) have been proposed with the above mentioned goals. The version previously proposed by one of the present authors (/5/,/8/,/9/)

*Supported in part by the Deutsche Forschungsgemeinschaft

incorporates the first requirement. Further it allows for a non-constant potential in the plane wave region as do the formulations of Leigh /3/, Schlosser and Marcus /4/, and De Cicco /6/. General potentials have been studied by Marcus /7/, and others (/8/,/9/,/10/). In the 1937 formulation of the APW method the wave functions of the valence electrons will be orthogonal to those of the core states if the latter are also determined by the APW method. A nonlocal potential taking into account many body aspects has been considered by two of us (/8/,/11/).

II. Description of the Method

Instead of solving the Schrödinger equation directly we transform it into the equivalent variational principle, thus obtaining a standard eigenvalue problem.

A. The MAPW Trial Functions

The trial functions outside the APW spheres are expanded into plane waves.

$$\psi_{\vec{k}}^{II}(\vec{r}) = \sum_j v_j e^{ik_j \vec{r}} , \text{ for } r > r_i \quad (1)$$

Inside the spheres we use products of radial functions and spherical harmonics up to the value L of the azimuthal quantum number l and products of spherical bessel functions and spherical harmonics from L+1 to ∞ .

$$\begin{aligned} \psi_{\vec{k}}^I(\vec{r}) = & \sum_n \sum_{l=0}^L (2l+1) i^{l+2m} A_{nlm} R_{nl}(r) Y_{lm}(\vec{r}^o) \\ & + \sum_j \sum_{l=L+1}^{\infty} v_j (2l+1) i^{l+2m} Y_{l,-m}(\vec{k}_j^o) j_l(k_j r) Y_{lm}(\vec{r}^o), \end{aligned} \quad (2)$$

for $r \leq r_i$.

Hereby \vec{k}_j is defined by $\vec{k}_j = \vec{k} + \vec{K}_j$, where \vec{K}_j denotes a reciprocal lattice vector.

The functions R_{nl} are solutions of the radial differential equation with the spherical mean value of the potential

$$R''_{nl} + \frac{2}{r} R'_{nl} + \left[E_{nl} - \frac{l(l+1)}{r^2} - V_{sph}(r) \right] R_{nl} = 0. \quad (3)$$

As shown in a detailed discussion /12/ it is useful to consider the solutions of eq. (3) in the region $0 \leq r \leq r_o$, where $r_o > r_i$. The dependence of the R_{nl} 's on the parameters E_{nl} is used to construct an orthogonalized set for each value of l by choosing the E_{nl} 's appropriately. This is done by demanding that the logarithmic derivatives at $r=r_o$ have the same value for all n ,

$$R'_{nl}(E_{nl}, r_o) / R_{nl}(E_{nl}, r_o) = C_l.$$

We are left with one free parameter for each l only. The index n now counts the zeros of the R_{nl} 's in $0 < r \leq r_o$. It has proved advantageous to fix the C_l 's by choosing a value of E which has the order of magnitude of the expected energies of the valence band and by calculating $C_l = R'_l(E, r_o) / R_l(E, r_o)$.

The expressions for ψ_k , eqs. (1) and (2), may be rewritten yielding

$$\begin{aligned} \psi_k(\vec{r}) = & \sum_{j=1}^J v_j e^{ik_j \vec{r}} + \Theta(r_o - r) \sum_{l=0}^L (2l+1) i^{l+2m} Y_{lm}(\vec{r}^o) \\ & \times \left[\sum_{n=1}^N A_{nlm} R_{nl}(r) - \sum_{j=1}^J v_j Y_{l,-m}(k_j^o) j_l(k_j r) \right], \end{aligned} \quad (4)$$

$$\text{with } \Theta(r_o - r) = \begin{cases} 0 & \text{in region II} \\ 1 & \text{in region I} \end{cases}$$

We note that all sums are finite and, as the results show, extend to rather small upper limits. This proves advantageous for the calculation of matrix elements or a self-consistent potential with these wave functions.

B. The Conditions for the Continuity of the Wave Functions

From eq. (4) we deduce that the trial functions and their derivatives are continuous for $l > L$. In order to make the wave functions and their first derivatives continuous across the surfaces of the APW spheres, we only have to demand that for $l \leq L$ the two equations

$$\sum_n A_{nlm} R_{nl}(r_i) - \sum_j v_j Y_{l,-m}(k_j^o) j_l(k_j r_i) = 0, \quad (5)$$

$$\sum_n A_{nlm} R'_{nl}(r_i) - \sum_j v_j Y_{l,-m}(\vec{k}_j^o) k_j J'_l(k_j r_i) = 0 \quad (6)$$

hold for each pair l,m .

The presence of the second sum in eq. (2) makes it possible to work with a rather small value of L . Usually it is sufficient to limit L to 2.

C. The Variational Principle for a Local Potential

We assume the periodic potential $V(\vec{r})$ to be known. Its deviation from a muffin-tin form, $\tilde{V}(\vec{r})$, is expressed by

$$\tilde{V}(\vec{r}) = \begin{cases} V(\vec{r}) - V_{\text{sph}}(r) & r \leq r_i \\ V(\vec{r}) & \text{for } r > r_i \end{cases} . \quad (7)$$

Two representations are used,

i) an expansion in a Fourier series,

$$\tilde{V}(\vec{r}) = \sum_j \tilde{V}(\vec{k}_j) e^{i\vec{k}_j \cdot \vec{r}} , \quad (8)$$

and

ii) inside the APW sphere, an appropriate expansion in terms of lattice harmonics,

$$\tilde{V}(\vec{r}) = \sum_{l>0} \tilde{V}(l,r) K_l^l(\vec{r}^o) . \quad (9)$$

We proceed in the usual manner and make the expectation value of the energy stationary under the constraints of normalization and continuity of the wave function and its derivative, as described by eqs. (5) and (6). These conditions are incorporated by means of Lagrange multipliers $-E_k$, α_{lm} and β_{lm} , respectively. We get the following secular equations

$$\begin{aligned} \sum_j (H_{jj} - E_k) j_{jj} + \sum_n \sum_m^L H_{j,nlm} A_{nlm} - \sum_m^* Y_{l,-m}^*(\vec{k}_j^o) j_l(k_j r_i) \alpha_{lm} \\ - \sum_m^L Y_{l,-m}^*(\vec{k}_j^o) k_j J'_l(k_j r_i) \beta_{lm} = 0 \end{aligned} \quad (10)$$

for all j ;
and for all n,l,m

$$\sum_j H_{nlm,n'l'm'} j^v_j + \sum_n \sum_{l'=0}^L (H_{nlm,n'l'm'} - E_k) \delta_{ll'} \delta_{mm'} \Omega_{nlm,n'l'm'} \\ \times A_{n'l'm'} + R_{nl}(r_i) \alpha_{lm} + R'_{nl}(r_i) \beta_{lm} = 0 \quad . \quad (11)$$

The following abbreviations have been used:

$$\Omega_{jj'} = \Omega_o \delta_{jj'} \\ - 4\pi \sum_{l=0}^L (2l+1) P_l(\vec{k}_j^o \cdot \vec{k}_{j'}^o) \int_0^{r_i} j_l(k_j r) j_l(k_{j'} r) r^2 dr \quad , \quad (12)$$

$$\Omega_{nlm,n'l'm'} = 4\pi(2l+1) \frac{(l+m)!}{(l-m)!} \int_0^{r_i} R_{nl}(r) R_{n'l'}(r) r^2 dr \quad , \quad (13)$$

$$H_{jj'} = \frac{1}{2} (k_j^2 + k_{j'}^2) \Omega_{jj'} + 4\pi \int_0^{r_i} V_{sph}(r) \left[J_o(|\vec{k}_j - \vec{k}_{j'}| r) \right. \\ \left. - \sum_{l=0}^L (2l+1) P_l(\vec{k}_j^o \cdot \vec{k}_{j'}^o) j_l(k_j r) j_l(k_{j'} r) \right] r^2 dr + \Omega_o \tilde{V}(\vec{k}_j - \vec{k}_{j'}) \\ - \left(\sum_{l=0}^L \sum_{m'=0}^L + \sum_{l=0}^L \sum_{m'=L+1}^{\infty} + \sum_{l=L+1}^{\infty} \sum_{m'=0}^L \right) S_{jlm,j'l'm'} \quad , \quad (14)$$

$$H_{j,n'l'm'} = H_{n'l'm',j}^* = (2l'+1) i^{l'+2m'} \sum_{l=L+1}^{\infty} (2l+1) i^{l+2m} Y_{l,-m}(\vec{k}_j^o) \\ \times \sum_{l''>0} \int_{l''}^* Y_{l'm} K^{l''} d\omega \int_0^{r_i} j_l(k_j r) R_{n'l'}(r) \tilde{V}(l'', r) r^2 dr \quad , \quad (15)$$

$$H_{nlm,n'l'm'} = 2\pi(2l+1) \frac{(l+m)!}{(l-m)!} \delta_{ll'} \delta_{mm'} (E_{nl} + E_{n'l'}) \\ \times \int_0^{r_i} R_{nl}(r) R_{n'l'}(r) r^2 dr + \sum_{l''>0}^{l+l'} (2l+1)(2l'+1) i^{l'-l+2(m'-m)}$$

$$\times \int_{\Omega_0} Y_{lm}^* Y_{l'm'}^{K''} d\omega \int_0^{r_i} R_{nl}(r) R_{n'l'}(r) \tilde{V}(l'', r) r^2 dr , \quad (16)$$

$$S_{jlm, j'l'm'} = \sum_{l''=0}^{\infty} (2l+1)(2l'+1) i^{l'-l+2(m'-m)} Y_{l,-m}^*(k_j^o) Y_{l',-m'}(k_j^o) \\ \int_0^{r_i} j_l(k_j r) R_{n'l'}(r) \tilde{V}(l'', r) r^2 dr . \quad (17)$$

Ω_0 denotes the volume of the atomic cell.

The surface integrals extend over the unit sphere and can be evaluated using Clebsch-Gordan coefficients. Almost all sums are finite. The exceptions are found in eqs. (14) and (15), where several sums run from $L+1$ to ∞ . However, since the individual terms contain one of the integrals

$$\int_0^{r_i} j_l(k_j r) R_{n'l'}(r) \tilde{V}(l'', r) r^2 dr \quad \text{or} \quad \int_0^{r_i} j_l(k_j r) j_{l'}(k_j r) \tilde{V}(l'', r) r^2 dr$$

which in most cases turns out to diminish rapidly with growing values of l , convergence is ensured /14/.

D. Some Special Properties of the Formalism

Because of the fact that the radial functions $R_{nl}(r)$ do not depend on the eigenvalue E_K , there is no implicit dependence of the matrix elements on E_K . Therefore it is not necessary to examine the secular determinant as a function of the energy in order to find the zeros. We obtain a general eigenvalue problem of the form

$$A \vec{x} = E_K B \vec{x}$$

or, more explicitly, the following structure

$$\begin{pmatrix} H_{11} & H_{12} & C_1^+ & C_2^+ \\ H_{12}^+ & H_{22} & C_1 & 0 \\ C_1 & C_2 & 0 & \end{pmatrix} \begin{pmatrix} v \\ A \\ \alpha \\ B \end{pmatrix} = E \begin{pmatrix} \Lambda_1 & 0 & 0 \\ 0 & \Lambda_2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} v \\ A \\ \alpha \\ B \end{pmatrix} \quad (18)$$

where the submatrices H_{ij} generate the matrix element $(\psi H \psi)$, the Λ_i guarantee the normalization of ψ , and where finally the C_i

force the wave function and its gradient to be continuous.

The rank of this eigenvalue problem may actually be reduced. Of course, the goal is to retain a linear eigenvalue problem, because nonlinear eigenvalue problems can only be solved by directly searching for the zeros of the secular determinant. In our case, it is possible to eliminate twice as many equations as the vectors of the Lagrange multipliers α_{lm} and β_{lm} have components. The elimination can be done such that the Hermitean property of both matrices involved is conserved. Since all rows of C_1 and C_2 are linearly independent - otherwise any energy E would solve the eigenvalue problem - we can take a sequence of Gauss elimination steps to remove the submatrices C_1 and C_2 . These steps cannot change any element of C_1^+ or C_2^+ . Therefore it is possible to perform the transposed elimination steps for C_1^+ and C_2^+ . It would not be suitable to carry out the elimination analytically in our formulae since any of the arising denominators may vanish or may become very small leading to numerical risks. As a consequence of the elimination process all zero eigenvalues of the non-negative semidefinite matrix on the right hand side of eq. (18) are removed, so that its Cholesky decomposition /15/ becomes possible, and we get a Hermitean eigenvalue problem of standard form $D \vec{y} = E_K^- \vec{y}$. For most computers subroutines are available which yield, with high speed and accuracy, all energy eigenvalues as well as eigenvectors in a given energy interval simultaneously.

If we take as a realistic example a \vec{k} -point without any symmetry we have for instance

$$L=2, J=51 \text{ and } N=3$$

leading to a total rank of 96. After the reduction the rank will be 60.

Given a spherically symmetric potential for $r \leq r_i$, the APW method leads to exact solutions of the Schrödinger equation inside the sphere for the eigenvalue E_k^- , which in practice, however, can only be achieved iteratively. In the present version the partial waves both for $l \leq L$ and $l > L$ generally do not have this property, much like in the APW method for a non-spherical potential in $r \leq r_i$. Despite this fact our trial functions have the correct r -dependence near the atomic site, where

$$|v_{\text{sph}}(r) + l(l+1)/r^l| \gg E_K^-$$

The formalism developed above yields both valence states and core states which are orthogonal to one another. In the case of a low-lying core state an essential simplification is possible since the wave function falls off very rapidly inside the APW sphere and can be regarded to be zero at its surface. As a consequence the coefficients v_j and the Lagrange multipliers will turn out to be negligibly small. Furthermore, in the cases

of Ni, Cu, and Al we found that by choosing the C_1 's as described above, only one of the coefficients A will be essentially different from zero. This core state may approximately be determined by solving the Schrödinger equation of the corresponding atomic problem with the actual crystal potential. This procedure leads to \vec{k} -indepent wave functions which are orthogonal both to one another and to the valence functions. Noticing this property we can omit those partners R_{nl} which represent the low-lying core states in the ansatz for the valence states.

E. The Variational Principle for a Nonlocal Potential

The extension of the method to the case of a nonlocal potential in order to include many-body aspects has been investigated by two of the present authors (/8/,/11/). We briefly present the ideas and the results.

In the equation of motion for quasi-particles,

$$i \frac{\partial \psi}{\partial t} = (-\Delta + \hat{V}(\vec{r})) \psi_{\vec{k}}(\vec{r}, t) + \int M(\vec{r}, t, \vec{r}', t') \psi_{\vec{k}}(\vec{r}', t') d\vec{r}' dt' ,$$

the mass operator M describes the effects of exchange and correlation, whereas \hat{V} denotes the potential of the nuclei.

With several approximations, namely

- (i) the random phase approximation,
- (ii) an instantaneous interaction,
and
- (iii) the limit of long wave lengths,

we get the following equation with a screened exchange

$$E_{\vec{k}} \psi_{\vec{k}} = (-\Delta + \hat{V}(\vec{r}) + 2 \int \frac{S(\vec{r}')}{|\vec{r}-\vec{r}'|} d\vec{r}') \psi_{\vec{k}}(\vec{r}) - 2 \sum_{\vec{k}'}^{\text{occ.}} \psi_{\vec{k}'}^*(\vec{r}) \int \psi_{\vec{k}'}^*(\vec{r}') \frac{e^{-\alpha |\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|} \psi_{\vec{k}}(\vec{r}') d\vec{r}' , \quad (19)$$

$$\text{with } S(\vec{r}) = \sum_{\vec{k}'}^{\text{occ.}} \psi_{\vec{k}'}^*(\vec{r}) \psi_{\vec{k}'}(\vec{r}) \quad (20)$$

$$\text{and } \alpha = \left(\frac{8\pi}{\Omega_o} N(E_F) \right)^{1/2} , \quad (21)$$

where $N(E_F)$ is the density of states at the Fermi energy. Regarding α as a free parameter, the limit $\alpha = +\infty$ yields the Hartree approximation, whereas $\alpha = 0$ leads to the Hartree-Fock approximation. In the view of the restriction to long wave lengths the value of α , eq. (21) should be considered as an upper limit.

As the integro-differential equation is nonlinear, we linearize it by inserting wave functions of a lower order for the functions $\psi_{\vec{k}}$, in eqs. (19) and (20).

For the numerical calculations we then replace eq. (19) by the equivalent variational principle. We obtain an eigenvalue problem with the same structure as in the case of the local potential. Only the matrix elements $H_{z,z'}$ are changed because terms $A_{z,z'}$, which are defined in /8/, have to be added, z and z' standing for j or nlm .

III. Results

A. Local Potential

As the computation time is one of the interesting features of a band structure calculation, in Tab. 1 we give two figures for the TR 4 where the calculations have been performed and the corresponding estimates for the IBM 7090. These times include the simultaneous determination of all the eigenvalues and the wave functions of the matrix D for a fixed \vec{k} . However, no account is taken of several expressions such as e.g. the Fourier coefficients, which can be calculated once for all.

Tab. 1 Computation time

Parameters	Rank of matrix A	Rank of matrix D	our computer TR 4	corresp. IBM 7090
L=2, N=3, J=51	96	60	100 sec	30 sec
L=2, N=2, J=51	85	51	70 sec	20 sec

Using the muffin-tin potentials of Hanus /16/, Burdick /17/, and Segall /18/ the bandstructure and the wave functions were determined for paramagnetic Ni, Cu, and Al in the $\langle 001 \rangle$, $\langle 110 \rangle$, and $\langle 111 \rangle$ directions of \vec{k} -space limiting ourselves to L=2. In order to test the convergence the number of partners R_{nl} was increased from 2 to 7 and the number of plane waves from

15 to more than 100. As a typical example we show in the first three columns of Tab. 2 a \vec{k} -point in the middle of the Δ_1 -line of the first conduction band of Cu. We remark that we achieve an accuracy better than .005 Ry if we restrict ourselves to 2 n-values and 51 plane waves. The same accuracy was established in all cases considered.

Comparing our results with those of Burdick and Segall we found that the deviations of the energy values for Cu were constantly within the limits of accuracy of .01 Ry and less than .008 Ry for Al.

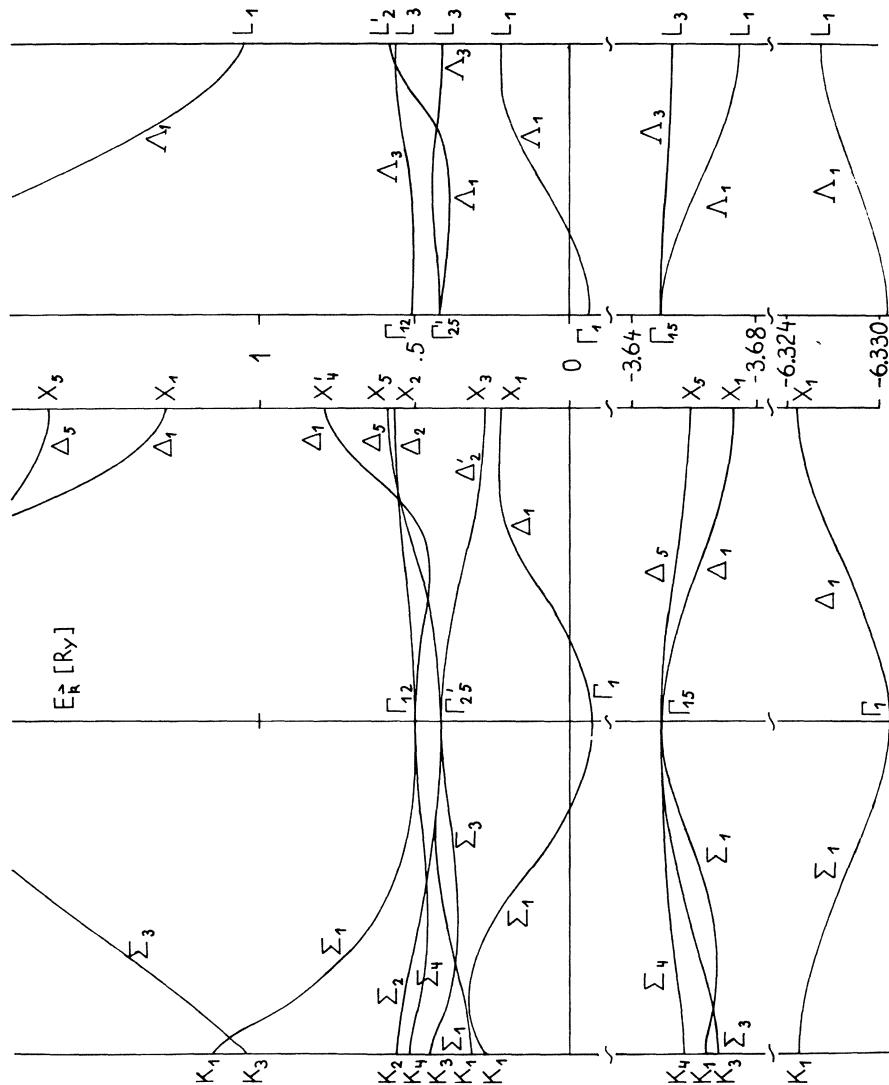
For Ni the \vec{k} -dependence of the energies for the three directions of high symmetry as determinated by the MAPW procedure is plotted in Fig. 1. As compared to Hanus' results we find for the conduction electrons throughout somewhat lower energy values. These deviations show a weak \vec{k} -dependence and are most pronounced at Γ_1 with .05 Ry. On the average these deviations are about .02 Ry. In the view of the excellent agreement in the cases of Cu and Al we have some confidence in the Ni results, yet cannot understand the discrepancy between our results and those of Hanus.

We also calculated the highest lying core bands. In Tab. 3 we list the energies for $\vec{k}=0$, $E_{\vec{k}=0}$, and the band widths, ΔE .

Tab. 2

Dependence of $E_{\vec{k}}$ and ϵ_{III}^2 on the number of plane waves, J, and the number of partners R_{nl}^{III} for a fixed l, N. L=2.

1	2	3	4
J	N	$E_{\vec{k}} \text{ [Ry]}$	$\epsilon_{III}^2 \text{ [Ry}^2]$
15	2	-.797439	.051
	3	-.845422	.021
	4	-.846226	.019
	5	-.846473	.022
	6	-.846584	.026
	7	-.846700	.030
51	2	-.845170	.0333
	3	-.846997	.0073
	4	-.847188	.0050
	5	-.847233	.0055
	6	-.847253	.0065
	7	-.847273	.0070
89	4	-.847309	.0046
	5	-.847371	.0013
	6	-.847376	.0011
	7	-.847378	.0012
	8	-.847380	.0010
113	4	-.847326	.00430
	5	-.847384	.00064
	6	-.847389	.00029
	7	-.847390	.00029



\vec{k} versus \vec{k} along the $<110>$, $<001>$, and $<111>$ directions in the Brillouin zone in Ni. The representation labels are in the usual BSW notation.

Tab. 3

Energies at T' , E_F , and band widths, ΔE , for the highest core bands of Ni, Cu, and Al.

Metal	Ni		Cu		Al	
corresp. atomic level	3s	3p	3s	3p	2s	2p
E_F (in Ry)	-6.9248	-4.2436	-8.3722	-5.4248	-7.0952	-4.2420
ΔE (in Ry)	.006	.02	.002	.01	.0003	.0018

Since it is well known that variational principles need not yield a reasonable approximation for the true eigenvectors even if the eigenvalue is converging well, we have to look for a criterium to judge the accuracy of the wave functions.

An obvious criterium would be

$$\sigma_I = \int_{\vec{K}} \psi_{\vec{K}}^* (H - E_{\vec{K}}) \psi_{\vec{K}} d\tau / \int_{\vec{K}} \psi_{\vec{K}}^* \psi_{\vec{K}} d\tau$$

since the value of this expression will vanish for the exact solution. However, it is easy to verify that this is true for any trial function.

Another test for the accuracy of the wave function would be to evaluate

$$\sigma_{II}(\vec{r}) = \frac{H\psi_{\vec{K}}(\vec{r})}{\psi_{\vec{K}}(\vec{r})} - E_{\vec{K}}$$

at every point of the cell. In regions where the wave function $\psi_{\vec{K}}$ is small a slight deviation of $\psi_{\vec{K}}$ from the exact value will lead to a large $\sigma_{II}(\vec{r})$. However, when calculating matrix elements these regions generally are expected to give negligible contributions. Therefore σ_{II} does not seem to be an appropriate measure. A more adequate criterium for the wave function seems to be the expectation value of the operator $(H-E)^2$. The natural generalization to cases where the higher derivatives of the wave function or the potential are discontinuous is the expression

$$\sigma_{III}^2 = \left(\int (H\psi_{\vec{K}})^* (H\psi_{\vec{K}}) d\tau - E_{\vec{K}}^2 \int \psi_{\vec{K}}^* \psi_{\vec{K}} d\tau \right) / \int \psi_{\vec{K}}^* \psi_{\vec{K}} d\tau$$

$$= \int |(H-E_{\vec{K}})\psi_{\vec{K}}|^2 d\tau / \int \psi_{\vec{K}}^* \psi_{\vec{K}} d\tau$$

In column (4) of Tab. 2 the values σ_{III}^2 are given for the same cases as considered above. It is remarkable that a rapid decrease of σ_{III}^2 is only achieved if the values of N and J are chosen in a well-balanced manner.

B. Nonlocal Potential

In the case of Lithium we did calculations which took into account the many body aspect as described above.

The linearization of the integro-differential equation (19) has been achieved by using core functions as determined by Herman and Skillman /19/ and using single OPW's for the wave functions of the conduction band.

The energy eigenvalues and the wave functions were determined for four values of α along the $\langle 001 \rangle$, $\langle 110 \rangle$, and $\langle 111 \rangle$ directions of k-space. In Tab. 4 some energy differences are listed, together with the density of states at the Fermi level,

Tab. 4

Comparison of energy differences, densities of states, and band widths of the conduction band

	MAPW				Ham	Kenney
	.5 α_{TF}	.7 α_{TF}	.9 α_{TF}	$\sqrt{2} \alpha_{TF}$		
$\Gamma_{15} - \Gamma_1$	1.5235	1.4502	1.4083	1.3616		1.2616
$\Gamma_{25}' - \Gamma_1$		1.7618	1.6931			1.7210
$\Gamma_{12} - \Gamma_1$	1.8823	1.8063	1.7618	1.7090		1.7840
$\Gamma_{25} - \Gamma_1$		1.8295	1.7850			1.8206
$H_{15} - \Gamma_1$.8113	.7456	.7093	.6695	.628	.6940
$H_{12} - H_{15}$.1447	.1441	.1436	.1420	.208	.1982
$H_1 - H_{15}$.5808	.5807	.5823	.5917		.3049
$P_4 - \Gamma_1$.6341	.5750	.5427	.5075	.495	.5593
$P_1 - P_4$.3622	.3615	.3626	.3701	.317	.1622
$N_1' - \Gamma_1$.3809	.3335	.3080	.2798	.271	.3392
$N_1' - N_1$.2126	.2117	.2117	.2149	.209	.1254
$\frac{N(E_F)}{N_o(E_F)}$	1.36	1.45	1.58	1.70	1.67	
ΔE_F	.347	.305	.283	.258	.252	.3016

All energies are measured in Ry. The densities of states are normalized by the free electron density.

$N(E_F)$, and the band width of the conduction band, ΔE_F . These are compared with the results of Ham /20/ and Kenney /21/.

The screening parameter is written in the form $\kappa = \alpha \kappa_{TF}$, where κ denotes the Thomas-Fermi screening constant.

Measurements of the electronic specific heat by Martin /22/ yield a value of about 2 for $N(E_F) / N_0(E_F^0)$. Because of the martensitic transformation which Lithium undergoes at low temperatures, the quoted figure is very uncertain. Furthermore, a contribution by the electron-phonon interaction is included which was estimated by Grimvall /23/ to be 19 percent. Thus a likely prediction for the value of $N(E_F) / N_0(E_F^0)$ should be 1.6 which would correspond to $\kappa = 0.9 \kappa_{TF}$.

The energy difference between the Fermi level and the (almost) \vec{k} -independent core level can be deduced from measurements of the K-edge of the photoelectric absorption /24/. The experimental value of 4.01 Ry corresponds to $\kappa = 0.55 \kappa_{TF}$.

IV. Summary

A method for the calculation of the energy eigenvalues and the wave functions of solids has been presented, allowing for a general potential and yielding the wave functions of the valence and core bands orthogonal to one another. This may be achieved in the APW method as well.

In the present formalism, the wave functions and their first gradients are continuous everywhere in the cell. As a consequence the formalism is more complicated than the APW method. However, we hope to obtain better wave functions at the cost of an increase in the computation time. The method has been extended to take into account many body aspects.

Results have been presented for Li only. For Al, Cu and paramagnetic Ni we have done a large number of tests which partially have been reported. The calculation for Cu and Ni are still performed.

References

- 1 J.C. Slater, Phys.Rev. 51, 846 (1937)
- 2 T.L. Loucks, Augmented-Plane-Wave Method, Benjamin, N.Y. (1967)
- 3 R.S. Leigh, Proc.Phys.Soc. A 69, 388 (1956)
- 4 H. Schlosser and P.M. Marcus, Phys.Rev. 131, 2529 (1963)
- 5 H. Bross, Phys.kondens.Materie 3, 119 (1964)
- 6 P.D. DeCicco, Ph.D.thesis, Dep. of Physics, MIT, 1965

- 7 P.M. Marcus, Int.J.Quant.Chem. 1S, 567 (1967)
8 H. Bross, Z.Physik 215, 485 (1968); ibid. 218, 109 (1969)
9 H. Bross, Helvetica Physica Acta 41, 717 (1968)
10 W.E. Rudge, Phys.Rev. 181, 1024, (1969)
11 G. Bohn, to be published
12 H. Bross, G. Bohn, G. Meister, W. Schubö, H. Stöhr,
Phys.Rev., to be published
13 See e.g. D.G. Bell, Rev.Mod.Phys. 26, 311 (1954)
14 If this should happen not to be the case an alternative
approach based on the Fourier expansion of $V(r)$ is to be
preferred.
15 See e.g. Schwarz, Rutishauser, Stiefer, Numerik symmetri-
scher Matrizen, Teubner, Stuttgart, 1968
16 J.G. Hanus, MIT Solid State and Molecular Theory Group,
Quart.Progress Report 44, 29 (1962), unpublished
17 G.A. Burdick, Phys.Rev. 129, 138 (1963)
18 B. Segall, Phys.Rev. 124, 1797 (1961)
19 F. Herman, S. Skillman: Atomic Structure Calculations,
Prentice Hall, 1963
20 S.H. Ham, Phys.Rev. 128, 82 + 2524 (1962)
21 J.F. Kenney, Quart.Prog.Rep. 53, (1954), SSMTG, MIT,
Cambridge
22 D.L. Martin, Phys.Rev. 124, 438 (1961)
23 G. Grimvall, Phys.kondens.Materie 6, 15 (1967)
24 D.J. Baker, D.H. Tomboulian, Phys.Rev. 128, 677 (1962)

GRADIENTS OF $E(\vec{k})$ FROM THE APW DETERMINANT

J. H. Wood

University of California
Los Alamos Scientific Laboratory
Los Alamos, New Mexico 87544

APW METHOD PROPER

Because our method of determining $\nabla_k(E(\vec{k}))$ was partially determined by the programming system we use for the 1957 APW method,^{1,2} we first briefly describe the methods we used to determine $E(\vec{k})$. In this scheme, the $E(\vec{k})$ are taken as the zeroes of the so-called APW determinant $D(E, \vec{k})$:

$$D(E(\vec{k}), \vec{k}) = 0 \quad (1)$$

The matrix M , of which D is the determinant, has as its elements

$$(H - E)_{ij} = \vec{k}_i \cdot \vec{k}_j \left[\delta(\vec{k}_i, \vec{k}_j) - \frac{4\pi R_s^2}{\Omega} \frac{j_1(|\vec{k}_j - \vec{k}_i| R_s)}{|\vec{k}_j - \vec{k}_i|} \right] \\ - E \left[\delta(\vec{k}_i, \vec{k}_j) - \frac{4\pi R_s^2}{\Omega} \cdot \frac{j_1(|\vec{k}_j - \vec{k}_i| R_s)}{|\vec{k}_j - \vec{k}_i|} \right] \\ + \frac{4\pi}{\Omega} \sum_{\ell} [(2\ell + 1) P_{\ell}(\cos \theta_{ij}) j_{\ell}(k_i R_s) j_{\ell}(k_j R_s)] R_s^2 \frac{u'_{\ell}(R_s, E)}{u_{\ell}(R_s, E)} \quad (2)$$

This is the matrix element of the operator $(H - E)$ between an APW of wave vector $\vec{k}_i = \vec{k} + \vec{K}_i$ and energy E and another APW of wave vector \vec{k}_j and energy E . H is the muffin tin Hamiltonian, Ω is the unit cell volume, R_s is the APW sphere radius and the u'_{ℓ}/u_{ℓ} are the (energy dependent) logarithmic derivatives for energy E evaluated at R_s . $u'_{\ell}/u_{\ell} \equiv [u'_{\ell}(r; E)/u_{\ell}(r; E)] r = R_s$. This scheme contrasts with the 1953 APW method⁴ in which the Bloch function is expanded in terms of APW's having different (fixed) energies. In that

method,⁵ it is necessary to construct two matrices, one of H and the other of ES , where S is the overlap of the (non-orthogonal) APW's.

Because the trial APW's in the 1937 APW method all have the energy parameter set to the trial E , it is convenient to set up only the single matrix of $(H-E)$. If one is interested only in $E(\vec{k})$ (and not the associated eigenvector), one needs only a subroutine which evaluates determinants. To accomplish this, we have employed Gaussian elimination with complete pivoting.² (With some further computation, this method can also yield the eigenvector if that is desired.)

We have belabored these points to show that, in general, we compute only one $E(\vec{k})$ at a time and have available (at most) only the eigenvector associated with a single $E(\vec{k})$. In the 1953 APW scheme, one solved for a complex of, say, n bands and had available a full $n \times n$ eigenvector matrix instead of just a single column.

COMPUTATION OF THE GRADIENT

In order to make as much use as possible of the existing computational machinery for the 1937 method, we attempted to obtain the gradient of $E(\vec{k})$ without any knowledge of the eigenvectors. Consider a k point of no symmetry (no sticking together of bands). Suppose we have determined a single band energy $E(\vec{k})$ so that $D(E(\vec{k}), \vec{k}) = 0$. Then the energy in that band at a point δk_x distant will change (by δE) from $E(\vec{k})$, so that the determinant remains zero:

$$0 = \delta E \frac{\partial D}{\partial E} + \delta k_x \frac{\partial D}{\partial k_x}$$

or

$$\delta E / \delta k_x = - (\partial D / \partial k_x) / (\partial D / \partial E). \quad (3)$$

This is the equation we have programmed. In order to evaluate the derivatives of the determinant, we require the derivatives of the matrix elements of Eq. (2). The one piece of information needed (beyond that in mathematical handbooks) is the energy derivatives of the logarithmic derivatives. These are:⁵

$$\frac{\partial}{\partial E} \left(\frac{u_\ell^i(r; E)}{u_\ell^j(r; E)} \Big|_{r=R_s} \right) = - \frac{1}{R_s^2 u_\ell^2(R_s; E)} \int_0^{R_s} r^2 u_\ell^2(r; E) dr \quad (4)$$

Then, the energy derivative of a matrix element is given by

$$\begin{aligned} \frac{\partial (H - E)_{ij}}{\partial E} &= - \left[\delta(\vec{k}_i, \vec{k}_j) - \frac{4\pi R_s^2}{\Omega} \frac{j_1(|\vec{k}_j - \vec{k}_i| R_s)}{|\vec{k}_j - \vec{k}_i|} \right] \\ &\quad - \frac{4\pi}{\Omega} \sum_\ell [(2\ell + 1) P_\ell(\cos \theta_{ij}) j_\ell(k_i R_s) j_\ell(k_j R_s)] \frac{1}{u_\ell^2(R_s, E)} \int_0^{R_s} r^2 u_\ell^2 dr. \end{aligned} \quad (5)$$

Denoting the x component of \vec{k}_i as k_{ix} , we also have

$$\begin{aligned} \frac{\partial(H - E)_{ij}}{\partial k_x} = & [k_{ix} + k_{jx}] \left[\delta(\vec{k}_i, \vec{k}_j) - \frac{4\pi R_s^2}{\Omega} \frac{j_1(|\vec{k}_j - \vec{k}_i| R_s)}{|\vec{k}_j - \vec{k}_i|} \right] \\ & + \frac{4\pi}{\Omega} \left\{ \sum_{\ell} (2\ell + 1) P_{\ell}(\cos \theta_{ij}) [j_{\ell}(k_i R_s) j'_{\ell}(k_j R_s) k_{jx}/k_j \right. \\ & \quad \left. + j'_{\ell}(k_i R_s) j_{\ell}(k_j R_s) k_{ix}/k_i] \cdot R_s^3 \frac{u'_{\ell}(R_s, E)}{u_{\ell}(R_s, E)} \right. \\ & \quad \left. + \frac{[(k_{ix} + k_{jx}) - (\vec{k}_i \cdot \vec{k}_j)(k_{jx}/k_j^2 + k_{ix}/k_i^2)]}{k_i k_j} \right. \\ & \quad \left. \times \sum_{\ell} [(2\ell + 1) j_{\ell}(k_i R_s) j_{\ell}(k_j R_s) P'_{\ell}(\cos \theta_{ij})] R_s^2 \frac{u'_{\ell}(R_s, E)}{u_{\ell}(R_s, E)} \right\}. \end{aligned} \quad (6)$$

In the program, we first determine an $E(\vec{k})$ (which is then the energy parameter for Eqs. (5) and (6)) and then we evaluate the four derivatives $\partial D/\partial E$, $\partial D/\partial k_x$, $\partial D/\partial k_y$, and $\partial D/\partial k_z$; the gradient is then calculated from Eq. (3). If the matrix M (of which D is the determinant) is $n \times n$, each derivative is the sum of $n \times n$ determinants. The first is the determinant of a matrix which is M with the first column (row) elements replaced by their derivatives; the second is the determinant of M with a similar replacement of the second column (row) elements, and so on.

This means that determination of the three components of the gradient involves the evaluation of $4n \times n \times n$ determinants, which is the chief drawback of the method. In a test with $n=34$, we found the gradient computation required 18.5 seconds per energy; finding the energies themselves from the zeroes of D required about 4 seconds per energy. At this writing it appears that the gradients are determined to the same precision as are the eigenvalues themselves.

USE OF THE METHOD

We set up the gradient calculation to provide input information for a Gilat-Raubenheimer⁶ density-of-states program. We had envisaged using the APW calculated energies and gradients directly; it now appears that this will be rather costly simply because the G-R scheme requires a rather fine grid in k -space. If one thinks of Janak's techniques⁷ of fitting to calculated $E(\vec{k})$ with an interpolation function, one might well now fit not only to the $E(\vec{k})$ but also demand a fit to the gradients. We plan to further

explore these ideas.

The gradients may be useful within the context of the APW program itself. At present, most APW codes fix \vec{k} and search the determinant D over a considerable number of energies. Each \vec{k} point is treated as a separate problem with no use made of previously determined $E(\vec{k})$. With the gradient available, it may well pay to estimate the energies at the new point from the energies and gradients at a previously determined point. This would greatly reduce the number of energies at which D has to be evaluated and this might more than pay for the expense of calculating the gradients.

REFERENCES

1. J. C. Slater, Phys. Rev. 51, 846 (1937).
2. L. F. Mattheiss, J. H. Wood, and A. C. Switendick in Methods In Computational Physics, Academic Press, New York (1968) give an exhaustive description of this method.
3. We restrict ourselves to a crystal containing one atom at the origin of each unit cell.
4. J. C. Slater, Phys. Rev. 92, 603 (1953); M. M. Saffren and J. C. Slater, Phys. Rev. 92, 1126 (1953).
5. D. J. Howarth, Phys. Rev. 99, 469 (1955).
6. G. Gilat and L. J. Raubenheimer, Phys. Rev. 144, 390 (1965).
7. J. F. Janak, Physics Letters 28A, 570 (1969).

APW PSEUDOPOTENTIAL FORM FACTORS FOR THE ALKALI METALS

Martin J. G. Lee

The James Franck Institute and Department of Physics

The University of Chicago, Chicago, Illinois 60637

ABSTRACT

Several authors have adapted the techniques of band structure calculation to analyze experimental Fermi surface data. The secular determinant that arises in a first principles band structure calculation is expressed in terms of suitably-chosen parameters, and these are adjusted to fit the data. Two alternative methods have been developed and exploited; the OPW pseudopotential method is based on the OPW secular determinant, while the phase shift method is based on the APW secular determinant. The relationship between these two methods in the analysis of experimental Fermi surface data is discussed. The APW pseudopotential form factors that describe the electron-ion interaction in the alkali metals are presented, and are compared with form factors derived from OPW based pseudopotentials.

INTRODUCTION

Several authors have adapted the techniques of band structure calculation for the interpretation of experimental Fermi surface data. One of the earliest of such calculations was by Ashcroft (1), who showed that the Fermi surface data for aluminum can be explained quantitatively in terms of a local pseudopotential. Subsequently, it has proved possible to interpret many of the physical properties of aluminum, among them the phonon spectrum (2), the electron-phonon mass enhancement (3), the resistivities of solid (4) and liquid (5), the crystal structure (6), and the superconducting transition temperature (7), on the basis of Ashcroft's pseudopotential. Recently the pseudopotential method has assumed some importance (8) in attempts to correlate

the physical properties of metals.

It is now known that a local pseudopotential analysis also gives a good description of the shapes of the Fermi surfaces of sodium and magnesium, the two elements that precede aluminum in the Periodic Table. It is nevertheless a matter of coincidence that the local pseudopotential approximation works so well for this group of metals. Ashcroft (9) has found that the shapes of the Fermi surfaces of potassium and rubidium can be interpreted in the local pseudopotential approximation only if very large V_{200} pseudopotential coefficients are assumed, while Lee and Falicov (10) have argued that the potassium data can be interpreted more naturally as an indication that the Fermi surface is significantly perturbed by low-lying d-like energy bands, an effect that can only be described by a nonlocal (angular-momentum dependent) pseudopotential. It is now believed that, even for simple metals, non-locality is an essential feature of the pseudopotentials that describe Fermi surface distortions.

Two distinct methods of constructing a nonlocal pseudopotential have been developed and exploited. The OPW pseudopotential, based on a transformation (11) of the Schrödinger equation that arises in the OPW method of band structure calculation, has been used with success to analyze experimental Fermi surface data for magnesium (12), zinc (13), and cadmium (14). More recently the method of phase shift analysis, based on the APW secular determinant, has been applied to interpret experimental data for the alkali metals (15). The shapes of the Fermi surfaces of these metals are well described by phase shift analysis, and the phase shifts so derived are related in a simple way to the position of each metal in the Periodic Table. It appears that, at least for the simple metals, OPW pseudopotential analysis and phase shift analysis are equivalent techniques by which to interpret Fermi surface data, although no direct comparison of the two techniques has so far been made.

In the present paper we discuss the method of phase shift analysis within the context of pseudopotential theory. We begin by outlining some relevant aspects of pseudopotential theory, stressing the role of screening in determining the low- q limit of the pseudopotential form factor. We find that by making an appropriate choice of the Fermi energy parameter in a phase shift analysis, one can generate an APW pseudopotential form factor that satisfies the condition of charge neutrality in the metal. In this way we demonstrate that equivalent OPW and APW pseudopotential form factors may be derived from experimental Fermi surface data. We conclude by presenting and discussing the APW pseudopotential form factors that we have obtained by phase shift analyses of the shapes of the Fermi surfaces of the alkali metals.

THE PSEUDOPOTENTIAL METHOD

The simplest picture of electron states in a metal is provided by the free-electron model. The lattice potential is assumed to be infinitesimally weak, and the energies of electron states are given by

$$E(\mathbf{k}) = (\frac{e^2}{\hbar^2} \frac{\mathbf{k}^2}{2m}). \quad (1)$$

The structural periodicity of the lattice potential implies that wave vectors which differ only by a reciprocal lattice vector are physically equivalent, and it is convenient to reduce the parabolic energy bands to the first Brillouin zone. The Fermi surfaces of the alkali metals lie entirely within the first Brillouin zone. It follows that, on the free-electron model, they are spherical.

Over the past five years, experimental studies (16-19) have shown that the shapes of the Fermi surfaces of the alkali metals are significantly aspherical. The measured area distortions normal to symmetry directions are set out in Table I. The shapes of the distorted surfaces depend on the interaction

TABLE I.

Experimental area distortions of the Fermi surfaces of the alkali metals normal to symmetry directions.

	$10^4(A_{100}-A_0)/A_0$	$10^4(A_{110}-A_0)/A_0$	$10^4(A_{111}-A_0)/A_0$
Li ^a	+ 90 ± 85	- 110 ± 85	+ 340 ± 85
Na ^b	- 2.04 ± 0.20	- 2.00 ± 0.20	+ 7.15 ± 0.40
K ^c	+14.15 ± 0.44	- 4.65 ± 0.22	+ 2.36 ± 0.12
Rb ^d	+ 38.4 ± 2.2	- 27.9 ± 2.2	+ 56.5 ± 2.2
Cs ^e	+ 110 ± 10	- 87 ± 10	+ 214 ± 10

^aReference 16

^dReference 18

^bReference 17

^eReference 19

^cReferences 18, 10

between the conduction electrons and the ion cores; pseudopotential techniques have been applied (10, 15, 17) to yield information about the electron-ion interaction in these metals.

In the pseudopotential method one associates with each lattice site a weak effective potential $U_{NL}(\underline{r})$ that gives the same conduction-electron energy bands as the true ionic potential (20, 21). The eigenfunctions of the Schrödinger equation become pseudo wave functions, from which the atomic-like oscillations in the region of each ion core have been smoothed away. There is some arbitrariness in the construction of such a pseudopotential, and an optimized pseudopotential may be defined, whose pseudo wave function is smoothed to the greatest possible extent. The pseudopotential defined in this way has one drawback: it may depend significantly on energy and on angular momentum. The energy dependence is of no consequence in Fermi surface studies, since the experimental data define the shape of a surface of constant energy. The angular-momentum dependence may be expressed by writing the pseudopotential $U_{NL}(\underline{r})$ in the form

$$U_{NL}(\underline{r} - \underline{r}_\alpha) = \sum_\ell \hat{P}_\ell(\underline{r}_\alpha) U_\ell(\underline{r} - \underline{r}_\alpha) \hat{P}_\ell(\underline{r}_\alpha) \quad (2)$$

where \underline{r}_α labels the center of the α th unit cell in the crystal, and where the projection operator $\hat{P}_\ell(\underline{r}_\alpha)$ selects the ℓ th spherical harmonic component of the wave function on which it operates. The lattice pseudopotential $V_{NL}(\underline{r})$ is then obtained by summing the ionic pseudopotentials associated with each lattice site in the metal

$$V_{NL}(\underline{r}) = \sum_\alpha U_{NL}(\underline{r} - \underline{r}_\alpha) . \quad (3)$$

The Schrödinger equation

$$[-(\hbar^2/2m) \nabla^2 + V_{NL}(\underline{r})] \phi_k(\underline{r}) = E(\underline{k}) \phi_k(\underline{r}) \quad (4)$$

may be solved by expanding the pseudo wave function $\phi_k(\underline{r})$ as a sum of plane waves

$$\phi_k(\underline{r}) = \sum_g \beta_{\underline{k}+\underline{g}} |\underline{k}+\underline{g}\rangle \quad (5)$$

where the wave vectors \underline{g} are vectors of the reciprocal lattice. It is easy to show that the electronic energy bands are solutions of the secular equation

$$\det \left\{ [(\hbar^2/2m) (\underline{k}+\underline{g})^2 - E(\underline{k})] \delta_{gg'} + \Gamma_{NL}(\underline{k}+\underline{g}, \underline{k}+\underline{g}') \right\} = 0 \quad (6)$$

where $\Gamma_{NL}(\underline{k} + \underline{q}, \underline{k} + \underline{q}')$ are matrix elements of the lattice pseudopotential between plane wave states:

$$\Gamma_{NL}(\underline{k}, \underline{k}') = \langle \underline{k} | V_{NL}(r) | \underline{k}' \rangle . \quad (7)$$

Each matrix element may be written as the product of a structure factor and a matrix element of the ionic pseudopotential (22). The result is

$$\Gamma_{NL}(\underline{k}, \underline{k}') = S(\underline{k} - \underline{k}') \times \langle \underline{k} | U_{NL}(r) | \underline{k}' \rangle . \quad (8)$$

Since the structure factor $S(\underline{k} - \underline{k}')$ vanishes except when its argument is equal to a reciprocal lattice vector, only matrix elements of the form $\Gamma_{NL}(\underline{k}, \underline{k} + \underline{q})$ enter into calculations of the energy bands of a metal, and of the shape of its Fermi surface.

The magnitude of the nonlocal matrix element $\Gamma_{NL}(\underline{k}, \underline{k}')$ depends on both the initial state and the final state, and calculations of the energy band structure involve the full set of matrix elements $\Gamma_{NL}(\underline{k}, \underline{k} + \underline{q})$. Nevertheless, the energy bands are most strongly influenced by matrix elements between those plane wave states whose energies are most nearly equal (23). It is convenient to define the form factor $U_{NL}(\underline{q})$ of a nonlocal pseudopotential as the subset of matrix elements that influences most strongly the shape of the Fermi surface. In the region $[(\underline{q}/2k_F) \leq 1]$ the form factor is defined as the matrix element for scattering between states on the Fermi sphere, while in the region $[(\underline{q}/2k_F) > 1]$ it is defined as the matrix element for backward scattering from a state on the Fermi sphere. Thus the form factor is given by

$$U_{NL}(\underline{q}) = \Gamma_{NL}(\underline{k}_F, \underline{k}_F + \underline{q}) \quad (9)$$

where $|\underline{k}_F + \underline{q}| = |\underline{k}_F| \quad \text{if } (\underline{q}/2k_F) \leq 1$

and $|\underline{k}_F + \underline{q}| = (|\underline{q}| - |\underline{k}_F|) \quad \text{if } (\underline{q}/2k_F) > 1 .$

The form factor of a nonlocal pseudopotential is an approximate but concise expression of the matrix elements that describe the scattering of electrons on the Fermi surface. In the present paper we shall compare different nonlocal pseudopotentials by evaluating and comparing their form factors.

We now consider how one might calculate the pseudopotential form factor for a metal from the known potential of the free ion. The first step is to construct a lattice potential by superimposing ionic potentials centered on each

lattice site in the metal. Then the strongly-attractive lattice potential is truncated within the region of each ion core, in order to generate the lattice pseudopotential. Finally the lattice pseudopotential is divided by a dielectric constant to take into account screening of the pseudopotential by the conduction electrons in the metal.

The ionic potential within the core may be truncated in either of two distinct ways, which lead to the OPW and the APW pseudopotentials.

Following Phillips and Kleinman (11), one may transform the Schrödinger equation which arises in the OPW formulation of the band structure problem to a plane-wave basis set. The electronic energy bands are solutions of a secular equation (6) in which the nonlocal matrix elements are given by

$$\Gamma_{\text{OPW}}(\underline{K}, \underline{K}') = \left\{ V(\underline{K} - \underline{K}') + \sum_c (E - E_c) \langle \underline{K} | b_c \rangle \langle b_c | \underline{K}' \rangle \right\}. \quad (10)$$

The first term in this expression is the Fourier transform of the self-consistent ionic potential. The second term, which involves the energies E_c of the core states and a projection onto the Bloch states $| b_c \rangle$ that are constructed out of core wave functions, arises because the conduction-electron wave functions must be orthogonal to the core states. This orthogonalization term is necessarily positive (i.e., repulsive), and tends to cancel the attractive core potential. The extent of the cancellation depends on energy and angular momentum. We shall often refer to the weak nonlocal effective potential $\Gamma_{\text{OPW}}(\underline{K}, \underline{K}')$ as the OPW pseudopotential.

Alternatively, one may describe the truncation of the lattice potential in the language of multiple-scattering theory. Slater (24) devised the APW method as a means of solving the Schrödinger equation for a lattice potential of muffin-tin form. The electronic energy bands are solutions of a secular equation of the form (6). The nonlocal matrix elements are given by

$$\begin{aligned} \Gamma_{\text{APW}}(\underline{K}, \underline{K}') &= (4\pi R_s^2 / \Omega) \left\{ [E - (\hbar^2/2m)\underline{K} \cdot \underline{K}'] \frac{j_1(|\underline{K} - \underline{K}'| R_s)}{|\underline{K} - \underline{K}'|} \right. \\ &\quad \left. + \frac{\hbar^2}{2m} \sum_\ell (2\ell + 1) P_\ell (\cos \theta_{\underline{K}, \underline{K}'}) j_\ell (KR_s) j_\ell (K'R_s) \times (\mathcal{R}_\ell' / \mathcal{R}_\ell)_{R_s} \right\} \end{aligned} \quad (11)$$

where R_s is the radius of the muffin-tin sphere, and Ω is the volume of the primitive unit cell. The logarithmic derivatives of the radial wave functions $\mathcal{R}_\ell(r)$ are conveniently expressed in terms of the phase shifts of the effective potential by the equation

$$\frac{\eta'_\ell}{\eta_\ell} = \left[\frac{j'_\ell(kr) - \tan \eta_\ell(E) y'_\ell(kr)}{j_\ell(kr) - \tan \eta_\ell(E) y_\ell(kr)} \right]_{r=R_s} \quad (12)$$

where $(\hbar^2 k^2 / 2m) = E$, and where the primes on the spherical Bessel functions denote derivatives with respect to r .

In a band structure calculation by the APW method, the phase shifts $\eta_\ell(E)$ are determined by integrating the radial Schrödinger equation within the muffin-tin sphere. But the nonlocal matrix elements depend only on the reduced phase shifts, from which all integer multiples of π have been subtracted. In solving the band structure problem, the true ionic potential may therefore be replaced by an APW pseudopotential, as is illustrated in figure 1. The pseudopotential, which is constructed to yield the same set of phase shifts, modulo π , as the ionic potential, may depend both on energy and on angular momentum.

We have seen how one may truncate the short range component of the ionic potential in order to construct the lattice pseudopotential. In a metal, the long-range component of the ionic potential must also be modified to take into account screening by the conduction electrons. No long-wavelength fluctuations of lattice density exist in a perfectly-ordered metal, and there-

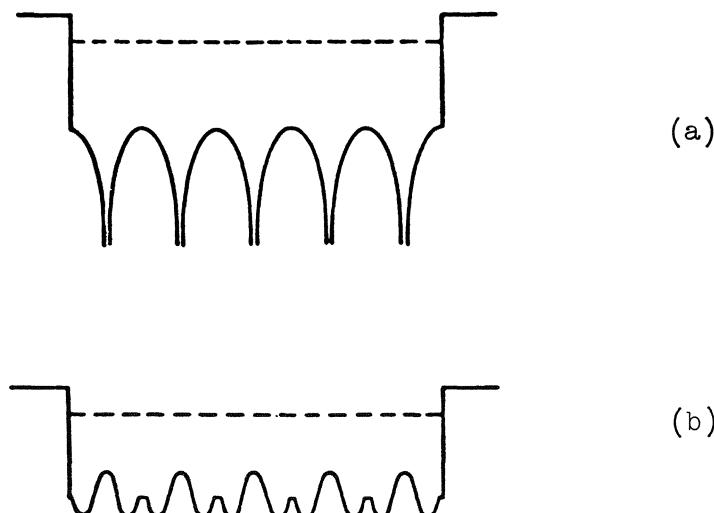


Figure 1. Schematic illustration of the construction of the APW pseudopotential. The ionic potential (a) is approximated by a potential of muffin-tin form. The lattice pseudopotential (b) is derived by truncating this potential within each ion core.

fore the long range components of the ionic pseudopotential (the low- q components of the form factor) cannot be determined by fitting Fermi surface data. Nevertheless, if the pseudopotential that is deduced from studies of the Fermi surface is also to be used to describe the various manifestations of the electron-phonon interaction, it must satisfy a condition imposed by charge neutrality in the metal. Shaw and Harrison (25) have shown that if a nonlocal pseudopotential is to describe correctly the interaction between conduction electrons and long-wavelength fluctuations of lattice density it must satisfy the condition

$$\lim_{q \rightarrow 0} U_{NL}(q) = -(2/3) E_F^* \quad (13)$$

where E_F^* is the Fermi energy measured with respect to the bottom of the band.

ANALYSIS OF EXPERIMENTAL DATA

We now consider how a nonlocal pseudopotential may be derived from experimental Fermi surface data. The matrix elements $\Gamma_{NL}(K, K')$ that appear in the secular equation (6) are expressed in terms of suitable parameters, and the secular equation is solved for the shape of a surface of constant energy in wave-vector space. The parameters are adjusted in order to bring the computed surface of constant energy into coincidence with the experimental data.

As an example of the OPW pseudopotential approach we consider an analysis of the Fermi surface of magnesium by Kimball, Stark and Mueller (12). The OPW pseudopotential was constructed as in equation (10). The Fourier coefficients of the screened ionic potential were treated as adjustable parameters, and were determined by fitting the Fermi surface data. The energies of the core states, and the overlap integrals between plane wave states and core states, were evaluated from atomic data. Their best choice of the adjustable parameters gave an excellent overall fit to the Fermi surface data. An alternative procedure based on the APW pseudopotential has been applied successfully to study the Fermi surfaces of the alkali metals (15). The phase shifts $\eta_\ell(E_F)$ that appear in the APW secular determinant were determined by fitting the Fermi surface data.

Both the OPW and the APW pseudopotentials seem capable of giving a satisfactory fit to the experimental data for simple metals, although a direct comparison of the two methods has yet to be attempted. The OPW pseudopotential has the advantage that it may be evaluated for any real-space potential, whereas the APW pseudopotential is restricted to a real-space potential of muffin-tin form. On the other hand, the APW pseudopotential

is somewhat easier to construct, since assumptions about the core states do not enter into the calculations. While the OPW and APW pseudopotential methods appear to be alternative ways of analyzing the shapes of the Fermi surfaces of simple metals, the relationship between these two methods is not immediately obvious. For example, the phase shifts obtained by fitting the Fermi surface data do not have to satisfy a Friedel sum rule, and in consequence the APW pseudopotential form factor does not necessarily satisfy the screening condition (13). We shall show that it is nevertheless possible to carry out a phase shift analysis of experimental Fermi surface data in such a way that the APW pseudopotential form factor satisfies the screening condition.

In a first-principles band structure calculation by the APW method, the value of the Fermi energy parameter E_F (the Fermi energy measured with respect to the constant potential in the region between the muffin-tin spheres) is determined by the condition that the number of electron states whose energies lie below E_F should be equal to the number of conduction electrons in the crystal. One might perhaps suppose that the value of the Fermi energy parameter which enters in a phase shift analysis can be determined from the known volume of the Fermi surface. That this is not so is demonstrated in figure 2. To the extent that the interstitial region of the lattice pseudopotential occupies a negligibly small fraction of the volume of each unit cell, a

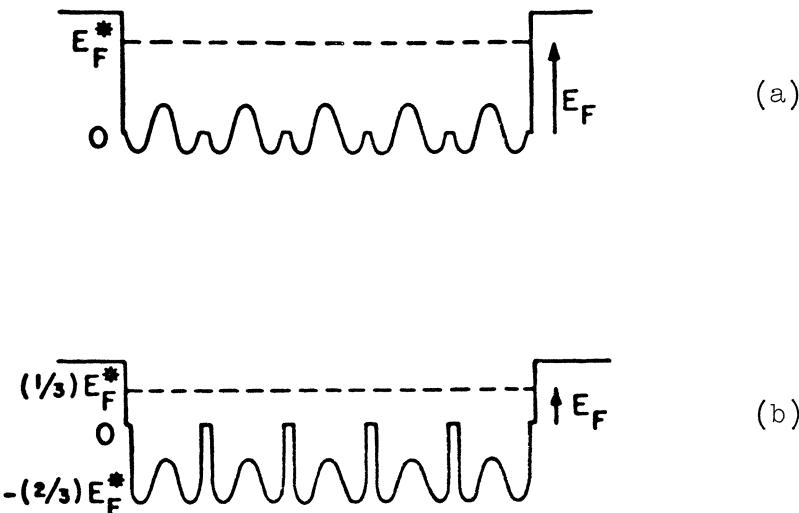


Figure 2. The principal effect of altering the Fermi energy parameter E_F in a phase shift analysis of experimental Fermi surface data is to change the zero of the energy scale. The low- q limit of the form factor of the APW pseudopotential in (a) is zero, while that of the pseudopotential in (b) is $-(2/3) E_F^*$.

particular choice of E_F determines only the zero of the scale on which energies are measured. It follows that E_F is a free parameter in the fit to the Fermi surface data. One can generate a family of APW pseudopotentials, each of which corresponds to a different choice of E_F , and each of which fits the experimental data equally well.

When the finite volume of the interstitial region is taken into account, it appears that APW pseudopotentials which correspond to different choices of E_F must generate somewhat different charge distributions in the metal. In particular, if E_F is substantially smaller than E_F^* (figure 2b), charge must be displaced away from the interstitial regions of the muffin-tin potential. Nevertheless, phase shift analyses of Fermi surface data for the alkali metals have been carried out successfully for a wide range of values of the Fermi energy parameter. In figure 3 we present APW pseudopotential form factors for lithium, that correspond to three different values of E_F . The discontinuous slope results from the change in the definition of the form factor when $(q/2k_F) = 1$. The similar behavior of three form factors in the structure region [$(q/2k_F) \gg 1$] is to be expected, since their behavior in this region is determined by fitting the Fermi surface data. But in the plasma region [$(q/2k_F) \ll 1$] the form factors differ significantly from one another, and

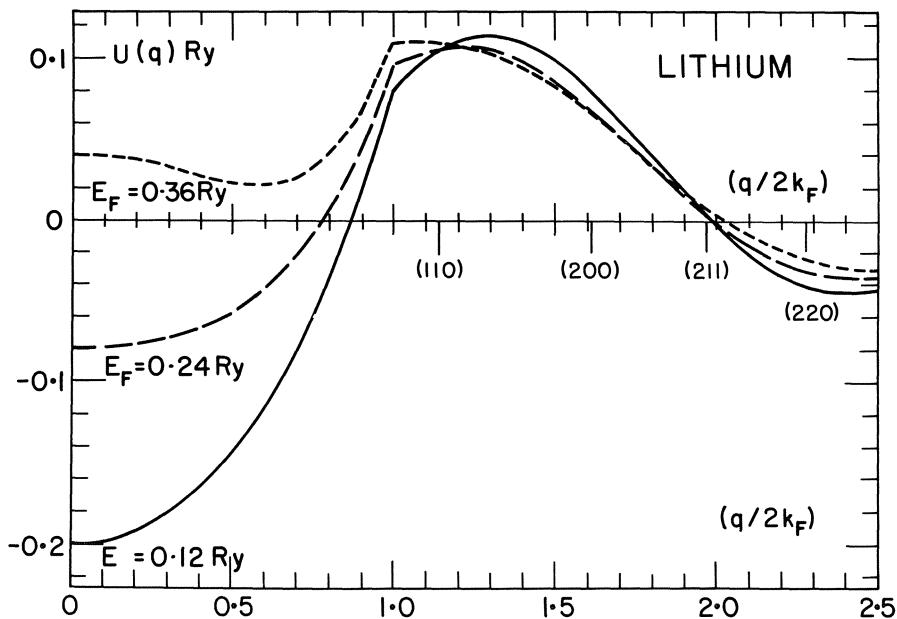


Figure 3. Alternative APW pseudopotential form factors for lithium. The form factors were obtained by fitting Fermi surface data with three different values of the Fermi energy parameter E_F .

by adjusting the otherwise undetermined Fermi energy parameter in a phase shift analysis, one can construct an APW pseudopotential whose form factor satisfies equation (13). It is shown in figure 2b that if we set

$$E_F = (1/3) E_F^* \quad (14)$$

then the APW pseudopotential form factor will satisfy the screening condition:

$$\lim_{q \rightarrow 0} U_{NL}(q) = -(2/3) E_F^*.$$

RESULTS AND DISCUSSION

We have developed computer programs to carry out the fitting procedure described above. We shall not present details of our numerical techniques, nor shall we discuss the convergence of our calculations, since such information has been presented elsewhere (15).

The Fermi energies of the alkali metals are not known with great accuracy, so in the present calculations we have set the Fermi energy parameter equal to one-third of the free-electron Fermi energy

$$E_F = (1/3) E_F^0. \quad (15)$$

Our results suggest that no serious error is introduced by this approximation. In Table II we have set out the phase shifts that are required to fit the Fermi

TABLE II.

Phase shift fit to Fermi surface data for the alkali metals.

	E_F	η_0	η_1	η_2
Li	0.11640	0.6785 ± 0.3322	0.1551 ± 0.0070	-0.0014 ± 0.0034
Na	0.07946	0.9495 ± 0.0040	0.1036 ± 0.0003	$+0.00204 \pm 0.00001$
K	0.05196	0.6532 ± 0.0030	0.0975 ± 0.0001	0.00713 ± 0.00004
Rb	0.04546	0.7194 ± 0.0170	0.0852 ± 0.0008	0.0077 ± 0.0001
Cs	0.03883	0.7087 ± 0.0499	0.0711 ± 0.0027	0.0094 ± 0.0004

surface data for the alkali metals when the Fermi energy parameter is given by equation (15). In figures 4 through 8 we present our APW pseudopotential form factors for the alkali metals, and compare them with form factors determined by other authors.

The experimental error in the measurements of the area distortions of the Fermi surface of lithium introduces a substantial uncertainty in our calculation of the form factor. Nevertheless, our form factor (figure 4) is in qualitative agreement with the form factors of Shaw (26) and of Harrison (27), and with an estimate of V_{110} derived independently of Cohen and Heine (8) from the same experimental data.

Our APW pseudopotential form factor for sodium (figure 5) is in good qualitative agreement with the form factor of the optimized pseudopotential of Shaw (26), and with the OPW-based form factor of Harrison (27). The value of V_{110} obtained from the APW form factor ($V_{110} = 0.0091 \pm 0.0002$ Ry)

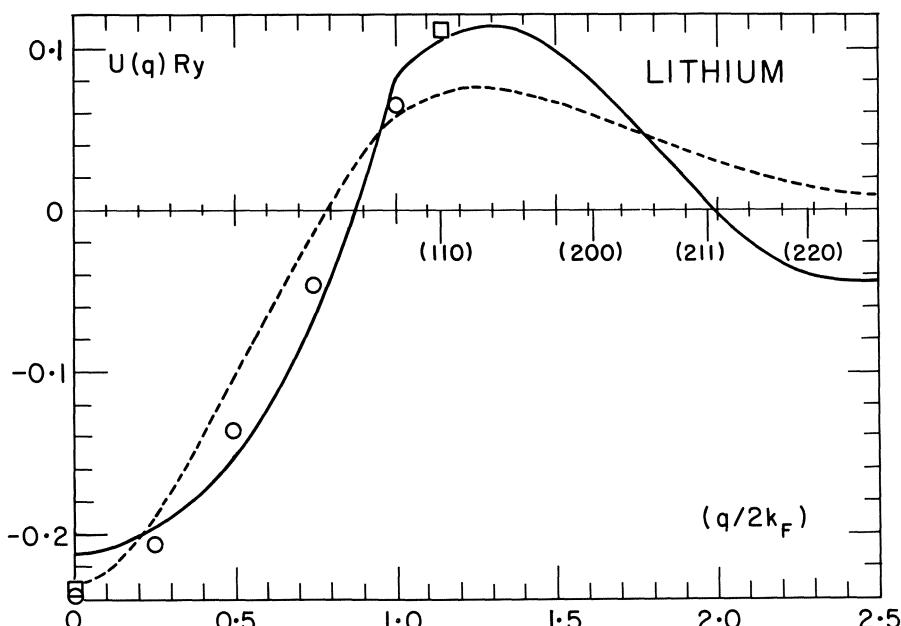


Figure 4. The APW pseudopotential form factor for lithium (continuous curve) is compared with the form factors of Shaw (26) (broken curve) and of Harrison (27) (open circles). The open square at $q=0$ is the (approximate) low- q limit $U(q) = -(2/3) E_F^0$. The open square at (110) is the V_{110} pseudopotential coefficient given by Cohen and Heine (8).

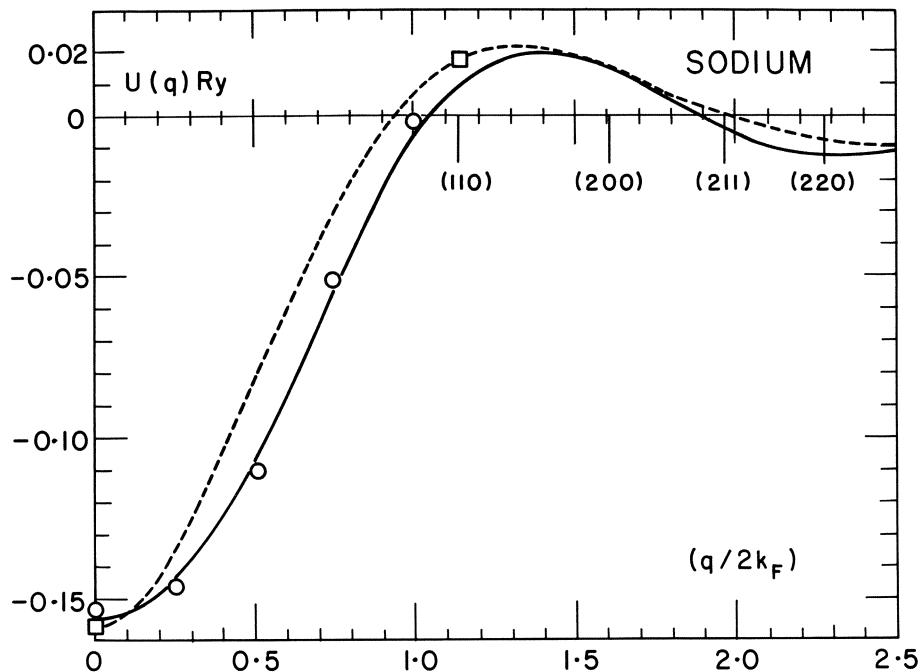


Figure 5. The APW pseudopotential form factor for sodium. The notation is the same as in figure 4, except that the open square at (110) is the V_{110} local pseudopotential coefficient given by Lee (17).

is smaller than the value derived by Lee (17) from a local pseudopotential analysis of Fermi surface data ($V_{110} = 0.0184 \pm 0.0008$ Ry). This discrepancy suggests that, even for sodium, the nonlocality of the matrix elements of the pseudopotential may be far from negligible.

The APW pseudopotential form factor for potassium is plotted in figure 6. In the low- q region it is in excellent agreement with the form factor of Shaw (26), but it departs substantially from that of Harrison (27). However, unlike the form factors of Animalu and Heine (28), Animalu et al (2), and Shaw (26), the APW form factor remains negative in the structure region; in particular, we find that V_{110} is negative.

There is some independent evidence to support this result. Band structure calculations for potassium (29) predict that the s-like state N_1 lies below the p-like state N'_1 , an ordering which is consistent only with a negative pseudopotential coefficient V_{110} . Moreover, Powell (30) and Smith (31) have interpreted discrepancies between the local pseudopotential coefficients that are

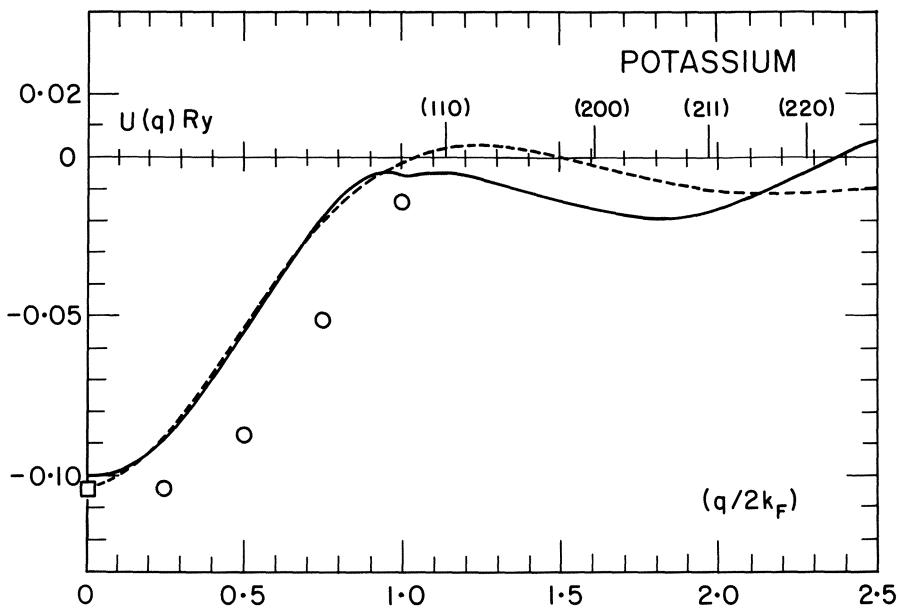


Figure 6. The APW pseudopotential form factor for potassium. The notation is the same as in figure 4.

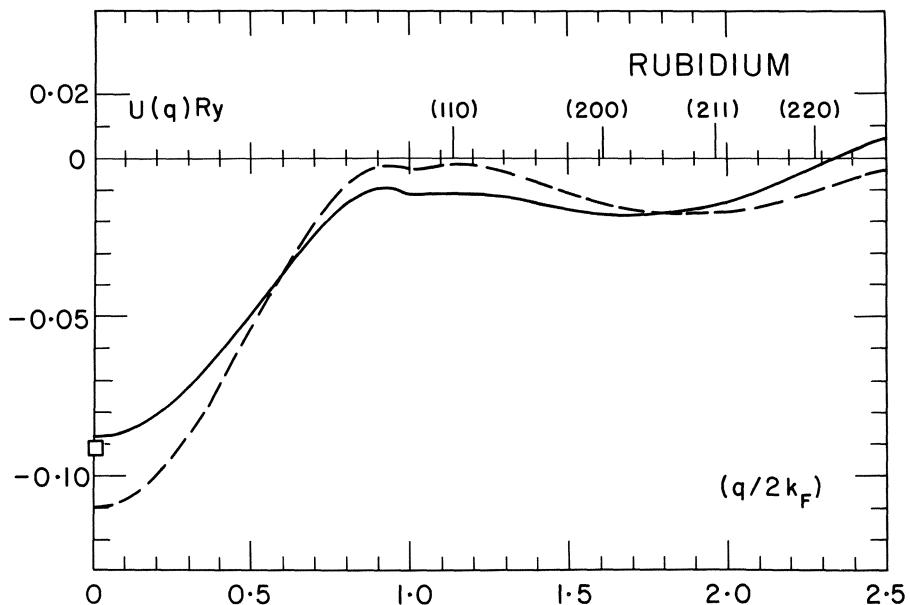


Figure 7. The APW pseudopotential form factor for rubidium (continuous curve) is compared with the form factor of Animalu and Heine (28). The open square at $q=0$ is $U(q) = -(2/3) E_F$.

required to fit experimental Fermi surface and optical data as evidence that V_{110} is positive in sodium but negative in potassium, in agreement with our result.

In figure 7 we display the APW pseudopotential form factor for rubidium. In the region of low q , our form factor satisfies the screening condition (equation 13) to within a few percent. Again we find V_{110} to be negative, as do Animalu and Heine (28). Finally, in figure 8 the APW pseudopotential form factor for cesium is presented. Our form factor satisfies the screening condition rather closely in the low- q limit; our value of V_{110} is much more strongly negative than is that of Animalu and Heine (28).

The numerical values of our pseudopotential coefficients are set out in Table III. Except for lithium, the errors in the experimental data do not introduce significant errors in the coefficients. We find that the value of V_{110} becomes progressively more negative with increasing atomic number, and that it changes sign between sodium and potassium. This trend was predicted by Cohen and Heine (32), and is associated with the decreasing density of the conduction electron gas with increasing atomic number in the alkali series.

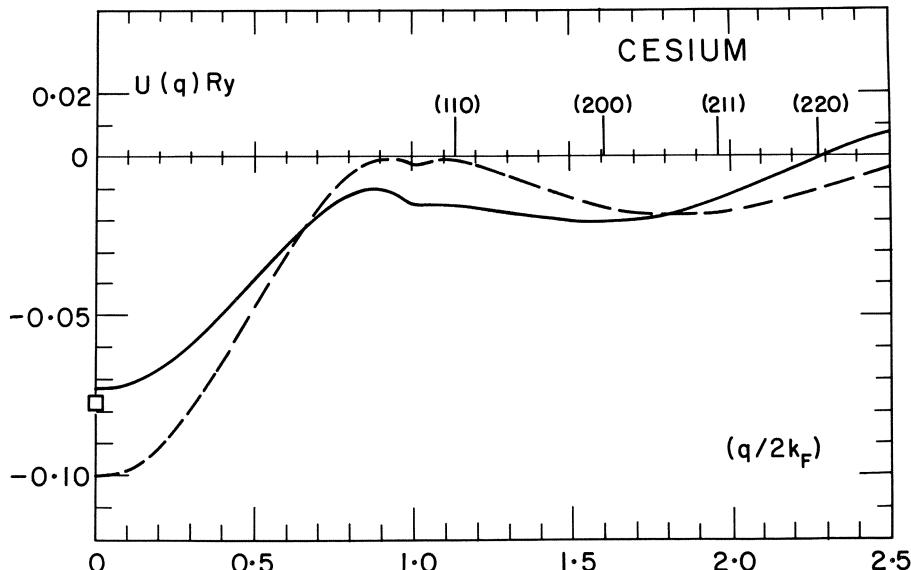


Figure 8. The APW pseudopotential form factor for cesium. The notation is the same as in figure 7.

TABLE III.
Fourier coefficients of the APW pseudopotential form factor (Rydbergs).

	v_0	v_{110}	v_{200}	v_{211}
Li	-0.206 ±0.013	0.103 ±0.039	0.084 ±0.083	0.005 ±0.044
Na	-0.1565 ±0.0003	0.0091 ±0.0002	0.0146 ±0.0001	-0.0044 ±0.0001
K	-0.1006 ±0.0001	-0.0054 ±0.0001	-0.0171 ±0.0001	-0.0171 ±0.0001
Rb	-0.0876 ±0.0006	-0.0112 ±0.0005	-0.0177 ±0.0003	-0.0143 ±0.0002
Cs	-0.0726 ±0.0017	-0.0162 ±0.0014	-0.0206 ±0.0005	-0.0133 ±0.0006

CONCLUSIONS

We have shown that, by an appropriate choice of the Fermi energy parameter that appears in a phase shift analysis of experimental Fermi surface data, one can construct an APW pseudopotential whose form factor satisfies the screening condition

$$\lim_{q \rightarrow 0} U(q) = -(2/3) E_F^*.$$

Our form factors for the alkali metals are in good qualitative agreement with those obtained by other authors, and this suggests that the depletion of the interstitial charge density that is implied by our pseudopotential may not be a serious objection to our procedure.

As a method of determining the pseudopotential form factors for simple metals, phase shift analysis of Fermi surface data has several advantages. Small changes in the phase shifts induce approximately linear changes in the distortions of the computed surfaces of constant energy, and this assures rapid convergence of our fitting procedure. The series of phase shifts is also rapidly convergent in angular momentum. For the alkali metals, only the s, p, and d phase shifts differ significantly from zero, and analysis in terms of phase shifts allows the most efficient use of the experimental data. Moreover, for a given choice of the Fermi energy parameter, the phase shifts are determined uniquely by fitting the Fermi surface data, and from the phase shifts one can compute the APW pseudopotential form factor for all values of q.

The APW pseudopotential is entirely analogous to the OPW pseudopotential. Our APW pseudopotential form factors are therefore an appropriate starting point for pseudopotential calculations of such properties of metals as the phonon spectrum, the mass renormalization, and the component of electrical resistivity that is associated with phonon scattering, all of which depend on the electron-phonon interaction. It seems that our demonstration of the relationship between the OPW and the APW pseudopotential form factors may be a step towards the wider application of the method of phase shift analysis.

ACKNOWLEDGMENTS

This work was begun during a short visit to the Royal Society Mond Laboratory, Cambridge, and was stimulated by discussions with Dr. Volker Heine of the Cavendish Laboratory. I wish to express my gratitude for their hospitality and assistance. In carrying out this research, I made use of facilities provided by ARPA for materials research at the University of Chicago.

REFERENCES

1. N. W. Ashcroft, *Phil. Mag.* 8, 2055 (1963).
2. A. O. E. Animalu, F. Bonsignori, and V. Bortolani, *Nuovo Cimento* 44B, 159 (1966).
3. N. W. Ashcroft and J. Wilkins, *Phys. Lett.* 14, 285 (1965).
4. J. P. Carbotte and R. C. Dynes, *Phys. Lett.* 25A, 532 (1967).
5. N. W. Ashcroft and L. J. Guild, *Phys. Lett.* 14, 23 (1965).
6. C. Hodges, *Phil. Mag.* 15, 371 (1967).
7. J. P. Carbotte and R. C. Dynes, *Phys. Lett.* 25A, 685 (1967).
8. For a review see M. L. Cohen and V. Heine, Solid State Physics, Vol. 24 (Academic Press, New York, 1970).
9. N. W. Ashcroft, *Phys. Rev.* 140, A935 (1965).
10. M. J. G. Lee and L. M. Falicov, *Proc. Roy. Soc. (London)* A304, 319 (1968).
11. J. C. Phillips and L. Kleinman, *Phys. Rev.* 116, 287 (1959).
12. J. C. Kimball, R. W. Stark, and F. M. Mueller, *Phys. Rev.* 162, 600 (1967).
13. R. W. Stark and L. M. Falicov, *Phys. Rev. Lett.* 19, 795 (1967).
14. R. W. Stark and L. M. Falicov, *Phys. Rev. Lett.* 19, 795 (1967).
15. M. J. G. Lee, *Phys. Rev.* 178, 953 (1969).
16. J. J. Donaghy and A. T. Stewart, *Phys. Rev.* 164, 391 (1967).
17. M. J. G. Lee, *Proc. Roy. Soc. (London)* A295, 440 (1966).
18. D. Shoenberg and P. J. Stiles, *Proc. Roy. Soc. (London)* A281, 62 (1964).
19. K. Okumura and I. M. Templeton, *Proc. Roy. Soc. (London)* A287, 89 (1965).

20. J. M. Ziman, *Adv. Phys.* 13, 89 (1964).
21. W. A. Harrison, Pseudopotentials in the Theory of Metals (W. A. Benjamin Inc., New York, 1966).
22. *Ibid.*, p. 31.
23. *Ibid.*, p. 288.
24. J. C. Slater, *Phys. Rev.* 51, 846 (1937); 92, 603 (1953).
25. R. W. Shaw and W. A. Harrison, *Phys. Rev.* 163, 604 (1967).
26. R. W. Shaw, *Phys. Rev.* 174, 769 (1968).
27. W. A. Harrison, *Phys. Rev.* 131, 2433 (1963).
28. A. O. E. Animalu and V. Heine, *Phil. Mag.* 12, 1249 (1965); see also reference 21, p. 310.
29. J. Callaway, *Phys. Rev.* 103, 1219 (1956); L. C. Allen, *Quarterly Report of the Solid State and Molecular Theory Group, MIT*, October 1958, p. 45; F. S. Ham, *Phys. Rev.* 128, 82 (1962); J. F. Kenney (private communication).
30. C. J. Powell (to be published).
31. N. V. Smith (to be published).
32. M. H. Cohen and V. Heine, *Adv. Phys.* 7, 395 (1958).

SELF-CONSISTENT ORTHOGONALIZED-PLANE-WAVE CALCULATIONS

R. N. Euwema, D. J. Stukel, and T. C. Collins

Aerospace Research Laboratories

Wright-Patterson Air Force Base, Ohio

I. INTRODUCTION

A natural way to describe mathematically a valence wave function in a periodic crystal is in terms of a Fourier series. However, convergence of such a plane-wave series is very poor because thousands of plane-wave terms are required to simulate the rapid oscillations of the wave function close to the atomic nuclei. To improve convergence, Herring¹ proposed the Orthogonalized-Plane-Wave (OPW) method in which the plane-wave terms making up the Fourier series are orthogonalized to all the tightly-bound, core-wave functions. This orthogonalization vastly improves the convergence because the core functions present in the valence wave function expansion correctly simulate the behavior of the valence wave function in the core regions; while the plane-wave terms adequately describe the overall crystalline behavior of the function.

The OPW method has been successfully applied to semiconductors by a number of workers. Herman and co-workers use a non-self-consistent potential obtained from isolated atom charge densities.² They neglect the rearrangement of the core and valence electrons in the crystalline environment, and compensate for this by empirically adjusting particular potential coefficients until certain energy differences match experiment.³ In our Self-Consistent OPW model⁴ (SCOPW), core and valence wave functions are calculated alternately in a more realistic crystalline environment. That is, given a potential, valence wave functions are obtained. Using this valence charge density, new self-consistent core wave functions are calculated. The new core charge density then contributes to a new crystal potential for the calculation of new valence wave functions, etc.

TABLE I-1. A comparison of SCOPW energies with experimental energies in eV.
 zb: zincblende symmetry
 w: wurtzite symmetry
 bg: direct band gap at Γ
 ibg: indirect band gap

Compound	Exp.	SCOPW*	Reference	Compound	Exp.	SCOPW*	Reference
Si	ibg	1.1	5	BeS	ibg	4.1	11
	ϵ_2	4.3	4.1	BeSe	ibg	3.4	11
Ge	ibg	5.3	5.0	BeTe	ibg	2.6	11
	bg	.8	1.2	ZnS(w)	bg	3.9	3.7
	ϵ_2	.9	1.1	ZnS(zb)	bg	3.8	3.7
	2.4	2.4	2.5	ϵ_2	5.7	5.8	12
BP	ibg	4.3	4.5		6.9	7.0	
	2.0	5.7	5.7		9.4	9.2	
BAs	ibg	1.5	1.8	CdS(w)	bg	2.6	2.5
	bg	2.4	1.5	CdS(zb)	bg	2.5	2.7
AlP	bg	2.2	2.1	CdSe(w)	bg	1.8	2.4
	ibg	2.3	2.3	CdSe(zb)	bg	1.9	2.2
AlAs	bg	3.7	2.5	ZnS(w)	bg	3.9	3.7
	ϵ_2	5.0	3.7	ZnS(zb)	bg	3.8	3.7
GaP	ibg	6.4	6.2	ϵ_2	5.7	5.8	12
	2.3	1.5	1.5		6.9	7.0	
GaAs	bg	3.1	3.1		9.4	9.2	
	ϵ_2	4.7	4.8	ZnSe(w)	bg	2.8	3.0
InP	bg	6.5	6.3	ZnSe(zb)	bg	2.8	2.8
	bg	1.3	1.3	ϵ_2	4.8	4.7	12
InAs	bg	.3	.9		6.4	6.0	
					8.4	8.0	

*Zincblende SCOPW band gaps have been corrected for spin-orbit splitting

The results of the SCOPW calculations have been very encouraging. Slater exchange band energies and imaginary part of the dielectric constant (ϵ_2) peak positions are usually within a few tenths of an eV of experiment, as can be seen in Table I-1. Input to the calculation consists only of lattice constants, nuclear charges and core-level quantum numbers. Absolutely no adjustments are made at any stage of the calculation. We are now applying this model to insulators and metals. In metals in which d states are important, it is necessary to extend the OPW model to a modified OPW method in which additional core-like states supplement the OPW basis set to aid the convergence of the d-like bands.¹³

In this paper, a detailed description is given of the self-consistent OPW method as developed at the Aerospace Research Laboratories over the last five years. Working equations, numerical methods, and a general outline of the procedure are given in textbook fashion. Illustrative examples will be taken from the semiconducting compounds calculations we have made using this method. Sources and magnitudes of numerical errors and physical approximations are discussed in order that one may investigate the applicability of the model to any particular crystalline system of interest.

In Section II, the general OPW formalism is presented, with special attention given to the form of the core and valence wave functions, the derivation and solution of the valence secular equation, and the convergence of the OPW Fourier expansion. Section III gives a brief discussion of the potential and core-shift calculations for Herman's isolated free atom model which serves as the zeroth iteration for the SCOPW model. In Section IV, the self-consistent aspects of the calculations are presented. These concern the treatment of the valence charge density, Brillouin zone averaging of the charge density and the calculation of crystalline core and valence potentials. A final overview of the SCOPW model is given in Section V.

II. GENERAL OPW FORMALISM

Wave Functions

In the OPW formalism,¹⁴ the electronic states of the crystal are divided into core and valence states which are handled quite differently. The core states are the tightly-bound, inner states. The following two assumptions are made about these states: (1) The core states at different atom sites do not overlap; and (2) the potential seen by the core states can be considered spherically symmetric. The second assumption implies that core wave functions can be written in the atomic form,

$$\psi_{nlm}(\vec{r}) = \frac{P_{nl}(r)}{r} Y_m^l(\theta, \phi), \quad 2.1$$

where n , l , and m are the principal, angular and magnetic quantum numbers, and Y_m^l is a spherical harmonic. The Bloch sum of these states then gives the crystalline wave function

$$\psi_{nlm}^k(\vec{r}) = \frac{1}{\sqrt{N}} \sum_{a=1}^N e^{i\vec{k} \cdot \vec{R}_a} \psi_{nlm}(\vec{r} - \vec{R}_a). \quad 2.2$$

In this expression, and throughout the article, one considers a crystal composed of N unit cells, each of volume Ω . Sums running from 1 to N thus run over one atom in each cell. We assume that there are t atoms per unit cell. Sums running from 1 to Nt are thus over all atoms in the crystal. In equation 2.2, the z -axis for the lm quantum numbers is taken in the \vec{k} direction.

The valence state wave functions are expanded in a modified Fourier series of plane waves, each of which has the form

$$\phi_k(\vec{r}) = \frac{1}{\sqrt{N\Omega}} e^{i\vec{k} \cdot \vec{r}}. \quad 2.3$$

These states are orthogonalized to all of the core states by using the valence projection operator

$$\hat{Q} = \hat{1} - \sum_{b,nlm} | b, nlm, \vec{k} \rangle \langle b, nlm, \vec{k} |, \quad 2.4$$

where the b sum is over one atom in each cell of the crystal. If the core states do not overlap from one atom site to another, \hat{Q} is

independent of \mathbf{k} and satisfies the familiar projection operator properties

$$\hat{Q}^+ = \hat{Q}, \quad \hat{Q}^2 = \hat{Q}. \quad 2.5$$

In order to calculate $\hat{Q} | \vec{k} \rangle$, the following mathematical relations are needed:

$$\begin{aligned} Y_o^\ell(\vec{r}, \vec{g}) &= \sqrt{4\pi/(2\ell+1)} \sum_m Y_m^*(\vec{k}, \vec{g}) Y_m^\ell(\vec{r}, \vec{k}) \\ &= \sqrt{(2\ell+1)/4\pi} P_\ell(\vec{r}, \vec{g}) \end{aligned} \quad 2.6$$

$$e^{ig \cdot \vec{r}} = \sum_\ell \sqrt{4\pi/(2\ell+1)} i^\ell j_\ell(gr) Y_o^\ell(\vec{r}, \vec{g}). \quad 2.7$$

P_ℓ is a Legendre polynomial, and the notation (\vec{r}, \vec{g}) stands for the angle between \vec{r} and \vec{g} . Using these relations, the matrix elements between core and plane-wave states can be calculated.

$$\begin{aligned} \langle b, n\ell m, \vec{k} | g \rangle &= \frac{1}{N\sqrt{\Omega}} \sum_{a=1}^N e^{i(\vec{g}-\vec{k}) \cdot \vec{R}_a} \int e^{ig \cdot (\vec{r}-\vec{R}_a)} \psi_{n\ell m}^*(\vec{r}-\vec{R}_a) d\vec{r} \\ &= \frac{1}{\sqrt{\Omega}} e^{i(\vec{g}-\vec{k}) \cdot \vec{R}_b} \int e^{ig \cdot \vec{r}} \psi_{n\ell m}(\vec{r}) d\vec{r} \\ &= \sqrt{4\pi/(2\ell+1)} e^{i(\vec{g}-\vec{k}) \cdot \vec{R}_b} Y_m^*(\vec{k}, \vec{g}) A_{n\ell}^b(g) \end{aligned} \quad 2.8$$

where

$$A_{n\ell}^b(g) = \sqrt{4\pi(2\ell+1)/\Omega} i^\ell \int r j_\ell(gr) P_{n\ell}(r) dr. \quad 2.9$$

To go from step 1 to step 2 in deriving equation 2.8, it is assumed that \vec{k} and \vec{g} belong to the same point in the first Brillouin zone so that $\vec{g} - \vec{k}$ is \vec{K} , a reciprocal lattice vector. Then since \vec{R}_a is equal to $\vec{R}_b + \vec{T}$, where \vec{T} is a lattice translation vector, $\vec{K} \cdot \vec{T}$ equals 2π . Hence, the sum contains N identical terms. An orthogonalized-plane-wave has the form

$$\hat{Q} |\vec{k}\rangle = |\vec{k}\rangle - \sum_{b,n\ell m} |b,n\ell m, \vec{k}\rangle \langle b,n\ell m, \vec{k}| \vec{k}\rangle , \quad 2.10$$

where the fact that \hat{Q} is independent of \vec{k} (equation 2.4) has been used, and where the \vec{k} for \hat{Q} has been chosen equal to that for $|\vec{k}\rangle$. Thus finally, we arrive at

$$\hat{Q} |\vec{k}\rangle = \frac{1}{\sqrt{N}} \left[\frac{e^{i\vec{k} \cdot \vec{r}}}{\sqrt{\Omega}} - \sum_{a=1}^t e^{i\vec{k} \cdot \vec{R}_a} \sum_{n\ell} A_{n\ell}^a(k) \psi_{n\ell 0}(\vec{r} - \vec{R}_a) \right] , \quad 2.11$$

where the z-axis for the angular core function, $\psi_{n\ell 0}^k$, is along \vec{k} .

The valence wave functions are then expanded in these orthogonalized-plane-waves,

$$|v, \vec{k}_0\rangle = \sum_k B_k \hat{Q} |\vec{k}\rangle , \quad 2.12$$

where the sum is over all \vec{k} vectors differing by a reciprocal lattice vector from \vec{k}_0 which is in the first Brillouin zone. The expansion coefficients, B_k , are chosen by means of a variational procedure to minimize the one-electron energies calculated from the crystal Hamiltonian.

Secular Equation

One wants to minimize

$$\lambda = \frac{\langle v, \vec{k}_0 | H | v, \vec{k}_0 \rangle}{\langle v, \vec{k}_0 | v, \vec{k}_0 \rangle} = \frac{\sum_{k,g} B_k^* H_{kg} B_g}{\sum_{\ell m} B_\ell^* U_{\ell m} B_m} \quad 2.13$$

where

$$H_{kg} = \langle \hat{Q} \vec{k} | H | \hat{Q} \vec{g} \rangle , \quad U_{\ell m} = \langle \hat{Q} \vec{\ell} | \hat{Q} \vec{m} \rangle . \quad 2.14$$

Zeroing the partial derivative of λ with respect to B_k^* gives the familiar secular equation for a non-orthogonal basis,

$$\sum_g H_{kg} B_g = \lambda \sum_g U_{kg} B_g , \quad 2.15$$

which in matrix form is

$$\hat{H} B = \lambda \hat{U} B .$$

The most obvious way to solve this matrix eigenvalue problem is to first diagonalize U , and then, after some matrix manipulation, to diagonalize a modified H matrix. This procedure requires two time-consuming diagonalizations; the first of which requires all the eigenvectors. A more efficient method is to use the Choleski decomposition of U into the product of a lower triangular matrix and its hermitian adjoint.¹⁵

$$U = L L^+ ; L_{ij} = 0 \text{ for } i < j \quad 2.16$$

One thus has the simpler eigenvalue problem,

$$(L^{-1} H L^{-1+}) L^+ B = \lambda L^+ B , \quad 2.17$$

which can be solved with a single matrix diagonalization requiring only the finally desired eigenvectors. Efficient techniques for converting a positive definite hermitian matrix to tri-diagonal form, and for finding the eigenvectors of a hermitian matrix can be found in the literature.¹⁶

The matrix elements of H and U can be readily obtained. For the core states,

$$H | b, n\ell m, \vec{k} > = E_{n\ell}^b | b, n\ell m, \vec{k} > , \quad 2.18$$

where $E_{n\ell}^b$ is the core energy for the electron which has quantum numbers $n\ell$ and is situated on the b^{th} atom of the unit cell. For plane-wave states, we have

$$\langle \vec{k} | H | \vec{g} > = k^2 \delta_{\vec{k}, \vec{g}} + V(\vec{g}-\vec{k}) . \quad 2.19$$

In this paper, atomic units are used for everything except energy units, where Rydbergs are used instead of Hartrees (double Rydbergs). Thus equation 2.19 has k^2 rather than $k^2/2$. The $V(\vec{g}-\vec{k})$ are obtained by Fourier transforming the potential in reciprocal-lattice-vector plane waves,

$$V(\vec{r}) = \sum_{\vec{k}} V(\vec{k}) e^{-i\vec{k}\cdot\vec{r}}$$

$$V(\vec{k}) = \frac{1}{N\Omega} \int V(\vec{r}) e^{i\vec{k}\cdot\vec{r}} d\vec{r} . \quad 2.20$$

The matrix elements are then

$$\begin{aligned}
 U_{kg} &= \langle \hat{Q} \vec{k} | \hat{Q} \vec{g} \rangle = \langle \vec{k} | \hat{Q}^* \hat{Q} | \vec{g} \rangle = \langle \vec{k} | \hat{Q} | \vec{g} \rangle \\
 &= \langle \vec{k} | \vec{g} \rangle - \sum_{b,n\ell m} \langle \vec{k} | b, n\ell m, \vec{k} \rangle \langle b, n\ell m, \vec{k} | \vec{g} \rangle \\
 &= \delta_{k,g} - \sum_{b=1}^t e^{i(\vec{g}-\vec{k}) \cdot \vec{R}_b} \sum_{n\ell} A_{n\ell}^{b*}(k) A_{n\ell}^b(g) P_\ell(\vec{k}, \vec{g}) \quad 2.21
 \end{aligned}$$

and

$$\begin{aligned}
 H_{kg} &= \langle \hat{Q} \vec{k} | H | \hat{Q} \vec{g} \rangle = \langle \vec{k} | \hat{Q}^* \hat{H} \hat{Q} | \vec{g} \rangle = \langle \vec{k} | H \hat{Q} | \vec{g} \rangle \\
 &= \langle \vec{k} | H | \vec{g} \rangle \\
 &- \sum_{b,n\ell m} \langle \vec{k} | b, n\ell m, \vec{k} \rangle E_{n\ell}^b \langle b, n\ell m, \vec{k} | \vec{g} \rangle \\
 &= k^2 \delta_{k,g} + V(\vec{g} - \vec{k}) \\
 &- \sum_{b=1}^t e^{i(\vec{g}-\vec{k}) \cdot \vec{R}_b} \sum_{n\ell} E_{n\ell}^b A_{n\ell}^{b*}(k) A_{n\ell}^b(g) P_\ell(\vec{k}, \vec{g}). \quad 2.22
 \end{aligned}$$

OPW Convergence

As mentioned before, the division of the electronic states of the crystal into core and valence states is governed primarily by the two requirements that the core wave functions at different atom sites do not overlap and that the valence wave function OPW Fourier expansion converges with a reasonable number of OPW's. The division that is most suitable for group IV, III-V and II-VI compounds is shown in Table II-1 for the first three rows of the periodic table using the group IV elements as examples. Core overlap is no problem with this choice, as can be seen in Figure II-1 for the case of ZnS. Less than 0.02 of an electron overlaps from one core into neighboring core regions. One cannot extend the core states beyond the double line in Table II-1, however, because large core overlap would occur.

Convergence of the valence OPW expansion is never complete with less than 229 OPW's and usually more are required. Three factors are involved in OPW convergence.¹⁷ (1) The presence or absence of core wave functions in the OPW,¹⁸ (2) the relative core

TABLE II-1. The double line indicates the division between core and valence states for representative first, second and third row compounds. The same division holds for all group IV, III-V, and II-VI compounds.

C	1s		2s	2p								
Si	1s		2s	2p		3s	3p					
Ge	1s		2s	2p		3s	3p	3d		4s	4p	

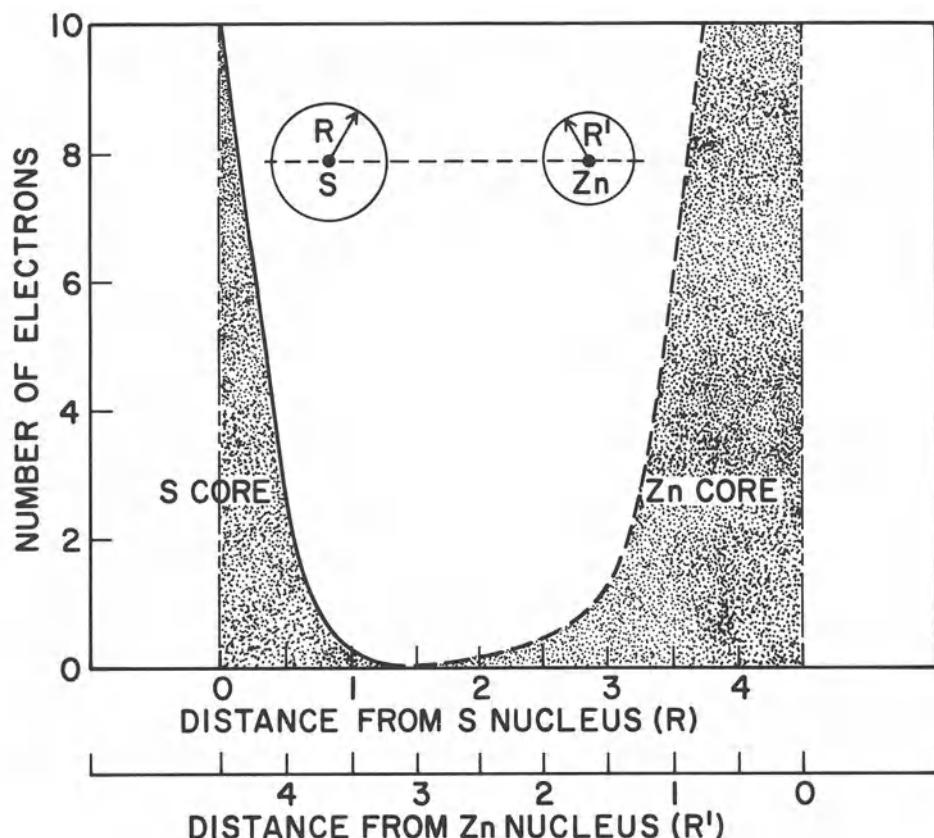


Fig. II-1. The number of core electrons outside radius r is plotted against r for Zn and S separated by their nearest neighbor distance in zincblende ZnS.

sizes of the atoms in the unit cell, and (3) the amount of localization of the valence wave function. One usually obtains comparable convergence over the Brillouin zone by choosing some maximum magnitude of k , k_{\max} , and including all those OPW's which have k less than k_{\max} . The Fourier series thus has a minimum wave length, λ_{\min} , associated with it, where

$$\lambda_{\min} = 2\pi / k_{\max} .$$

A good measure of the minimum distance that can be simulated by such a Fourier series is

$$d = \lambda_{\min} / 2 = \pi / k_{\max} . \quad 2.23$$

This suggests that convergence studies can conveniently be displayed by plotting the valence one-electron energies against this distance, d . It is then interesting to show the extent of the core region on the same plot by superimposing a plot of

$$\sigma = 4\pi r^2 \rho_{\text{core}} \quad 2.24$$

against r .

The convergence properties of C, Si, ZnS, BeTe and ZnO are shown in Figures II-2 through II-6. The first factor involved in OPW convergence, the presence or absence of core wave functions in the OPW's, is illustrated with C and Si in Figures II-2 and II-3. For Si, which has p states in the core, the convergence is excellent; while for C, which has no p state in the core, the convergence is very poor. The C core consists of a single s state. The Γ_{1v} (whose OPW's contain this ls state) has relatively good convergence. Convergence is complete at around 700 OPW's. The C Γ_{15} OPW's contain no core states to aid convergence. The OPW series becomes a plane-wave series - a straight Fourier series, and consequently, convergence is very slow. We suspect that complete convergence for C would require about 5,000 plane waves. The contrast between the C and Si convergence is dramatic. Si has 1s and 2s core states to aid the Γ_1 convergence, and it has a 2p core state to aid Γ_{15} convergence.

The second convergence factor, the relative core sizes of anion and cation, is illustrated by contrasting the convergence of Si, in which all cores have the same size, with the convergence of ZnS, BeTe and ZnO, shown in Figures II-4 through II-6. In the case of BeTe and ZnO, the convergence is also affected by the lack of p states in the Be and O cores. We can see from the figures that the convergence of the Γ_{15} band depends critically upon the anion core size, with a weaker dependence upon the cation core size.

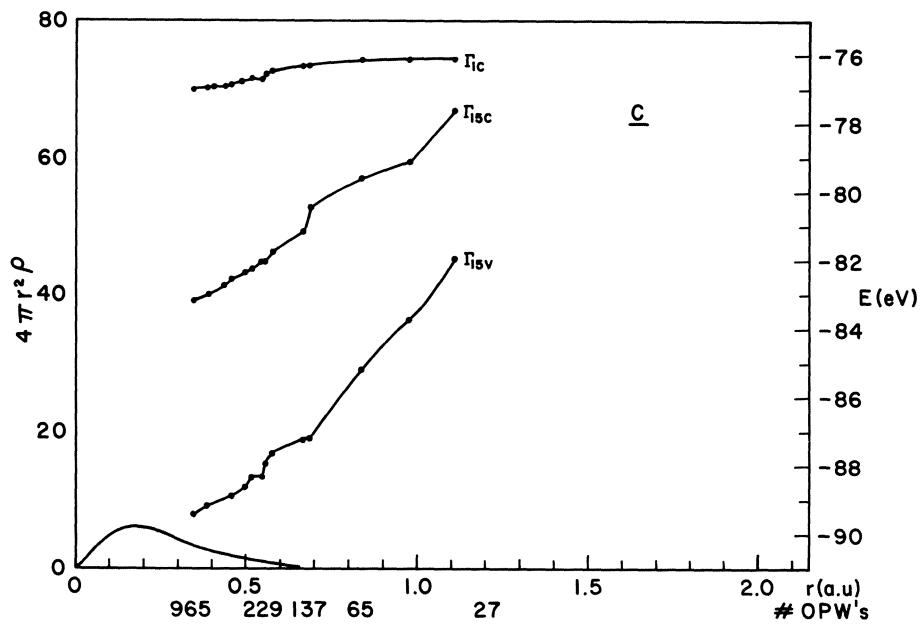


Fig. III-2. Diamond eigenvalues in eV are plotted against d (Eq. 2.23). σ (Eq. 2-24) is plotted against r for the C core.

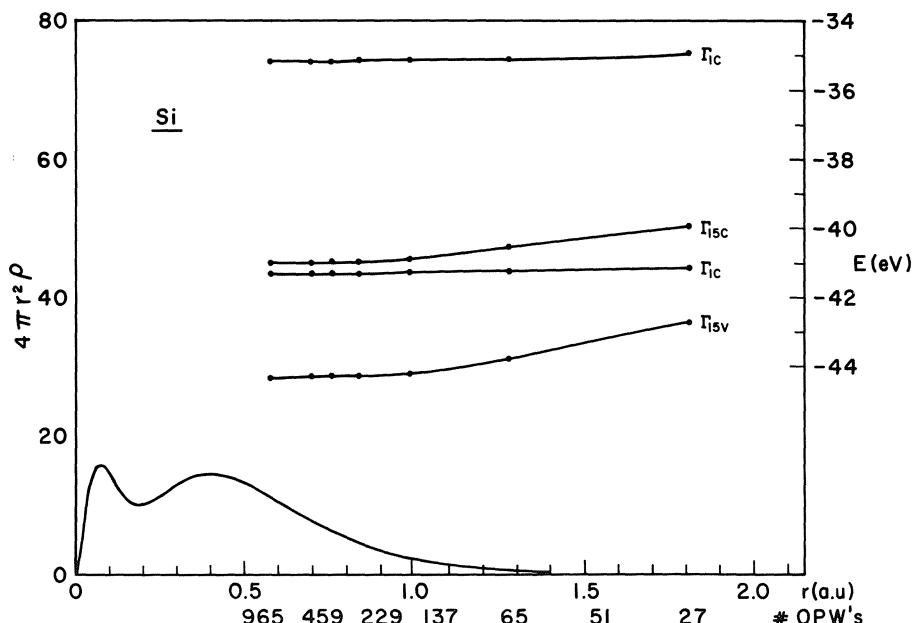


Fig. III-3. OPW convergence for Si is shown. (See caption to Fig. III-2).

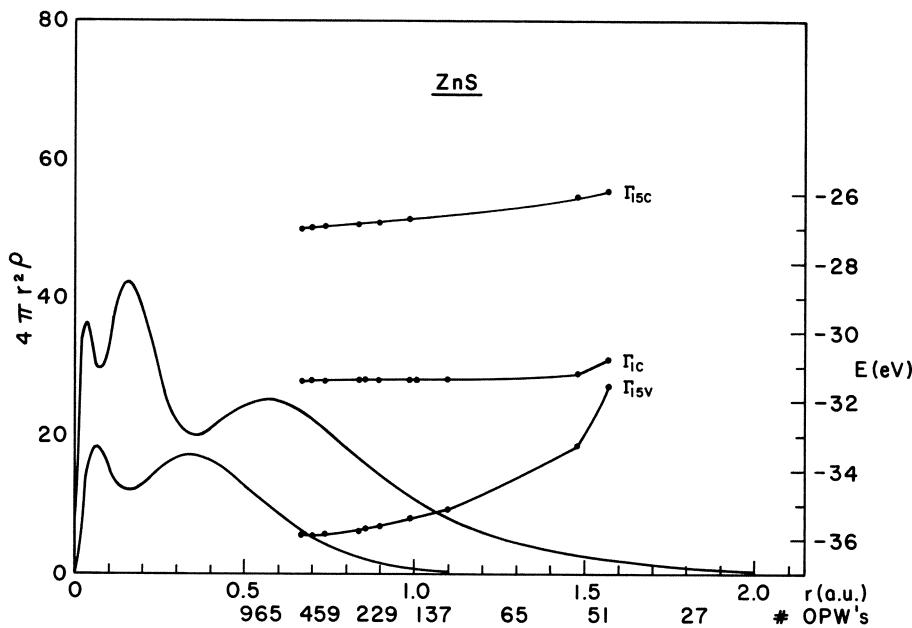


Fig. II-4. OPW convergence for ZnS is shown. (See caption to Fig. II-2).

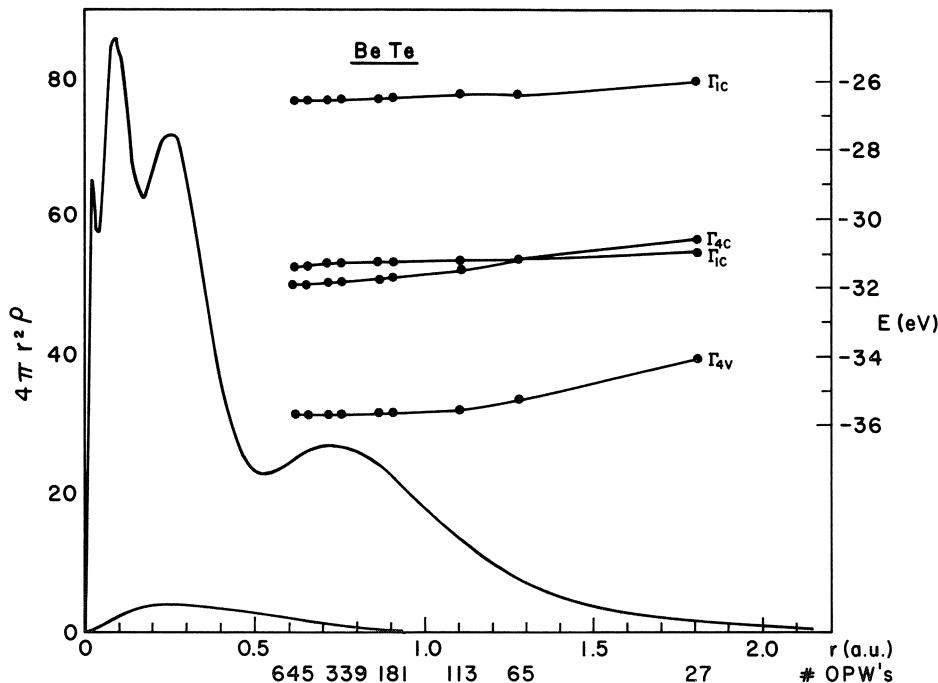


Fig. II-5. OPW convergence for BeTe is shown. (See caption to Fig. II-2).

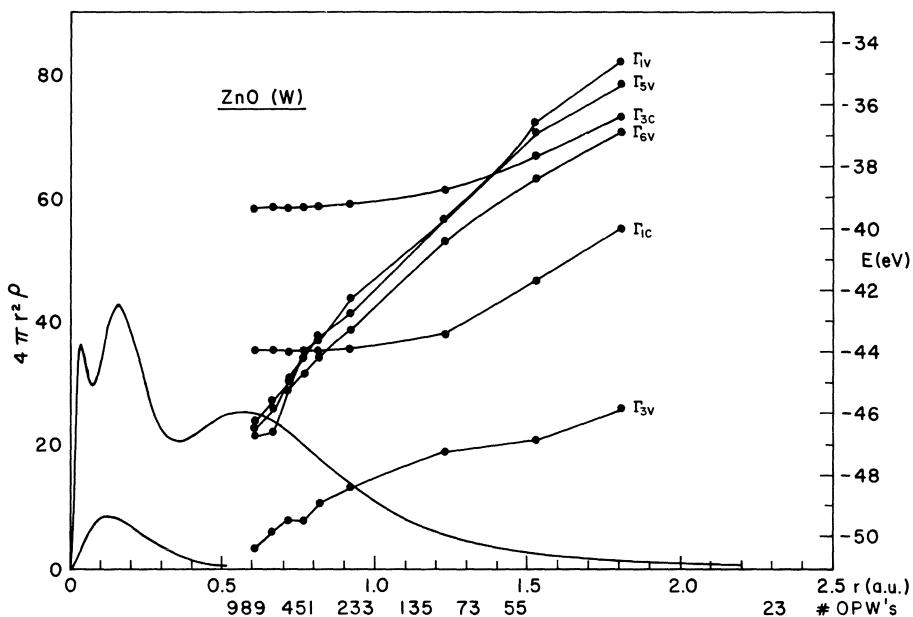


Fig. II-6. OPW convergence for ZnO is shown. (See caption to Fig. II-2).

Hence, the convergence of large anion-small cation compounds like BeTe is much better than the convergence of small anion-large cation compounds like ZnO.

Another factor affecting OPW convergence is the localization of the valence charge. The more localized the valence charge density, the poorer the convergence. Since solid Argon is in the same row as Si, one might expect similar convergence behavior. Actually, Argon's convergence is much worse and twice as many plane waves are required for Argon as are necessary for Si. The same poor convergence is observed for the NaCl compounds which are highly ionic.

Exchange Approximations

The largest fundamental problem in crystalline electronic calculations involves exchange and correlation approximations. If one ignores the correlation effects, one must consider the crystalline Hartree-Fock equations. Even the Hartree-Fock equations are very difficult to solve for electronic states in crystals because of the very complicated non-local, state-dependent exchange term. Consequently, various local approximations for this term have been proposed. The three most widely used approximations are

due to Slater,¹⁹ Kohn-Sham²⁰ Gaspar²¹ and Liberman.²² All of these exchange approximations have been used in our SCOPW model.²³

In all three approximations, it is assumed that the exchange potential at any point of a crystal or atom can be approximated locally by the exchange potential of electrons in a degenerate electron gas, with a Fermi momentum given in terms of the local density, $\rho(r)$, by the usual expression

$$k_F(r) = [3\pi^2\rho(r)]^{1/3} . \quad 2.25$$

For a degenerate electron gas, an electron with momentum k has an exchange energy in Rydbergs of

$$V_X(r) = -\frac{4}{\pi} k_F F [k(r)/k_F(r)] , \quad 2.26$$

where $F(\eta)$ is defined by the equation

$$F(\eta) = \frac{1}{2} + \frac{1-\eta^2}{4\eta} \ln [| (1+\eta)/(1-\eta) |] . \quad 2.27$$

Slater then averaged the exchange energy over the electron states in the filled sea to obtain

$$V_{XS}(r) = -\frac{3}{\pi} [3\pi^2\rho(r)]^{1/3} . \quad 2.28$$

Kohn-Sham, and independently Gaspar, derived an exchange potential equal to that of an electron at the Fermi surface,

$$V_{XG}(r) = -\frac{2}{\pi} [3\pi^2\rho(r)]^{1/3} . \quad 2.29$$

Slater and co-workers then proposed the generalized exchange potential,²⁴

$$V_{X\alpha}(r) = -\alpha \frac{3}{\pi} [3\pi^2\rho(r)]^{1/3} . \quad 2.30$$

Liberman used the Thomas-Fermi momentum of an electron with energy E_n ,

$$k(r) = \sqrt{2m [E_n - V(r)]} , \quad 2.31$$

together with the Fermi momentum, (equation 2.25) in equation 2.27. He arrived at

$$v_{xL}(r) = -\frac{4}{\pi} [3\pi^2 \rho(r)]^{1/3} F [\sqrt{2m(E_n - V(r))}/(3\pi^2 \rho(r))^{1/3}], \quad 2.32$$

where $V(r)$ is the potential energy, Coulomb plus exchange. Because of the presence of E_n , the energy of the state in question, the Liberman exchange is state-dependent.

These various approximations to the exchange potential have been extensively studied in free-atom calculations. It is found that Liberman's one-electron eigenvalues agree most closely with the Hartree-Fock eigenvalues. Kohn-Sham-Gaspar's wave functions are much better (compare more closely with Hartree-Fock) than Slater's and are slightly better than Liberman's. But Slater's eigenvalues are much closer to the experimental ionization energies than either of the others.

These exchange potentials have also been studied in SCOPW calculations. Table II-2 shows for ZnS how the eigenvalues vary with the choice of the exchange potential. We find the same relationship of the eigenvalues in the crystalline calculations as is found in the atomic calculations. Liberman's eigenvalues are too spread out to match experimental ionization energies, Kohn-Sham-Gaspar's are too compressed, while Slater's eigenvalues match experimental ionization energies quite closely. Slater's "exchange" seemingly contains some correlation as well. Table II-3 shows the theoretical x-ray form factors (four Ω times the Fourier components of the electron charge density) for ZnSe obtained by using different exchange approximations in the SCOPW model.²⁵

As we will see later, the eigenvalue shifts due to using different exchange approximations are of the same order of magnitude as the shifts due to the rearrangement of charge which occur when

TABLE II-2. The SCOPW energy eigenvalues obtained using the Liberman, Slater, and Kohn-Sham exchange approximations are given for cubic ZnS in eV.

	E_L	E_S	E_G
Γ_{15v}	.88	4.75	10.36
Γ_{1c}	6.48	8.46	13.08
X_{5v}	-1.93	3.12	8.26
X_{1c}	8.09	9.66	13.78
L_{3v}	-.15	4.13	9.56
L_{1c}	8.10	9.62	14.21

TABLE II-3. The experimental and SCOPW ZnSe Fourier components of the charge density in electrons per crystallographic unit cell are given for the different exchange approximations.

hkl	Exp.	L	S	G
111	158.55±1.6	156.17	157.68	156.22
200	14.86±0.55	11.44	11.72	11.15
220	189.80±2.0	189.05	191.14	189.48
311	129.10±2.0	125.03	126.40	125.50
222	11.63±0.6	9.82	9.61	9.50
400	162.31±1.8	161.39	163.57	162.51
331	109.55±2.0	109.39	111.13	110.34
420	12.09±0.7	11.29	10.85	11.08
422	140.56±1.8	142.82	145.37	144.43

one iterates to self-consistency. Hence, because the SCOPW model treats valence charge rearrangement correctly, it can be used to study different exchange approximations in crystalline calculations.

The following development will be in terms of the X_α exchange. The incorporation of Liberman's exchange into the SCOPW programs is discussed at length elsewhere.²³

Symmetry

Because of the slow convergence of the OPW series, it is very important to make maximum use of symmetry in the self-consistent calculation. A detailed discussion of symmetry in crystalline calculations can be found in Slater's book.²⁶ The specific application of symmetry to our OPW calculations is given elsewhere in this book.²⁷ Very briefly, the familiar group representation projection operator

$$\hat{P} = \frac{d}{g} \sum_{\hat{R} \in G} \Gamma_{11}^* (R) \hat{R}$$

operates on an OPW to give a symmetrized OPW. Use of this projection results in the block diagonalization of the secular equation. Typical symmetrized matrix sizes are given in Table II-4. Because P satisfies the relations

$$\hat{P}^2 = \hat{P} = \hat{P}^+ , [\hat{P}, H] = 0 , [\hat{P}, \hat{Q}] = 0 , \quad 2.33$$

the calculations of the H and U arrays simplify. For example,

TABLE II-4. The matrix sizes for various irreducible representations when 113, 229, and 459 OPW's are used.

Number of Plane Waves	Representation							
	Γ_1	Γ_{15}	X_1	X_3	X_5	L_1	L_3	$W_1 - W_4$
113	13	19	20	20	26	32	34	27
229	22	37	39	39	54	60	74	56
459	39	72	76	76	113	112	142	111

$$H_{gk} = \langle \overset{\leftrightarrow}{PQg} | H | \overset{\leftrightarrow}{PQk} \rangle = \langle \vec{g} | \overset{\leftrightarrow}{QPH} | \overset{\leftrightarrow}{PQk} \rangle \quad 2.34$$

$$= \langle \vec{g} | H | \overset{\leftrightarrow}{PQk} \rangle .$$

Calculations of the valence charge density are also shortened because only that part of $|\psi_v|^2$ having the full symmetry of the crystal contributes. Therefore, again the \hat{P} operator (but not $\hat{Q}!$) need appear on only one side. In addition, only one of the three degenerate Γ_{15v} wavefunctions (for example) need be calculated as all three contribute equally to the valence charge density.

III. THE ISOLATED ATOM MODEL

The discussion so far has presumed the availability of core states and a crystalline potential. In this section, we will briefly discuss a non-self-consistent model which supplies this information. The model, suggested by Herman and co-workers,² uses isolated-atom core states and potentials. Adjustments must be made to various Fourier coefficients of the potential and core energies before the resulting band structure will match experiment. But Herman's model does serve as a starting point for our self-consistent model. It also allows one to quickly and easily perform various OPW studies. For example, our OPW convergence studies on new compounds are made with Herman's isolated-atom model.

In this model, one first finds the various wave functions and energies for the isolated atoms. One then imagines these isolated-atom charge densities packed in the crystalline lattice. The crystalline potential can be obtained by solving the equations

$$\nabla^2 V_z(\vec{r}) = 8\pi \rho_z(\vec{r}) = 8\pi \sum_{a=1}^{Nt} z_a \delta(\vec{r} - \vec{R}_a) \quad 3.1$$

$$\nabla^2 V_c(\vec{r}) = 8\pi \rho_{el.}(\vec{r}) = - 8\pi \sum_{a=1}^{Nt} \rho_a (\vec{r} - \vec{R}_a) \quad 3.2$$

$$V_{xa}(\vec{r}) = - 6\alpha [3/(32\pi^2)]^{1/3} \sum_{a=1}^{Nt} [4\pi \rho_a (\vec{r} - \vec{R}_a)]^{1/3} \quad 3.3$$

Z_a is the nuclear charge of atom a while ρ_a is the electron charge density. α is the exchange constant. The sum is over all atoms in the crystal. Since everything is expressed in atomic units except energies (which are in Rydbergs), equations 3.1 and 3.2 have 8π rather than the more familiar 4π . The signs are governed by the fact that the V 's are electron potential energies rather than electrostatic potentials. The resulting potential energy obtained from equation 2.20 is

$$V(\vec{K}) = \frac{1}{\Omega} \sum_{a=1}^t e^{i\vec{K} \cdot \vec{R}_a} [\frac{8\pi}{K^2} (-Z_a + \int \rho_a(r) j_o(Kr) 4\pi r^2 dr) - 6\alpha (6\pi)^{1/3} \int (4\pi \rho_a(r))^{1/3} j_o(Kr) r^2 dr] . \quad 3.4$$

There is an apparent singularity in equation 3.4 when K is equal to zero. However, if $j_o(Kr)$ is expanded for small Kr ,

$$j_o(Kr) = 1 - \frac{1}{6} (Kr)^2 + \dots$$

one obtains

$$V(K=0) = \frac{1}{\Omega} \sum_{a=1}^t [- \frac{4\pi}{3} \int \rho_a(r) 4\pi r^4 dr - 6\alpha (6\pi)^{1/3} \int (4\pi \rho_a)^{1/3} r^2 dr] . \quad 3.5$$

To do this, one uses the fact that

$$Z_a = \int \rho_a 4\pi r^2 dr .$$

If one wants a slightly more realistic crystal model where isolated ions (rather than atoms) are packed in the crystal lattice, one has to be careful to include Madelung terms proportional to the ionicity,

$$Z_a = \int \rho_a 4\pi r^2 dr .$$

It will be clear how this is to be done after we have discussed such terms in the framework of our self-consistent model.

Herman has thus presented a fairly realistic crystalline potential which does, however, neglect charge rearrangement as the isolated atoms are brought together to form a crystal. For the core wave functions, he suggests using the corresponding isolated-atom wave functions. This is a very good approximation only for the lower core states. To illustrate this, Table III-1 shows Slater ZnS core-plane wave orthogonality coefficients (equation 2.9) for both the isolated-atom and for the SCOPW core functions. The differences are significant for the outer core states.

In Herman's model, the core energies to be used in equation 2.22 differ from the isolated-atom energies because of the overlap of the valence atomic wave functions from one atom site to the next. Herman corrects for this overlap by adding up all the isolated-atom total potentials at a site due to neighboring sites.

$$E_{nl}^b = E_{nl}^b + \sum_{\substack{a=1 \\ a \neq b}}^{Nt} V(\vec{R}_a - \vec{R}_b) \quad 3.6$$

After the first few shells, the main contribution to the core shifts comes from the $p^1/3$ exchange term of the potential. It may seem unsatisfactory that the core shifts depend so much upon the very unrealistic exchange tails. But, in fact, the dependence is

TABLE III-1. Fourier transforms of core wave functions (equation 2.9) are shown for $K=(000)$ and $K=(311)$ for both the isolated-atom and the SCOPW ZnS core functions.

	$A_{nl}(000)$		$A_{nl}(311)$		
	IA	SCOPW	IA	SCOPW	
S	1s	.014493	.014491	.013976	.013974
	2s	-.11976	-.11915	-.074662	-.074568
	2p	0	0	.068226	.067762
Zn	1s	.0054931	.0054931	.0054389	.0054389
	2s	-.038539	-.038539	-.034627	-.034627
	2p	0	0	.012800	.012800
	3s	.18663	.18649	.079054	.079059
	3p	0	0	-.12273	-.12252
	3d	0	0	.13162	.12867

not critical because the exchange tail contributes the same amount at large r to $V(K=0)$. Thus the exchange tail shifts all core and valence eigenvalues by the same amount, and does not affect relative energy differences. To see this, one specializes equation 2.20 to $K=0$ to get

$$V(K=0) = \frac{1}{N\Omega} \sum_{a=1}^{Nt} \int V(\vec{r} - \vec{R}_a) d\vec{r} = \frac{1}{\Omega} \sum_{b=1}^t \int V(\vec{r} - \vec{R}_b) d\vec{r} \quad 3.7$$

$$V(K=0) = \sum_{b=1}^t \sum_{a=1}^N V(\vec{\xi}_a - \vec{R}_b) . \quad 3.8$$

The mean value theorem for integration was used to replace the integral over the crystal by a sum over cells, ξ_a being in the a th cell. We see that the distant neighbor part of equation 3.8 is identical to that of equation 3.6 once $V(\vec{r})$ varies so slowly that $V(\vec{\xi}_a) \approx V(\vec{R}_a)$. The model results are thus not sensitive to the long-range exchange tails. To utilize this fact, however, the integrals in equation 3.7 should be carried out just as far as the core-shift sums in equation 3.6. This caution is especially important in the relativistic version of the model.²⁸

IV. SELF-CONSISTENT OPW MODEL

In a self-consistent OPW calculation, valence wave functions are calculated at selected high-symmetry Brillouin zone points and the ensuing charge densities are averaged over the Brillouin zone. The valence charge density is spherically symmetrized about each inequivalent atom site so that new mutually self-consistent core wave functions can be calculated in this frozen-valence environment. One then calculates a new crystalline potential from the (non-spherically-symmetrized) valence charge density and from the new core charge density. New core-plane wave orthogonality coefficients are also calculated from the new core wave functions. From this information we calculate new valence wave functions and energies to complete the self-consistent iteration cycle. The rest of this article will be devoted to elaborating upon this paragraph. Table IV-1 summarizes the self-consistent cycle.

Self-consistent charge rearrangement is important in crystalline calculations. A rather extreme example of the importance of self-consistency occurs when an external charge is placed in a crystal. The non-self-consistent potential due to this charge is $2/r$, while a proper self-consistent calculation, in which electrons (and in this example, ions as well) are allowed to rearrange, would give a potential of $2/\epsilon r$, with the electronic contribution to the

TABLE IV-1. The flow diagram for an SCOPW calculation.

1. Isolated-atom starting model
2. Valence wave functions and energies
3. Stop if band energies changed less than 0.01 eV from previous iteration
4. Valence charge density
5. Crystal potential spherically symmetrized about core sites in r space for core potential only
6. Core wave functions and energies
7. Core-plane wave orthogonality integrals
8. Crystalline potential in K space for valence calculation
9. Go to 2 above

dielectric constant ϵ being around 5 to 12 for the semiconductors we have calculated. To illustrate valence charge rearrangement from the isolated atoms to the crystal, Table IV-2 shows a comparison of isolated-atom, self-consistent OPW and experimental band gaps. Table IV-3 gives a comparison of isolated-atom and self-consistent OPW Fourier components of the charge density for Si and ZnSe. It can be seen that isolated-atom to SCOPW shifts are as large as the Slater to Kohn-Sham-Gaspar exchange approximation shifts. Consequently, before any conclusions are drawn about the efficacy of various exchange-correlation approximations, one must do self-consistent calculations.

Valence Charge Density

In describing the SCOPW model, we will first discuss the treatment of the valence charge density. The valence wave functions are all sums of terms like equation 2.11, i.e., combinations of plane-wave and core terms. The valence charge density is thus

TABLE IV-2. The direct band gaps (in eV) are given based on the isolated atom and the SCOPW model for different exchange approximations

Compound	Exp.	Isolated Atom		SCOPW	
		S	G	S	G
ZnS	3.84	4.30	3.68	3.71	2.72
ZnSe	2.83	3.16	2.78	2.94	2.68
GaAs	1.54	1.32	1.49	1.61	1.97

TABLE IV-3. Isolated atom and SCOPW x-ray form factors calculated for Slater (S) and Kohn-Sham-Gaspar (G) exchanges are compared with experiment

hkl	Exp.	Si ($\Omega \rho(K)$)			
		IA-S	SC-S	IA-G	SC-G
111	11.12 ± .04	10.93	10.88	10.53	10.70
220	8.78 ± .09	8.99	8.76	8.71	8.63
311	8.05 ± .07	8.41	8.09	8.15	7.99
222	.22 ± .04	0.00	0.22	0.00	0.19
400	7.40 ± .14	7.73	7.53	7.48	7.42
ZnSe($4\Omega \rho(K) $)					
111	158.55 ± 1.6	157.51	157.68	155.70	156.22
200	14.86 ± .55	11.42	11.72	11.45	11.15
220	189.80 ± 2.0	191.92	191.14	189.23	189.48
311	129.10 ± 2.0	127.28	126.40	125.45	125.50
222	11.63 ± .6	10.10	9.61	10.65	9.50
400	162.31 ± 1.8	163.99	163.57	161.43	162.51

$$\rho_v(\vec{r}) = \sum_{\text{valence states}} \left| \sum_k (p_w - c) \right|^2 = \rho_{pw}(\vec{r}) + \rho_{vc}(\vec{r}), \quad 4.1$$

where

$$\rho_{pw}(\vec{r}) = \sum_{\text{valence states}} \left| \sum_k (pw) \right|^2. \quad 4.2$$

ρ_{pw} is that part of the valence charge density involving plane wave terms, and ρ_{vc} is that part involving core terms. It is very natural to handle ρ_{pw} in K-space.

$$\rho_{pw}(\vec{r}) = \sum_v \rho_{pw}(\vec{k}) e^{-i\vec{k}\cdot\vec{r}} = \sum_v \sum_{k,g} B_k^v * e^{i(\vec{g}-\vec{k})\cdot\vec{r}} B_g^v \quad 4.3$$

A convenient approximation is to spherically symmetrize $\rho_{vc}(r)$ about the atom sites.

$$\rho_{vcs}(\vec{r}) = \sum_{a=1}^{Nt} [\rho_{vcs}^a(\vec{r} - \vec{R}_a)] \quad 4.4$$

This is not a necessary approximation. We have calculated the Fourier transform of the charge density both ways - exactly and approximately. The results are shown in Table IV-4. Self-consistent Slater GaP band energies calculated with ρ_{VC} and with the correct ρ_{VC} are given in Table IV-5. The correct calculation, which is extremely time-consuming, changes the energy eigenvalues less than 0.1 eV. The deviations are so small that we will use this time-saving approximation in the following development. The relative integrated sizes of ρ_{PW} and ρ_{VC} are given in Table IV-6 for Slater ZnS. The integral of ρ_{PW} is ten times as large as that

TABLE IV-4. The first column gives GaP SCOPW form factors, $4\pi\rho(K)$, when ρ_{VC} is used in the iteration cycle. Column two gives the form factor difference between $\rho_{VC}(K)$ and $\rho_{VCs}(K)$. The last column gives the self-consistent form factors when ρ_{VC} is used in the iteration cycle.

hkl	ρ_{VCs} (self-consistent)	$\delta\rho_{VC}$	ρ_{VC} (self-consistent)
111	107.049, -47.236	-.016, .098	107.031, -47.195
200	58.927	.024	58.908
220	129.038		129.063
311	85.688, 33.851		85.688, 33.857
222	50.996, -1.786		50.990, -1.796
400	109.442		109.441

TABLE IV-5. The first column gives Slater GaP SCOPW band energies in eV when ρ_{VC} is used in the iteration cycle. For the second column, ρ_{VCs} is used.

	ρ_{VC}	ρ_{VCs}	ρ_{VC}	ρ_{VCs}
Γ_{15c}	9.07	9.09	X _{1v}	-1.90
Γ_{1c}	6.83	6.88	X _{3v}	-4.71
Γ_{15v}	4.44	4.51	L _{3c}	10.05
Γ_{1v}	-7.38	-7.32	L _{1c}	6.87
X _{1c}	7.34	7.36	L _{3v}	3.49
X _{3c}	7.08	7.06	L _{1v}	-1.61
X _{5v}	2.15	2.20	L _{1v}	-5.57

TABLE IV-6. Integrals over a unit cell for ZnS of ρ_{VC} and ρ_{PW} . The sum correctly gives 8 valence electrons.

$$\int \rho_{VC}^S d^3r = - .6735 \quad \int \rho_{VC}^{Zn} d^3r = - .2697 \quad \int \rho_{PW} d^3r = 8.9432$$

of ρ_{vc} and is treated exactly (non muffin tin). The charge densities ρ_c , ρ_{vcs} , and ρ_{pws} (ρ_{pw} spherically symmetrized about an atom site) are shown as a function of r in Figure IV-1.

Brillouin Zone Average of Valence Charge Density

All the electrons throughout the Brillouin zone contribute equally to the valence charge density. Again, because of computer space and time limitations, this integral over the Brillouin zone must be approximated by a weighted sum over selected high symmetry points. For semiconductors we choose the relative weighting factor at a zone point proportional to that volume of the Brillouin zone closer to that point than to any other of the considered points. The zincblende Brillouin zone is shown in Figure IV-2. The weightings based upon the nearest volume criterion are shown in Table IV-7. For most of our calculations, we average over the four zone points Γ , X, L, and W. To test the adequacy of this approximation, we have iterated to self-consistency using only the Γ point, only the L point, the Γ -X-L points, the Γ -X-L-W points and finally the Γ -X-L-W- Δ - Σ points.^{5,12} The last set of points forms a regular mesh in K-space and is thus a very natural choice. The

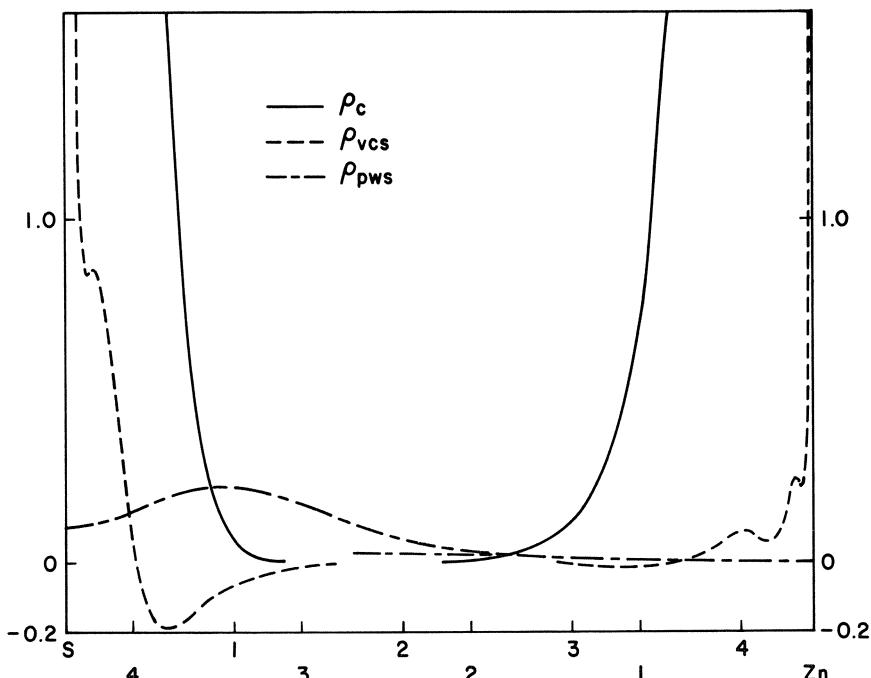


Fig. IV-1. ρ_c , ρ_{vcs} , and ρ_{pws} are shown as functions of r for Slater ZnS, with the Zn and S nuclei separated by their nearest neighbor distance.

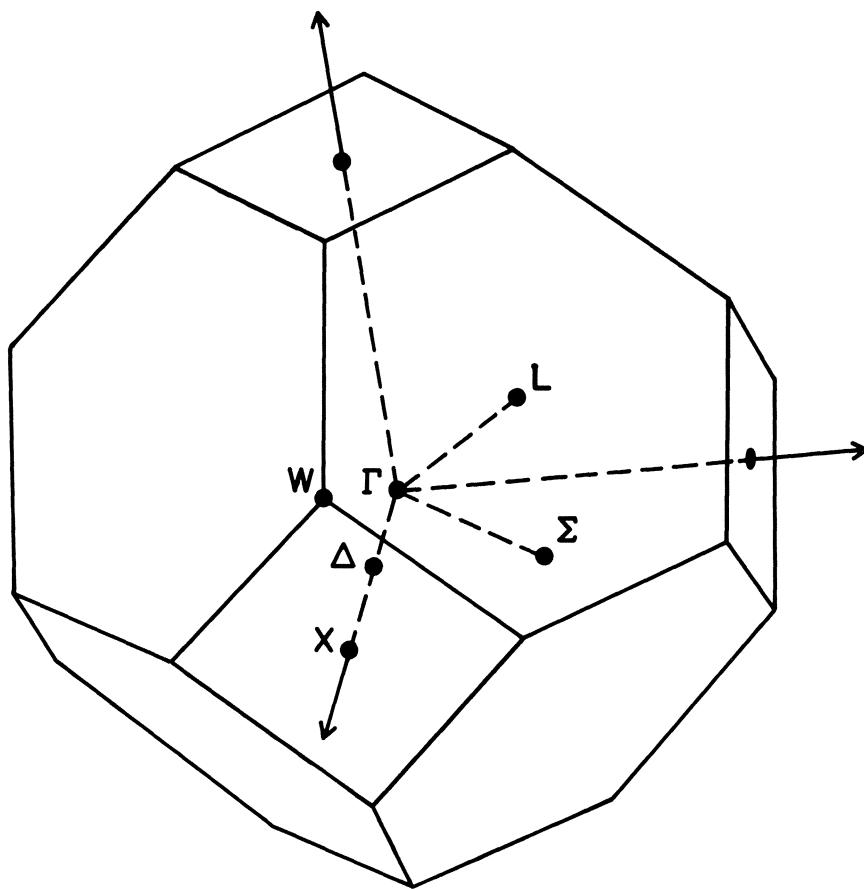


Fig. IV-2. The zincblende Brillouin zone with the high symmetry points $\Gamma(0\ 0\ 0)$, $X(1\ 0\ 0)$, $L(1/2\ 1/2\ 1/2)$, $W(1\ 1/2\ 0)$, $\Delta(1/2\ 0\ 0)$ and $\Sigma(1/2\ 1/2\ 0)$.

TABLE IV-7. Zonal weighting factors used in averaging the electron density over the Brillouin zone.

Point	Γ	L	$\Gamma-X-L$	$\Gamma-X-L-W$	$\Gamma-X-L-W-\Delta-\Sigma$
Γ	1.0	0.0	0.125	0.1258	0.03125
X	0.0	1.0	0.375	0.1731	0.09375
L	0.0	0.0	0.5	0.3600	0.1250
W	0.0	0.0	0.0	0.3411	0.1875
Δ	0.0	0.0	0.0	0.0	0.1875
Σ	0.0	0.0	0.0	0.0	0.3750

resulting self-consistent band results are shown for ZnSe and for Si in Table IV-8. The differences between the 4 and 6 point results are much less for Si than for ZnSe. Differences between the four and six point results for other compounds are usually smaller than those for ZnSe. We see that the Γ point, which is at the center of the zone, is not nearly as representative from the point of view of charge densities as is the L point, which is at an outer face of the zone. The differences between 4 and 6 point energies are around 0.01 eV for Si and 0.1 eV for ZnSe, indicating that zone point averaging has still not converged for ZnSe.

Once we have a suitably averaged valence charge density, we are ready to consider the calculation of new core-state wave functions and energies. Since the core states are calculated from a spherically symmetrized potential, we first spherically symmetrize

TABLE IV-8. Each column gives self-consistent band energies in eV obtained by using only the high symmetry points indicated at the top of the column, with the weightings of Table IV-7.

	Γ	L	Γ_{XL}	Γ_{XLW}	$\Gamma_{XLW}\Delta\Sigma$
<u>ZnSe</u>					
Γ_{15c}	4.64	6.42	6.35	6.66	6.75
Γ_{1c}	1.65	2.75	2.70	2.94	3.01
Γ_{15v}	0.0	0.0	0.0	0.0	0.0
Γ_{1v}	-11.86	-11.93	-11.83	-11.82	-11.81
L_{3c}	5.62	7.11	7.05	7.32	7.41
L_{1c}	2.08	3.59	3.51	3.79	3.88
L_{3v}	-.91	-.66	-.67	-.64	-.63
L_{1v}	-.572	-4.53	-4.57	-4.40	-4.34
L_{1v}	-10.46	-10.81	-10.80	-10.84	-10.85
<u>Si</u>					
Γ_{15c}				2.78	2.79
$\Gamma_{2'c}$				2.76	2.75
$\Gamma_{25'v}$				0.0	0.0
Γ_{1v}				-11.75	-11.74
L_{3c}				3.81	3.83
L_{1c}				1.59	1.60
$L_{3'v}$				-1.18	-1.18
L_{1v}				-6.76	-6.75
$L_{2'v}$				-9.53	-9.53

ρ_{pw} and the corresponding Coulomb potential V_{pw} about each equivalent atom site, using equation 4.3.

$$\begin{aligned}\rho_{pws}^a(\vec{r} - \vec{R}_a) &= \sum_K \rho_{pw}(K) e^{-i\vec{K} \cdot \vec{R}_a} j_o(K|\vec{r} - \vec{R}_a|) \\ V_{pws}^a(\vec{r} - \vec{R}_a) &= 8\pi \sum_K \frac{\rho_{pw}(K)}{K^2} e^{-i\vec{K} \cdot \vec{R}_a} j_o(K|\vec{r} - \vec{R}_a|) \quad 4.5\end{aligned}$$

It is computationally better to calculate the Coulomb potential arising from ρ_{pw} from equation 4.5 rather than from r-space integrals over ρ_{pws} . The densities ρ_c and ρ_{vcs} are already spherically symmetrized.

Madelung Potential

Neighboring cores mask out part of their nuclear charge leaving a part

$$Z_{pw}^a = Z^a - \int (\rho_c + \rho_{vcs}) d\vec{r}. \quad 4.6$$

We must consider the effect of this Z_{pw}^a at all the atom sites of the crystal, together with the charge-compensating $\rho_{pw}(K=0)$. (The contributions of all $\rho_{pw}(K)$ for K not equal to zero are taken into account in equation 4.5). This problem is clearly a Madelung problem to be solved using Ewald sums in both r and K spaces. We shall quickly sketch out the appropriate discussion. A more detailed discussion can be found in reference 29. We must solve the equation

$$\begin{aligned}\nabla^2 V_m &= 8\pi\rho = 8\pi \sum_{a=1}^{Nt} [\delta(\vec{r} - \vec{R}_a) - 1] \\ &= \frac{8\pi}{N\Omega} \sum_{a=1}^{Nt} \sum_{\vec{K} \neq 0} e^{-i\vec{K} \cdot (\vec{r} - \vec{R}_a)} \quad 4.7\end{aligned}$$

A particular solution to this equation is then

$$V_m(\vec{r}) = -\frac{8\pi}{N\Omega} \sum_{a=1}^{Nt} \sum_{\vec{K} \neq 0} \frac{1}{K^2} e^{-i\vec{K} \cdot (\vec{r} - \vec{R}_a)}. \quad 4.8$$

Note that the particular solution has $V_m(K=0)$ set equal to zero. This will not be true for the actual crystalline situation.

However, since the same term enters the valence potential, it results in an overall shift of core and valence states, and can be rigorously neglected if only energy differences are desired. We will discuss later how to calculate $V_m(K=0)$. Various mathematical manipulations are performed on equation 4.8.

$$V_m(\vec{r}) = -\frac{8\pi}{N\Omega} \sum_{a=1}^{Nt} \sum_{\vec{K} \neq 0} \int_0^\infty dy e^{-K^2 y - i\vec{K} \cdot (\vec{r} - \vec{R}_a)} \\ = V_1 + V_2 . \quad 4.9$$

Completing the K sum by adding and subtracting a $K=0$ term so that a Fourier transform can be performed, we obtain

$$V_1(\vec{r}) = -\frac{8\pi}{N\Omega} \int_0^\eta dy \sum_{a=1}^{Nt} \left[\sum_{\vec{K}=0}^{\infty} e^{iK^2 y - i\vec{K} \cdot (\vec{r} - \vec{R}_a)} - 1 \right] \\ = \frac{8\pi\eta t}{\Omega} - 8\pi \sum_{a=1}^{Nt} \int_0^\eta \frac{dy}{(4\pi y)^{3/2}} e^{-\frac{(\vec{r} - \vec{R}_a)^2}{4y}} \quad 4.10$$

In a straightforward manner,

$$V_2(\vec{r}) = -\frac{8\pi}{\Omega} \sum_{\vec{K} \neq 0} \sum_{a=1}^t e^{-i\vec{K} \cdot (\vec{r} - \vec{R}_a)} \int_\eta^\infty e^{-K^2 y} dy \\ = -\frac{8\pi}{\Omega} \sum_{\vec{K} \neq 0} \frac{e^{-K^2 \eta}}{\eta} \sum_{a=1}^t e^{-i\vec{K} \cdot (\vec{r} - \vec{R}_a)} \quad 4.11$$

The final Madelung potential,

$$V_m(\vec{r}) = \frac{8\pi\eta t}{\Omega} - \sum_{a=1}^{Nt} \frac{2}{|\vec{r} - \vec{R}_a|} \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{e^{-x^2}}{|\vec{r} - \vec{R}_a|} dx \\ - \frac{8\pi}{\Omega} \sum_{\vec{K} \neq 0} \frac{e^{-K^2 \eta}}{K^2} \sum_{a=1}^t e^{-i\vec{K} \cdot (\vec{r} - \vec{R}_a)} , \quad 4.12$$

involves sums of error functions over ion sites and sums over reciprocal lattice vectors, and is independent of η . The value of

η merely determines the relative contributions from the r -space and the K -space terms. If the potential is desired at site b and other b type sites all have a unit charge, a limiting procedure gives

$$V_m(\vec{R}_b) = \frac{8\pi n t}{\Omega} + \frac{2}{\sqrt{\pi\eta}} - \sum_{a \neq b}^N \frac{2}{|\vec{R}_a - \vec{R}_b|} \sqrt{\pi} \int_0^\infty \left| \frac{\vec{R}_a - \vec{R}_b}{2\sqrt{\eta}} \right| e^{-x^2} dx \\ - \frac{8\pi}{\Omega} \sum_{K \neq 0} \frac{e^{-K^2\eta}}{K^2} \sum_{a=1}^t e^{-i\vec{K} \cdot (\vec{R}_b - \vec{R}_a)} \quad 4.13$$

Madelung constants defined for unit charge and unit lattice constant are given for the zincblende, sodium chloride and ideal wurtzite ($c=1$, $a=\sqrt{3}/8$) lattice in Table IV-9. For these X-Y type lattices, M_{11} is the potential at an X site due to unit charges on all other X, while M_{12} gives the potential at an X site due to unit charges on all Y sites. The potential at a nucleus due to Z_p^a (equation 4.6) plus the compensating $\rho_{pw}(K=0)$ background is thus known.

The uniform background also produces a parabolic potential energy about each nucleus. Briefly,

$$\nabla^2 V_p(r) = \frac{1}{r^2} \frac{d}{dr} (r^2 \frac{d}{dr}) V_p(r) = -8\pi \rho_{pw}(K=0) \\ V_p(r) = -\frac{4\pi}{3} \rho_{pw}(K=0) r^2 \quad 4.14$$

TABLE IV-9. Madelung constants in Rydbergs for XY compounds with unit lattice constants (for wurtzite, $c=1$, $a=\sqrt{3}/8$). M_{11} is the potential at an X site due to a unit charge on all other X sites. M_{12} gives the potential at an X site due to a unit charge on all Y sites.

Symmetry	M_{11}	M_{12}
Zincblende	9.16972	1.60378
Sodium chloride	9.16972	4.05385
Wurtzite	10.58786	1.83415

This parabola is only valid inside the nearest neighbor distance. Further, the whole situation gets unduly complicated when the neighboring core is reached. Consequently, we stop the r -dependence of V_p at a radius where the nearest-neighbor core becomes significant.

Final Core Potential

The final potential used for the calculation of new core states is then

$$\begin{aligned}
 V^a(r) = & 4\pi \left[-\frac{2Z^a}{r} + \frac{2}{r} \int_0^r 4\pi s^2 (\rho_c + \rho_{vcs}) ds \right. \\
 & \left. + 2 \int_r^\infty 4\pi s (\rho_c + \rho_{vcs}) ds \right] - 6\alpha(3/8\pi)^{1/3} (\rho_c + \rho_{vcs} + \rho_{pws})^{1/3} + V_{pws}(r) \\
 & + \sum_{b=1}^t M_{ab} [Z_b - \int_b 4\pi s^2 ds] - \frac{4\pi}{3} \rho_{pw} (K=0) r^2 , \quad 4.15
 \end{aligned}$$

where the Madelung sum is over types of atoms in the unit cell, and the parabola is cut off as discussed. The core states are then iterated until they are mutually self-consistent.

New core-plane wave orthogonality integrals (equation 2.9) are then calculated from the new core states. In order to calculate these rapidly for the large number of k vectors, a Chebyshev polynomial fit is made along segments of the k -axis.

$\rho^{1/3}$ Fourier Transform

The remaining job is to calculate the Fourier transform of the new crystalline potential. This is straightforward for the Coulomb potential because $\rho_{pw}(K)$ is already in K space, while $\rho_c + \rho_{vcs}$ are spherically symmetrized about ion sites. It is less straightforward, however, for the exchange potential. For that we need the Fourier transform of $\rho^{1/3}(r)$. Because the largest part of $\rho^{1/3}$ is spherically symmetrized about atom sites, and can be accurately Fourier transformed, we consider the quantity.

$$\rho^{1/3} = [\rho^{1/3} - (\rho_c + \rho_{vcs})^{1/3}] + (\rho_c + \rho_{vcs})^{1/3} \quad 4.16$$

The last term will be integrated over a radial atomic mesh as in equation 3.4. The first term of equation 4.16 will be treated using a finite Fourier series which is a straightforward 3-dimensional generalization of the following theorem. Given $f(x_n)$ for the mesh

$$x_n = n/(2N+1), -N \leq n \leq N,$$

the finite Fourier transforms are

$$g(K) = \frac{1}{2N+1} \sum_{n=-N}^N f(x_n) e^{2\pi i K x_n}, -N \leq K \leq N \quad 4.17$$

and

$$f(x_n) = \sum_{K=-N}^N g(K) e^{-2\pi i K x_n}. \quad 4.18$$

The three-dimensional generalization of this theorem is straightforward. To use the theorem, one calculates the quantity

$$f(\vec{r}_{nlm}) = \rho^{1/3} - (\rho_c + \rho_{vcs})^{1/3} \quad 4.19$$

at the points in the irreducible segment of the crystal cell,

$$\vec{r}_{nlm} = (n \vec{T}_1 + l \vec{T}_2 + m \vec{T}_3) / (2N+1), -N \leq n, l, m \leq N, \quad 4.20$$

where the \vec{T}_i are translation vectors of the crystallographic space group. The finite Fourier transform of equation 4.19 is then easily obtained from the three-dimensional generalization of equation 4.17. In our SCOPW calculations, we usually use N equal to 11. There are then $(23)^3 = 12167$ points in the zincblende cell of which 650 are distinct and not connected by symmetry operations. Consequently, equation 4.19 must be calculated at 650 points by summing the K series (equation 4.3) for ρ_{pw} and adding up the contribution of $\rho_c + \rho_{vcs}$ from near atoms.

Final Valence Potential

In a way very analogous to that used to obtain equation 3.4, we then obtain for the new crystal potential,

$$\begin{aligned}
 V(\vec{K}) = & \frac{1}{\Omega} \sum_{a=1}^t e^{i\vec{K} \cdot \vec{R}_a} \left[\frac{8\pi}{K^2} (-Z_a + \int (\rho_c + \rho_{vcs}) 4\pi r^2 j_o(Kr) dr) \right. \\
 & - 6\alpha(6\pi)^{1/3} \int (4\pi (\rho_c + \rho_{vcs}))^{1/3} r^2 j_o(Kr) dr \left. \right] + \frac{8\pi}{K^2} \rho_{pw}(\vec{K}) \\
 & - 6\alpha \left(\frac{3}{8\pi} \right)^{1/3} [\rho^{1/3} - (\rho_c + \rho_{vcs})^{1/3}] (\vec{K}) . \quad 4.21
 \end{aligned}$$

Note that $\rho_{pw}(K)$, which represents about 0.9 of the valence charge density, does contain non-spherical terms. The muffin tin approximation is *not* made. For $K=0$, we again have a singularity coming from the Coulomb terms. Again, one can expand $j_o(Kr)$ as before. That part of Z_a equal to

$$\int (\rho_c + \rho_{vcs}) dr$$

just cancels the singular term of the Coulomb integral. Ionic charge

$$Z^a - \int (\rho_c^a + \rho_{vcs}^a) dr$$

and a uniform background compensating charge $\rho_{pw}(K=0)$ are left over. We encountered this charge configuration before when discussing the Madelung terms for the core calculation. We ignored the contribution there. To be consistent, we must do likewise here. If we were to take it into account in both places, all core and valence energies would merely be shifted by the same amount. We thus get for the $K=0$ Fourier component of the valence potential,

$$\begin{aligned}
 V(K=0) = & \frac{1}{\Omega} \sum_{a=1}^t \left[-\frac{4\pi}{3} \int (\rho_c + \rho_{vcs}) 4\pi r^2 dr \right. \\
 & - 6\alpha (6\pi)^{1/3} \int (4\pi (\rho_c + \rho_{vcs}))^{1/3} r^2 dr \left. \right] \\
 & - 6\alpha \left(\frac{3}{8\pi} \right)^{1/3} [\rho^{1/3} - (\rho_c + \rho_{vcs})^{1/3}] (K = 0) . \quad 4.22
 \end{aligned}$$

Smoothing Procedures

We use a smoothing procedure to obtain the "best" crystal potential coefficients $V^n(K)$ to use for the n th iteration. They

are obtained from the newly calculated $V(K)$ and from the values of $V^{n-1}(K)$ and $V^{n-2}(K)$ from the previous iterations. If $V^{n-2}(K)$, $V^{n-1}(K)$, and $V(K)$ are not monotonic increasing or decreasing, we use

$$V^n(K) = .25V^{n-2}(K) + .5V^{n-1}(K) + .25V(K) \quad 4.23$$

while otherwise we use

$$V^n(K) = (1-\alpha) V^{n-1}(K) + \alpha V(K), \quad 4.24$$

where α is around .8 for groups IV, III-V, and II-VI, and is .1 or .2 for the more difficult NaCl, KCl, Ar and Kr. Because of the different $V(K=0)$ convention for the isolated atom model as compared with the SCOPW model, the full new potential is used for the first iteration, and equation 4.24 with $\alpha = .5$ is used for the second iteration.

The same technique is used in the core state iterations with $V(K)$ replaced by $\rho(r)$. When trouble is encountered (perhaps the outer core state is not bound), equation 4.24 is used to back up towards the previous iteration's successful $\rho(r)$. If there is still trouble, the exchange constant is increased until a solution is found, and is then slowly decreased again.

Madelung $V(K=0)$

Let us now consider more carefully $V(K=0)$ for the case of a unit charge within the unit cell plus a compensating background. This term was arbitrarily assumed zero in the previous development. Consider a large, but finite, crystal made up of N cells. We must calculate

$$V(K=0) = \frac{1}{N\Omega} \int_{\text{crystal}} V(\vec{r}) d\vec{r} \approx \frac{1}{\Omega} \int_{\text{cell}} V(\vec{r}) d\vec{r} \quad 4.25$$

where the equality is approximate because surface cells will have a different $V(r)$ than interior cells. We assume that the crystal is large enough that the large majority of the cells are interior cells. Now $V(r)$ is obtained by integrating over all the charge in the crystal. Since we integrate the field point of $1/r$ over a cell, we can instead freeze the field point and smear the source point over the cell. In doing this, we find that the interior cells do not contribute to $V(K=0)$ as the point positive charge,

when smeared over the cell, just cancels the uniform background negative charge. The whole contribution comes from the boundary regions and is very dependent upon the charge density of the boundary cells. In the idealized case of a uniform background charge up to the boundary, the smeared charge distribution at a boundary becomes that of Figure IV-3. We want the potential in the interior due to this charge density. Assuming a long, narrow crystal, the problem is a straightforward one-dimensional problem with the solution

$$V(K=0) = -\frac{\pi a^2}{3} \rho_{pw} \quad (K=0) \quad 4.26$$

where the lattice constant is a . In actual situations, with non-orthogonal basis vectors and distorted surface charge, the treatment is more complicated, but it is clear in principle how to proceed. This calculation would have to be done correctly if one wanted to give meaning to the absolute values of the energy eigenvalues in order to calculate work functions and the like.

Iteration Experience

The SCOPW iteration cycle is summarized in Table IV-10 in which approximate 7094 II times are also given. It can be seen from the table that the largest amount of time goes into finding new valence wave functions and eigenvalues. The necessity to use scratch tapes for this part of the problem for 459 plane waves (matrices up to 142 by 142) causes the dramatic increase in time over that required for 229 plane waves.

To achieve self-consistency starting from the isolated-atom model potential requires about ten iterations (self-consistency to the extent that the valence eigenvalues change less than 0.01 eV from one iteration to the next). Table IV-11 gives some interesting quantities for the zeroth iteration (the isolated-atom model), plus the first, ninth, and tenth iterations for Slater's exchange approximation using 259 OPW's in ZnS. Only energy differences are meaningful when comparing the isolated-atom values with the SCOPW energies because of the different convention regarding $V(K=0)$. Thus, the entries with asterisks in the table have been given relative to the Γ_{15v} value in each calculation.

One notes first of all that the changes in all the quantities from the isolated-atom model to the first iteration of the SCOPW model are much larger than the changes from the first to the tenth iteration. This, of course, points out again that valence charge rearrangement is very significant in energy band calculations. One of the quantities which varies the most from the isolated-atom

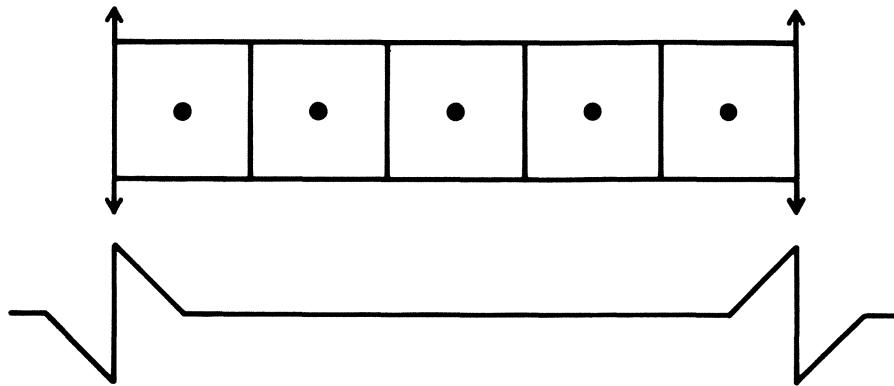


Fig. IV-3. An idealized crystal with one ion per unit cell. The average charge distribution resulting from smearing all charge over a unit cell is also shown.

TABLE IV-10. A self-consistent cycle is summarized with typical IBM 7094 II times in minutes for 229 and 459 OPW's.

7094 II Time in Minutes

229 pw 459 pw

1.	11	82	Valence ψ and E
2.			Valence charge density
	7	15	$\rho_{vc}(r)$
	1.5	6	$\rho_{pw}(K)$
3.	.5	.7	Spherically symmetrize ρ_{pw} for core calculation
4.	.5	.5	New core ψ
5.	3	5.5	$\rho(r)$ over crystal mesh
6.	3	3	Fourier transform $\rho^{1/3}(r)$ over crystal mesh
7.	1	3	Crystal potential $V(K)$
8.	1	3	Core-Plane Wave orthogonality coefficients, $A_{nl}^a(k)$

GO TO 1

— — —
28.5 118.7

model is the 3d energy level of Zn. It lowers relative to the valence bands by about 10 eV as the valence charge density delocalizes, causing less shielding of the nuclear charge.

While convergence is obtained in approximately ten iterations for Group IV, III-V, and II-VI semiconductors, many more iterations are necessary for I-VII compounds such as NaCl and KCl, and for Ar and Kr. In the I-VII's, core overlap becomes too large if the core-valence division of Table II-1 is used, and in the inert gases, one has to consider the entire filled outer shell. In these cases, up to twice as many valence electrons are involved and the valence charge rearrangement progresses very slowly towards self-consistency.

Energy Band Interpolation Through the Zone

Once the band energies have been obtained at selected high symmetry points, it is highly desirable to have a fast, convenient way of interpolating these band energies throughout the Brillouin zone. We have discussed the method we use elsewhere.³⁰ Briefly, separate pseudopotential fits are made to each set of high symmetry point band energies to an accuracy of 0.05 eV. The resulting pseudopotential coefficients vary up to 20 percent throughout the zone. The Shankland³¹ "smoothed" interpolation procedure is used to interpolate the pseudopotential coefficients throughout the zone. Diagonalization of the matrix constructed from the pseudopotentials and the kinetic energy terms then yields the desired interpolated band energies. In this way, the band energies along symmetry lines are generated. The absorptive part of the dielectric constant, ϵ_2 , and the valence and conduction band density of states can also be obtained by calculating the band energies and their gradients with respect to k over a coarse mesh within the first Brillouin zone. A fine mesh is then placed around each coarse mesh point and the energies and gradients are used to find the linearly extrapolated band energies over the fine mesh. At each fine mesh point, energy differences and pseudopotential optical matrix elements (for ϵ_2) can be used to build the appropriate histogram of phase space versus energy. The resulting histogram shapes do not match experiment because of the poor pseudopotential matrix elements and because of the neglect of electron-phonon and electron-hole interactions, but the peaks do occur at the proper energies.

TABLE IV-11. Various quantities are given for the zeroth (isolated atom), first, ninth, and tenth iterations for 259 OPW Slater ZnS. Quantities with an * have had the Γ_{15v} energy subtracted from them.

	IA	IA	1	9	10
L_{1c} (eV)*	7.598		5.019	4.864	4.861
L_{3v} (eV)*	-•.373		-•.389	-•.590	-•.590
Γ_{1c} (eV)*	4.359		3.395	3.595	3.593
Γ_{15v} (eV)*	0.00		0.00	0.00	0.00
Γ_{1v} (eV)*	-11.856		-12.577	-11.718	-11.718
$E(Zn3d)$ (eV)*	-4.283		-7.035	-14.187	-14.187
$E(Zn1s)(\Gamma_V)$ *	-697.487		-697.693	-697.535	-697.535
$E(S1s)(\Gamma_V)$ *	-179.012		-178.795	-178.995	-178.995
$V(000)(\Gamma_V)$ *	-1.0654		-1.0699	-•.4887	-•.4888
Γ_{15v} (eV)	-35.725		-3.241	4.683	4.684
$V(002)(\Gamma_V)$	-•.0632		-•.0694	-•.0698	-•.0698
$V(111)(\Gamma_V)$	-•.4542, •.4009		-•.4622, •.3954	-•.4749, •.3851	-•.4749, •.3850
$\int \rho_{pw}$			9.1528	8.9401	8.9400
$Zn \int \rho_{vc}$			-•.2936	-•.2699	-•.2700
$S \int \rho_{vc}$			-•.8593	-•.6701	-•.6700
$\rho_{pw}(111)$			•.00187, -•.01470	.00189, -•.01342	•.00190, -•.01342
$\rho_{pw}(002)$			-•.01050	-•.00955	-•.00955

*quantities are given relative to the Γ_{15v} level

V. SUMMARY

In this article, we have given detailed theoretical and calculational bases of our SCOPW model. This model generates a realistic (non-muffin-tin plus charge rearrangement) crystal potential associated with the assumed Hamiltonian of the system. Numerical examples were presented to give the reader a feeling for the magnitude of the various approximations made within the model. The major mathematical approximations are: (1) convergence of the OPW series; (2) neglect of core overlap; (3) spherical symmetrization of ρ_{vc} ; and (4) Brillouin zone averaging of the valence charge density. When these approximations have been investigated and shown to be small, one can then use the calculational model to test the two physical approximations of the model: (1) neglect of relativity and (2) use of a local approximation to simulate the many-body exchange-correlation-relaxation effects with a one-electron effective Hamiltonian. Dr. G. Wepfer has joined us in developing a relativistic version of the SCOPW model. When this work is complete, we will have isolated the exchange-correlation-relaxation problem.

Each of the various mathematical approximations contributes roughly a tenth of an eV uncertainty to the final energy eigenvalues. The SCOPW eigenvalues should thus be within a few tenths of an eV of the correct mathematical eigenvalues of the effective Hamiltonian. Non-relativistic calculations on many different semiconducting compounds have shown that the SCOPW eigenvalues obtained using the Slater exchange usually agree within a few tenths of an eV of the experimental excitation energies. The SCOPW charge densities are rather poor for all the best-known exchange approximations.

The SCOPW model is a very useful tool for studying the exchange-correlation problem, and for calculating band structures. To the extent that it gives an accurate mathematical picture of the electrons in a perfect crystal, it should prove useful in many other applications as well.

ACKNOWLEDGMENTS

We gratefully acknowledge the significant contributions that Dr. D. G. Shankland has made to the SCOPW model. He brought the Choleski decomposition to our attention, and furnished us with programs to calculate L and L^{-1} . He also provided us with a very efficient matrix diagonalization program, contributed to the Madelung calculations, and showed us his interpolation procedure.

Major J. S. DeWitt was instrumental in setting up the pseudo-potential programs which interpolate the SCOPW high-symmetry-point band energies throughout the Brillouin zone in the wurtzite crystals.

We also thank Dr. F. Herman and co-workers for many useful exchanges of information and comparison of isolated-atoms results. The Herman-Skillman atomic programs³² are an integral part of the SCOPW system. The atomic programs are used per se in our isolated-atom coding; and, the differential equation solver, SCHEQ, is used for the calculation of new core states.

We have received high-quality programming assistance from Messrs. L. Zaytoun, A. O'Hare, D. Carpenter, J. Thieman, and R. Schultz. The Wright-Patterson Computing and Information Systems Division personnel, under the direction of Mr. D. Zonars, have been highly cooperative.

Finally, we would like to thank Mrs. J. Ingvoldstad for her very intelligent editorial advice, and for her patient and highly-competent typing of this manuscript. We are also indebted to Mr. E. Johnson for drafting the figures for this manuscript.

REFERENCES

1. C. Herring, "A New Method for Calculating Wave Functions in Crystals", Phys. Rev. 57, 1169 (1940)
2. F. Herman, S. Skillman, "Theoretical Investigation of the Energy Band Structure of Semiconductors", Proc. Int'l Conf. on Semiconductor Phys., Prague, 1960
(Publishing House of the Czechoslovak Acad. Sci.)
3. F. Herman, R. L. Kortum, C. D. Kuglin, J. P. Van Dyke, and S. Skillman, "Electronic Structure of Tetrahedrally-Bonded Semiconductors: Empirically Adjusted OPW Energy Band Calculations", Methods in Computational Physics, B. Alder, S. Fernbach, M. Rotenberg, eds. (Academic Press, New York, 1968), Vol. 8, pp 193.
4. R. N. Euwema, T. C. Collins, D. Shankland, and J. S. DeWitt, "Convergence Study of a Self-Consistent Orthogonalized-Plane-Wave Band Calculation for Hexagonal CdS", Phys. Rev. 162, 710 (1967)

5. D. J. Stukel, R. N. Euwema, "Self-Consistent Orthogonalized-Plane-Wave Energy-Band Study of Silicon", Phys. Rev. B1, 1635 (1970)
6. D. J. Stukel, "Self-Consistent Energy Bands and Related Properties of Boron Phosphide", Phys. Rev. B, 15 April (1970)
7. D. J. Stukel, "Electronic Structure and Optical Spectrum of Boron Arsenide", Phys. Rev. B, 15 June (1970)
8. D. J. Stukel, R. N. Euwema, "Electronic Band Structure and Related Properties of Cubic AlP", Phys. Rev. 186, 754 (1969)
9. D. J. Stukel, R. N. Euwema, "Energy Band Structure of Aluminum Arsenide", Phys. Rev. 188, 1193 (1969)
10. T. C. Collins, D. J. Stukel, R. N. Euwema, "Self-Consistent Orthogonalized-Plane-Wave Band Calculation on GaAs", Phys. Rev. B1, 724 (1970)
11. D. J. Stukel, "Energy Band Structure of BeS, BeSe and BeTe", Phys. Rev. B (1970)
12. D. J. Stukel, R. N. Euwema, T. C. Collins, F. Herman, R. L. Kortum, "Self-Consistent Orthogonalized-Plane-Wave Energy-Band Models for Cubic ZnS, ZnSe, CdS, and CdSe", Phys. Rev. 179, 740 (1969)
13. R. A. Deegan, W. D. Twose, "Modifications to the Orthogonalized-Plane-Wave Method for Use in Transition Metals: Electronic Band Structure of Niobium", Phys. Rev. 164, 993 (1967)
F. A. Butler, F. K. Bloom, E. Brown, "Modifications of the Orthogonalized-Plane-Wave Method Applied to Copper", Phys. Rev. 180, 744 (1969)
14. T. O. Woodruff, "The Orthogonalized Plane-Wave Method", Solid State Physics 4, 367, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1957)
15. B. Wendroff, Theoretical Numerical Analysis, (Academic Press, 1966)
J. Wilkinson, Algebraic Eigenvalue Problem, (Oxford University Press, 1966)
16. J. H. Wilkinson, "Householder's Method for the Solution of the Algebraic Eigenproblem", Computer Journal 3, 23 (1960)
17. R. N. Euwema, D. J. Stukel, "OPW Convergence of Some Tetrahedral Semiconductors", Phys. Rev. B2, 15 June (1970)

18. J. Callaway, "Orthogonalized Plane Wave Method", Phys. Rev. 97, 933 (1955)
V. Heine, "The Band Structure of Aluminum I, II, III", Proc. Roy. Soc. A240, 340, 354, 363
F. Herman, "Calculation of Energy Band Structure of Diamond and Ge Crystals by Method of Orthogonalized Plane Waves", Phys. Rev. 93, 1214 (1955)
19. J. C. Slater, "A Simplification of the Hartree-Fock Method", Phys. Rev. 81, 385 (1951)
20. W. Kohn, L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects", Phys. Rev. 140, 1133 (1965)
21. R. Gaspar, "Über eine Approximation des Hartree-Fockschen Potentials durch eine Universelle Potentialfunktion", Acta. Phys. Acad. Sci. Hung. 3, 263 (1954)
22. D. A. Liberman, "Exchange Potential for Electrons in Atoms and Solids", Phys. Rev. 171, 1 (1968)
L. J. Sham, W. Kohn, "One-Particle Properties of an Inhomogeneous Interacting Electron Gas", Phys. Rev. 145, 561 (1966)
23. D. J. Stukel, R. N. Euwema, T. C. Collins, V. H. Smith, "Exchange Study of Atomic Krypton and Tetrahedral Semiconductors", Phys. Rev. B1, 779 (1970)
24. J. C. Slater, "Present Status of the X_α Statistical Exchange", M.I.T. Semi-Annual Progress Report No. 71, p 1 (1969)
25. P. M. Raccah, R. N. Euwema, D. J. Stukel, T. C. Collins, "Comparison of Theoretical and Experimental Charge Densities for C, Si, Ge, and ZnSe", Phys. Rev. B1, 756 (1970)
26. J. C. Slater, Quantum Theory of Molecules and Solids, Vol. II, (McGraw Hill Book Co., 1965)
27. G. Wepfer, T. C. Collins, R. N. Euwema, D. J. Stukel, "Symmetrization Techniques in Relativistic OPW Energy Band Calculations", this volume
28. We are grateful to G. Wepfer and J. Van Dyke for helpful discussions on this point
29. G. Liebfried, Encyclopedia of Physics, Vol. 7, Part 1, p 132 (Springer-Verlag, Berlin, 1955)

30. R. N. Euwema, D. J. Stukel, T. C. Collins, J. S. DeWitt, D. G. Shankland, "Crystalline Interpolation with Applications to Brillouin-Zone Averages and Energy-Band Interpolation", *Phys. Rev.* 178, 1419 (1969)
31. D. G. Shankland, "Interpolation and Fourier Transformation with Functions Having Specified Smoothness Properties", Boeing Document D6-29730, The Boeing Company, Renton, Wash. D. G. Shankland, "k-Space Interpolation of Functions of Arbitrary Smoothness", this volume
32. F. Herman, S. Skillman, Atomic Structure Calculations, (Prentice-Hall, Inc. 1963)

SYMMETRIZATION TECHNIQUES IN RELATIVISTIC OPW ENERGY BAND
CALCULATIONS

G. G. Wepfer, T. C. Collins, R. N. Euwema and
D. J. Stukel

Aerospace Research Laboratories

Wright-Patterson Air Force Base, Ohio

I. INTRODUCTION

Definite advantages can be attained by the use of crystalline symmetry properties in calculations of energy bands. In particular, dramatic savings in computer time and core storage requirements can be realized. In this paper, we give some observations relevant to the Orthogonalized-Plane-Wave (OPW) method. Many of these observations are equally relevant to other computational techniques. Mattheiss, Wood, and Switendick¹ have outlined the use of symmetry in the Augmented-Plane-Wave (APW) method. A well-known disadvantage of the OPW method is the slow convergence of the wave function expansion.² Our experience with group IV, III-V, and II-VI semiconducting crystals indicates that at least 229 plane waves must be used (usually many more) to obtain convergence. Hence, an eigenvalue problem of at least order 229 must be solved to find eigenvalues and eigenvectors. This is a time-consuming task on the computer. The problem is even more difficult in the relativistic OPW formalism where spin doubles the order of the matrices.

The introduction of symmetry properties, i.e., the use of symmetrized OPW's, gives a very real savings in calculation time, and permits the solution of large secular equations in spite of limited core size. As a general rule, the time necessary to diagonalize a large matrix of order N is proportional to N^3 . If the matrix can be reduced to block diagonal form with M equal blocks along the diagonal, the order of each block is N/M . The time required to diagonalize each block is proportional to $(N/M)^3$. There are no more than M blocks to solve since only one block needs to be diagonalized for each irreducible representation. Therefore,

the ratio of the time necessary to solve the block diagonal form to the time necessary to solve the original matrix can be taken as

$$t \propto \frac{Mx(\frac{N}{M})^3}{N^3} = \frac{1}{M^2} . \quad 1.1$$

As an example, there are five irreducible representations at the Γ point in the zincblende symmetry (T_d^2). Since the sum of the dimensionality of the representations at the Γ point is ten, the Hamiltonian matrix can be reduced to ten blocks along the diagonal, five of which must be solved. If the blocks were all of the same order (which they are not), one might expect a calculation time which is .005 of that required to diagonalize the original matrix.

There are additional advantages to the use of symmetrized OPW's. An obvious, but useful benefit is that the results of symmetrized calculations automatically characterize the states as to symmetry. Furthermore, two quantities of particular interest, e.g., the charge density and the crystal potential, have the full symmetry of the crystal. Therefore, all of the Γ_1 - like information that can be extracted from the square of the basis functions of a multi-dimensional representation is contained in any one basis function. Consequently, only one row of any multi-dimensional representation need be considered.

The purpose of this paper is to present the techniques by which well-known symmetry properties are incorporated into the OPW formalism. The background information necessary to accomplish this and many of the calculational techniques are readily available.^{1,3} Luehrmann has provided a very complete set of tables for forming linear combinations of symmetrized plane waves for all crystalline symmetries.⁴ The actual computational techniques will be emphasized here, in order to provide a prescription for the use of symmetry properties. The zincblende crystalline symmetry will be used throughout as a specific example. Following the development of the methods for the single and the double T_d groups (non-relativistic and relativistic OPW treatment), specific examples of comparative times are given in the final section. The savings in computer time are most impressive in the relativistic OPW formalism. In this work, symmetrization is essential because of computer core storage and time limitations.

II. OPW FORMALISM

The fundamental problem in calculating energy bands in crystalline solids is, of course, to solve the eigenvalue-eigenvector problem

$$H_{\text{eff}} \psi_i(\vec{r}) = E_i \psi_i(\vec{r}) . \quad 2.1$$

From the variational principle, we have

$$W[\psi_{\text{TRIAL}}] = \frac{\langle \psi_{\text{TRIAL}} | H_{\text{eff}} | \psi_{\text{TRIAL}} \rangle}{\langle \psi_{\text{TRIAL}} | \psi_{\text{TRIAL}} \rangle} \geq E_0 \quad 2.2$$

where E_0 is the lowest eigenvalue of H_{eff} and ψ_{TRIAL} is any normalizable function satisfying the continuity and boundary conditions which are imposed on a wave function. This principle is used to obtain estimates of E_0 and the associated wave function ψ_0 by using a ψ_{TRIAL} with disposable parameters, and varying the parameters so as to minimize $W[\psi_{\text{TRIAL}}]$. This minimum value is taken as an estimate of E_0 . With this choice of parameters, ψ_{TRIAL} is an estimate of ψ_0 . A solution with a higher energy, E_i , can be approximated in the same way by a trial function which is made orthogonal to each of the true wave functions associated with a lower energy.

In the OPW method, the trial function, ψ_{TRIAL} , is taken as a finite, linear combination of plane waves which have been made orthogonal to all the crystal core-wave functions by the subtraction of appropriate multiples of the core-wave functions (hence, the term "orthogonalized-plane-wave"). The coefficients in this linear combination are the variable parameters in the variational technique. The plane waves in the non-relativistic formalism are taken as

$$\chi = \frac{1}{\sqrt{N\Omega}} e^{i\vec{k}\cdot\vec{r}} \quad 2.3$$

where Ω is the volume of a unit cell and N is the number of unit cells in the crystal. An orthogonality operator \hat{O} is defined as

$$\hat{O} = \hat{1} - \sum_c |\psi_c\rangle \langle \psi_c| \quad 2.4$$

where the sum is over all core states on all atoms. This operator has the property that, if it is applied to a plane wave, it creates an orthogonalized-plane-wave,

$$OPW = \hat{O}|\chi\rangle = |\chi\rangle - \sum_c |c\rangle \langle c| \chi \rangle . \quad 2.5$$

These functions are used in a trial function of the form

$$|\psi_{TRIAL}\rangle = \sum_{j=1}^s a_j \hat{O} |x_j\rangle . \quad 2.6$$

Minimizing $W[\psi_{TRIAL}]$, one finds that the matrix eigenvalue-eigenfunction equations which determine the a_j are

$$\sum_{j=1}^s [\langle \hat{O}x_i | H_{eff} | \hat{O}x_j \rangle - \lambda \langle \hat{O}x_i | \hat{O}x_j \rangle] a_j = 0 . \quad 2.7$$

This is a matrix equation of the form

$$H_{eff} \underline{B} = \lambda \underline{U} \underline{B} . \quad 2.8$$

If one uses wave vectors (values of \bar{k}) in the OPW trial function which have special crystalline symmetry properties, the \underline{U} and H_{eff} matrices in Eq. 2.8 factor into submatrices. This is accomplished by classifying the OPW's contained in the trial function of Eq. 2.6 and finding symmetric linear combinations in accordance with the principles of the representation theory of space groups. Each submatrix can be associated with a trial wave function of a particular symmetry. The eigenvalues and eigenvectors of Eq. 2.8 are then found by separately solving the various submatrices. By means of this symmetry factorization, the order of the matrix equations which must be solved is greatly reduced.

In operating on a plane wave with all the operators of the T_d^2 space group, the result will yield not more than 24 different plane waves; one coming from each of the operations of the T_d point group. If \bar{k} is a general point in reciprocal space, with no special symmetry properties, there will be exactly 24 different plane waves, their wave vectors forming a star. These 24 plane waves form basis functions for a 24-dimensional, irreducible representation of the space group. If \bar{k} is a point with special symmetry in the Brillouin zone, the representation becomes reducible. In fact, some stars may contain fewer than 24 distinct wave vectors. The subgroup of operations of the space group which transforms a wave vector into itself is the group of the wave vector. One can make linear combinations of basis functions corresponding to the same wave vector, and obtain basis sets for irreducible representations of the group of the wave vector.

These linear combinations of plane waves are the symmetrized plane waves required for the symmetrized OPW method. Their usefulness lies in the fact that the Hamiltonian does not have matrix elements between different, irreducible representations of the group of the wave vector.

The symmetrized plane waves are found by the projection operator method. In the non-relativistic OPW method, the projection operator is defined as

$$\hat{P} = \frac{d}{g} \sum_{R \in G} \Gamma_p^*(R)_{rr} \hat{R} . \quad 2.9$$

The summation is over all operations \hat{R} of the space group G . $\Gamma_p(R)_{rr}$ is the matrix element of the p th irreducible, unitary representation of the group G . The multiplicative factor involves d , the dimensionality of the p th representation, and g , the order of the group, and insures that $\hat{P}^2 = P$.

The operation of such a projection operator on a plane wave yields a sum of plane waves

$$\begin{aligned} \hat{P} e^{i\vec{k} \cdot \vec{r}} &= \frac{d}{g} \sum_{R \in G} \Gamma_p^*(R)_{rr} \hat{R} e^{i\vec{k} \cdot \vec{r}} \\ &= \frac{d}{g} \sum_{R \in G} \Gamma_p^*(R)_{rr} e^{i\vec{k} \cdot \hat{R}\vec{r}} \\ &= \frac{d}{g} \sum_{R \in G} \Gamma_p^*(R)_{rr} e^{i\vec{R}^{-1} \vec{k} \cdot \vec{r}} \\ &= \frac{d}{g} \sum_{R \in G} \Gamma_p^*(R^{-1})_{rr} e^{i\hat{R}\vec{k} \cdot \vec{r}} \quad 2.10 \end{aligned}$$

The convention used here, with regard to the effect of a rotation on the spatial part of the plane wave function, is consistent with that used by Slater.³ His tables can then be used for the matrix elements $\Gamma_p(R^{-1})_{rr}$. It should be noted that a different convention is often used.^{1,4}

As was pointed out earlier, it is necessary to find only one basis set for an irreducible representation because physically meaningful quantities (such as the Hamiltonian, charge density, and potential) are invariant under all the operations of the space group. One is interested in matrix elements of the form

$$\langle (\hat{P} \hat{\chi})_{pr} | \hat{T} | (\hat{P} \hat{\chi})_{p'r'} \rangle .$$

One can make use of the orthogonality theorem to show that

$$\begin{aligned} & \langle (\hat{P} \hat{\chi})_{pr} | \hat{T} | (\hat{P} \hat{\chi})_{p'r'} \rangle \\ &= \frac{g}{d} \delta_{pp'} \delta_{rr'} \langle \chi | \hat{T} | (\hat{P} \hat{\chi})_{pr} \rangle . \end{aligned} \quad 2.11$$

Hence, there are no matrix elements of \hat{T} between symmetrized OPW's transforming according to different, irreducible representations. In addition, there are no matrix elements between symmetrized OPW's transforming according to different rows of the same irreducible representation. We have chosen, therefore, a projection operator of the form

$$\hat{P} = \frac{d}{g} \sum_{R \in G} \Gamma_p^* (R^{-1})_{11} R^{-1} \quad 2.12$$

Equation 2.11 shows that this projection operator needs to appear on only one side of expressions involving the \underline{U} and \underline{H} matrices (Eq. 2.8). In self-consistent calculations where $\rho(k)$ is needed, the projection operator again needs to appear on only one side of $|\psi|^2$ if Γ_1 symmetry is explicitly projected out.

The symmetrization techniques are similar in the relativistic OPW formalism. The Dirac Hamiltonian for an electron in a potential field $V(\bar{r})$ is

$$\hat{H} = -\frac{2}{q} i \hat{\underline{\alpha}} \cdot \bar{\nabla} + \frac{2}{q^2} \hat{\underline{\beta}} + \hat{\underline{\underline{I}}}_4 V(\bar{r}) , \quad 2.13$$

where q is the fine structure constant and $\hat{\underline{\alpha}}$, $\hat{\underline{\beta}}$, and $\hat{\underline{\underline{I}}}_4$ are the operators

$$\hat{\underline{\alpha}} = \begin{pmatrix} 0 & \bar{\sigma} \\ \sigma & 0 \end{pmatrix}, \quad \hat{\underline{\beta}} = \begin{pmatrix} \hat{I}_2 & 0 \\ 0 & -\hat{I}_2 \end{pmatrix}, \quad \hat{\underline{\underline{I}}}_4 = \begin{pmatrix} \hat{I}_2 & 0 \\ 0 & \hat{I}_2 \end{pmatrix}.$$

The three 2×2 Pauli matrices are represented by $\vec{\sigma}$, and \hat{I}_2 is the 2×2 unit matrix. In order to find the eigenvalues of this Hamiltonian by the method of orthogonalized-plane-waves, it is assumed that the eigenfunctions are Slater determinants of one-electron wave functions, as in the non-relativistic problem. The states are four component functions in which the equations satisfied by the large and small components of the radial wave functions are found by the variational principle. The complete, one-electron wave functions are again taken to be a finite linear combination of plane waves which are orthogonalized to the core states.

Following the development of Herman, Kortum, Ortenburger, and Van Dyke⁶, the functions chosen for the OPW method are the plane-wave spinors with wave vector \vec{k} and spin s

$$|\vec{k}, s\rangle = \frac{1}{\sqrt{N\Omega}} Q^{(s)}(\vec{k}) e^{i\vec{k} \cdot \vec{r}} \quad 2.14$$

In this expression,

$$Q^{(s)}(\vec{k}) = \begin{pmatrix} \sqrt{(E_0(k) + 2/q^2)/2E_0(k)} \chi^{(s)} \\ \sqrt{(E_0(k) - 2/q^2)/2E_0(k)} (\vec{\sigma} \cdot \vec{k}/k) \chi^{(s)} \end{pmatrix}$$

The free particle energy is given by

$$E_0(k) = \frac{2}{q^2} [1 + (kq)^2]^{1/2}$$

and $\chi^{(s)}$ is a spin eigenfunction

$$\chi^{(1/2)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \chi^{(-1/2)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

These spin functions are commonly called α and β respectively. This designation will be used here, and should not be confused with the matrices in the Dirac Hamiltonian.

The appropriate projection operator is altered somewhat from the non-relativistic case, Eq. 2.12. The plane-wave spinors are composed of a spatial part and a spin part. Hence, one also must consider the effect of the operations on a spin function. A convenient form is

$$\hat{P} = \frac{d}{g} \sum_{R \in G} \Gamma_p * (R^{-1})_{11} (\hat{R}^r)^{-1} (\hat{R}^s)^{-1} \quad 2.15$$

where the R^r operates on the spatial part of a plane wave, $e^{ik \cdot r}$, and R^s operates on a spin function, $\chi(s)$.

Of course, the appropriate double groups have to be used in generating the symmetrized OPW's in the relativistic problem. The application of symmetry in the relativistic formalism is particularly useful because the sizes of the matrices which appear in Eq. 2.8 are doubled due to two spin states associated with each plane wave.

III. CALCULATION TECHNIQUES

k-Vectors

There is some work which is preliminary to finding symmetrized OPW's and associating them with the proper irreducible representation. It is necessary to create reciprocal lattice vectors which set up the plane waves as in Eq. 2.3 or Eq. 2.14. As indicated in the discussion on the projection operator, we will be operating on k-vectors with the operations of the appropriate point group. In particular, the discussion will be restricted to the T_d group of the zincblende crystalline symmetry.

We generate 1500 k-vectors at each of the high symmetry points Γ , X, L, and W in the Brillouin zone. A k-vector at the Γ point is defined by

$$\bar{k}_h = \frac{2\pi}{a} (h_1 \bar{b}_1 + h_2 \bar{b}_2 + h_3 \bar{b}_3)$$

where a is the crystalline lattice constant; h_1 , h_2 , and h_3 are integers and \bar{b}_1 , \bar{b}_2 , \bar{b}_3 are the fundamental vectors of the reciprocal lattice. In Cartesian coordinates,

$$\bar{k}_h = k_x \bar{i} + k_y \bar{j} + k_z \bar{k} ,$$

where

$$k_x = \frac{2\pi}{a} (-h_1 + h_2 + h_3) ,$$

TABLE I. \mathbf{k} -vectors at Γ point

Number	Star	k_x	k_y	k_z	Number	Star	k_x	k_y	k_z
1	1	-0	-0	-0	28	6	1	1	3
2	2	1	1	1	29		1	3	1
3		1	-1	-1	30		3	1	1
4		-1	-1	1	31		1	-1	-3
5		-1	1	-1	32		-1	-3	1
6	3	-1	-1	-1	33		-3	1	-1
7		-1	1	1	34		1	-3	-1
8		1	1	-1	35		-3	-1	1
9		1	-1	1	36		-1	1	-3
10	4	0	0	2	37		-1	-1	3
11		0	2	0	38		-1	3	-1
12		2	0	0	39		3	-1	-1
13		0	-0	-2	40	7	-1	-1	-3
14		-0	-2	0	41		-1	-3	-1
15		-2	0	-0	42		-3	-1	-1
16	5	2	-0	2	43		-1	1	3
17		-0	2	2	44		1	3	-1
18		2	2	-0	45		3	-1	1
19		-0	-2	-2	46		-1	3	1
20		-2	-2	-0	47		3	1	-1
21		-2	-0	-2	48		1	-1	3
22		-2	0	2	49		1	1	-3
23		0	2	-2	50		1	-3	1
24		2	-2	0	51		-3	1	1
25		-2	2	0	52	8	2	2	2
26		2	0	-2	53		2	-2	-2
27		0	-2	2	54		-2	-2	2
					55		-2	2	-2

TABLE I. k-vectors at Γ point (Continued)

Number	Star	k_x	k_y	k_z	Number	Star	k_x	k_y	k_z
56	9	-2	-2	-2	84		3	1	-3
57		-2	2	2	85		1	-3	3
58		2	2	-2	86		-3	3	1
59		2	-2	2	87		3	-3	1
					88		-3	1	3
60	10	0	0	4	89		1	3	-3
61		0	4	0					
62		4	0	0	90	13	2	-0	4
63		0	-0	-4	91		-0	4	2
64		-0	-4	0	92		4	2	-0
65		-4	0	-0	93		2	4	-0
					94		4	-0	2
66	11	3	1	3	95		-0	2	4
67		1	3	3	96		-0	-2	-4
68		3	3	1	97		-2	-4	-0
69		1	-3	-3	98		-4	-0	-2
70		-3	-3	1	99		-0	-4	-2
71		-3	1	-3	100		-4	-2	-0
72		-3	-1	3	101		-2	-0	-4
73		-1	3	-3	102		-2	0	4
74		3	-3	-1	103		0	4	-2
75		-3	3	-1	104		4	-2	0
76		3	-1	-3	105		-2	4	0
77		-1	-3	3	106		4	0	-2
					107		0	-2	4
78	12	-3	-1	-3	108		0	2	-4
79		-1	-3	-3	109		2	-4	0
80		-3	-3	-1	110		-4	0	2
81		-1	3	3	111		0	-4	2
82		3	3	-1	112		-4	2	0
83		3	-1	3	113		2	0	-4

$$k_y = \frac{2\pi}{a} (h_1 - h_2 + h_3) ,$$

$$k_z = \frac{2\pi}{a} (h_1 + h_2 - h_3) .$$

To determine the k-vectors at the X, L, and W points, the coordinates of each of these points are added to those of the k-vectors at the Γ point. The vectors are generated by a general sort program, and grouped by stars along with the values of k^2 . Table I gives the first 113 k-vectors at the Γ point for zincblende. This table is sufficient to include typical stars of all sizes, 1, 4, 6, 12, and 24. For future use, only the first member of each star is stored. In our treatment, this will be sufficient as will become clear shortly.

SCRAMBLE Subroutine

An essential part of the symmetrized OPW method is a subroutine which we call SCRAMBLE. This subroutine takes one member of a k-vector star and generates all the other members of the star with the group operations of the appropriate Brillouin zone point. It also combines $\Gamma_{11}^{**}(R^{-1})$ coefficients of duplicate k-vectors so that one gets the correct number of members of the star together with the symmetry coefficients of the plane waves. The subroutine thus generates the symmetrized plane waves from a given plane wave. In the relativistic SCRAMBLE, some additional data is needed to include the operations of the group on a spin function α or β as in Eq. 2.15. This information is included in the form of coefficients of α and β due to the operations of the group on α or β . The group operations mix spin states so that four data sets are needed. The output of the relativistic SCRAMBLE subroutine consists of all members of the star, the number of k-vectors in the star, and two projection operator coefficients $D(1,IC)$ and $D(2,IC)$ where the first multiplies plane waves with spin α and the second multiplies plane waves with spin β . The input and output information for the relativistic SCRAMBLE subroutine can be summarized as follows:

Input to Subroutine

1. Γ , X, L, or W high symmetry point
2. Desired representation at the high symmetry point
3. Starting \bar{k} of a star
4. Spin state (α or β) of starting member

Data Within Subroutine

1. Γ_{11}^* matrix elements for all representations of all groups (Γ , X, L, W).
2. Coefficients of α and β due to the operations of the group on α or β .

Output From Subroutine

1. All distinct members of the star
2. Number of distinct k-vectors in the star (IC)
3. Projection operators in the form of coefficients of spin states, so that

$$\begin{aligned}\hat{P} e^{ik \cdot \bar{r}} &= \frac{d}{g} \sum_{R \in G} \Gamma_{11}^* (R^{-1}) (\hat{R}_{\alpha}^{-1}) e^{i \hat{R} \bar{k} \cdot \bar{r}} \\ &= \sum_{I=1}^{IC} [D(1, I)\alpha + D(2, I)\beta] e^{i \hat{R}_I \bar{k} \cdot \bar{r}}\end{aligned}$$

The D's are complex numbers and depend on the input k-vector and spin.

Linearly Independent Vectors

In order to create a complete basis set of symmetrized OPW's generated by the projection operators for an irreducible representation, it is necessary to start with a sufficient number of k-vectors. The first members of each star are, in general, not enough. The first-member star list must be augmented. The use of every member of the star is redundant since some of the resulting functions will be linearly dependent. Therefore, one can look for the maximum linearly independent set of symmetrized vectors within each star. This number is related to the dimensionality of the irreducible representations and the size of the star. Perhaps the most straightforward, calculational technique is to operate on each member of a star, in turn, with a projection operator. This can be accomplished with the SCRAMBLE subroutine. The symmetrized plane wave which results is compared to the previous ones; if it is linearly independent, it is retained. This is a non-trivial process, but fortunately, it only has to be done once for each type star for each representation at the four high symmetry zone points.

The results of this process are given for the zincblende symmetry in Tables II, III, IV, and V. Tables II and III give a list of the starting members which can be used to generate linearly independent, symmetrized vectors for the Γ point for the first

TABLE II. Linearly independent k -vectors at the Γ point for the single group

Wave Vector	Star	k_x	k_y	k_z	Expanded First Star List
1	1	-0	-0	-0	1
2	2	1	1	1	2
6	3	-1	-1	-1	3
10	4	0	0	2	4
12		2	0	0	5
16	5	2	-0	2	6
17		-0	2	2	7
28	6	1	1	3	8
30		3	1	1	9
40	7	-1	-1	-3	10
42		-3	-1	-1	11
52	8	2	2	2	12
56	9	-2	-2	-2	13
60	10	0	0	4	14
62		4	0	0	15
66	11	3	1	3	16
67		1	3	3	17
78	12	-3	-1	-3	18
79		-1	-3	-3	19
90	13	2	-0	4	20
91		-0	4	2	21
92		4	2	-0	22
105		-2	4	0	23

thirteen stars for the single group and the double group. The vectors which are then used to form the symmetrized OPW's are entered in an expanded first star list. These tables are sufficient to establish the pattern for all Γ -point stars. Tables IV and V provide the same listings for the X, L, and W high symmetry points for both the single and double groups. The latter two tables are much shorter than those at the Γ point, but are sufficient to illustrate stars of all sizes at X, L, and W. Specifically, only 2, 4, and 8 member stars occur at the X point. Table IV indicates that for the non-relativistic problem (single group), one needs to use only the first member of 2 and 4 member stars and two members of 8 member stars. Similarly, one can use the first vector of 1 and 3 member stars at the L point and two vectors of a 6 member star. All of the stars at the W point have 4 members. The first k -vector of each star is all that is needed since all representations are one dimensional.

TABLE III. Linearly independent \mathbf{k} -vectors at the Γ point for the double group

Wave Vector	Star	k_x	k_y	k_z	Spin State	Expanded First Star List
1	1	-0	-0	-0	α	1
2	2	1	1	1	α	2
6	3	-1	-1	-1	α	3
10	4	0	0	2	α	4
10		0	0	2	β	5
11		0	2	0	α	6
16	5	2	-0	2	α	7
17		-0	2	2	α	8
17		-0	2	2	β	9
18		2	2	-0	α	10
28	6	1	1	3	α	11
29		1	3	1	α	12
29		1	3	1	β	13
30		3	1	1	α	14
40	7	-1	-1	-3	α	15
41		-1	-3	-1	α	16
41		-1	-3	-1	β	17
42		-3	-1	-1	α	18
52	8	2	2	2	α	19
56	9	-2	-2	-2	α	20
60	10	0	0	4	α	21
60		0	0	4	β	22
61		0	4	0	α	23
66	11	3	1	3	α	24
67		1	3	3	α	25
67		1	3	3	β	26
68		3	3	1	α	27
78	12	-3	-1	-3	α	28
79		-1	-3	-3	α	29
79		-1	-3	-3	β	30
80		-3	-3	-1	α	31
90	13	2	-0	4	α	32
91		-0	4	2	α	33
91		-0	4	2	β	34
92		4	2	-0	α	35
92		4	2	-0	β	36
93		2	4	-0	α	37
94		4	-0	2	α	38
95		-0	2	4	α	39

TABLE IV. Linearly independent k -vectors at X, L, and W for the single group

Wave Vector	Star	k_x	k_y	k_z	Expanded First Star List
X POINT					
1	1	1	-0	-0	1
3	2	-0	-1	1	2
7	3	1	0	2	3
8		-1	2	-0	4
L POINT					
1	1	.5	.5	.5	1
2	2	-.5	-.5	-.5	2
3	3	1.5	-.5	-.5	3
6	4	.5	.5	-1.5	4
9	5	-.5	1.5	1.5	5
12	6	.5	-1.5	-1.5	6
15	7	1.5	1.5	1.5	7
16	8	.5	.5	2.5	8
19	9	-.5	-.5	-2.5	9
22	10	-1.5	-1.5	-1.5	10
23	11	-1.5	.5	2.5	11
24		.5	-1.5	2.5	12
W POINT					
1	1	1	.5	-0	1
5	2	-0	1.5	-1	2
9	3	2	-.5	-1	3

Relevant Stars for Representations

The one remaining task is to record which of the linearly independent vectors (or plane-wave spinors in the relativistic problem) should be used with the different, irreducible representations at the four high symmetry points. This is accomplished with an array which gives the members of the expanded first star list to be used for each representation. This array of relevant stars includes enough vectors from each star to form a complete set for a given representation; and, the vectors have been tested to insure that they do not project to zero. Examples of such are shown in Tables VI and VII for the Γ point and L point. The designation of the irreducible representations is that of Bouckaert, Smoluchowski, and Wigner⁷ and is also used by Slater.³

TABLE V. Linearly independent k-vectors at X, L, and W for the double group

Wave Vector	Star	k_x	k_y	k_z	Spin State	Expanded First Star List
X POINT						
1	1	1	-0	-0	α	1
1		1	-0	-0	β	2
3	2	-0	-1	1	α	3
4		0	1	1	β	4
7	3	1	0	2	α	5
8		-1	2	-0	α	6
8		-1	2	-0	β	7
10		1	2	0	α	8
L POINT						
1	1	.5	.5	.5	α	1
2	2	-.5	-.5	-.5	α	2
3	3	1.5	-.5	-.5	α	3
4		-.5	1.5	-.5	α	4
6	4	.5	.5	-1.5	α	5
7		.5	-1.5	.5	α	6
9	5	-.5	1.5	1.5	α	7
10		1.5	-.5	1.5	α	8
12	6	.5	-1.5	-1.5	α	9
13		-1.5	.5	-1.5	α	10
15	7	1.5	1.5	1.5	α	11
16	8	.5	.5	2.5	α	12
17		.5	2.5	.5	α	13
19	9	-.5	-.5	-2.5	α	14
20		-.5	-2.5	-.5	α	15
22	10	-1.5	-1.5	-1.5	α	16
23	11	-1.5	.5	2.5	α	17
24		.5	-1.5	2.5	α	18
25		.5	2.5	-1.5	α	19
25		.5	2.5	-1.5	β	20
W POINT						
1	1	1	.5	-0	α	1
2		0	-.5	1	β	2
5	2	-0	1.5	-1	α	3
6		1	-1.5	-0	β	4

TABLE VI. Relevant stars for single group representations

Γ_1 (1DIM)	Γ_2 (1DIM)	Γ_{12} (2DIM)	Γ_{25} (3DIM)	Γ_{15} (3DIM)	L_1 (1DIM)	L_2 (1DIM)	L_3 (2DIM)
1	20	4	2	6	1	11	3
2	36	6	3	8	2	13	4
3	40	8	5	10	3	21	5
4		10	6	16	4		6
6		14	7	18	5		8
8		16	8	20	6		9
10		18	9	21	7		11
12		20	10	22	8		12
13		23	11		9		13
14			12		10		14
16			13		11		
18			15		13		
20			16		15		
			17		17		
			18		18		
			19		19		
			20		20		
			21		21		
			22		22		

TABLE VII. Relevant stars for double group representations

$\Gamma_6(2\text{DIM})$	$\Gamma_7(2\text{DIM})$	$\Gamma_8(4\text{DIM})$	$L_4(2\text{DIM})$	$L_5(1\text{DIM})$
1	2	2	1	3
2	3	3	2	5
3	4	5	3	7
4	7	6	4	9
7	10	7	5	12
8	11	8	6	14
11	14	9	7	17
12	15	10	8	18
15	18	11	9	21
16	19	12	10	
19	20	13	11	
20	21	14	12	
21	24	15	13	
24	27	16	14	
25	28	17	15	
28	31	18	16	
29	32	19	17	
32	33	20	18	
33	35	22	19	
35	37	23	20	
37	40	24		
40		25		
		26		
		27		
		28		
		29		
		30		
		31		
		32		
		33		
		34		
		35		
		36		
		37		
		38		
		39		
		40		

Summary of Technique

The results of the symmetrization process are utilized in the following way: One considers a particular high symmetry point in the Brillouin zone, and chooses an irreducible representation at that point. The relevant star list indicates which member of the expanded first star list will form a linearly independent set of

TABLE VIII. Matrix sizes and times to obtain energy eigenvalues for GaAs (Γ point)

	Number of Plane Waves	Matrix Size	7094 Time
Unsymmetrized	27	54x54	2 Min. 45 Sec.
	65	130x130	17 Min. 50 Sec.
Symmetrized	27	Γ_6 Γ_7 Γ_8	6x6 5x5 8x8 3 Min. 50 Sec.
	65	Γ_6 Γ_7 Γ_8	13x13 12x12 20x20 6 Min. 40 Sec.
	113	Γ_6 Γ_7 Γ_8	21x21 20x20 36x36 10 Min. 52 Sec.
	229	Γ_6 Γ_7 Γ_8	41x41 40x40 74x74 28 Min. 17 Sec.
	229	Γ_1 Γ_2 Γ_{12} Γ_{25} Γ_{15}	25x25 3x3 21x21 42x42 21x21 2 Min. 15 Sec.
Non-Relativistic			

starting vectors for each star. The use of the resulting symmetrized vectors is equivalent to the use of the entire star in an unsymmetrized treatment. Each designated vector is used in the SCRAMBLE subroutine to create one symmetrized basis function of the irreducible representation.

IV. TIME COMPARISONS

The savings in computer time that can be realized by the use of symmetrized OPW's is demonstrated in Table VIII. The times given are those necessary to set up the matrices and solve a determinant equation of the form of Eq. 2.7. The number of plane waves listed is the total number of k-vectors involved, or the equivalent number for the symmetrized case. The matrix sizes are twice as large due to spin in the relativistic OPW formalism. When only a very few plane waves are used, the time is limited by other factors, e.g., tape handling. However, with 65 plane waves,

there is a significant saving in calculation time when symmetry is used to reduce the secular equations.

The last entry in this table gives the 7094 time for 229 plane waves in a non-relativistic OPW calculation. One pays a very severe penalty in time to go to the relativistic formalism. In fact, without the use of symmetry or some other technique to factor the matrices, the times would be prohibitive.

We have concluded in our work that symmetrization can be included in the OPW method in a systematic and transparent way. The benefits far outweigh the time and effort necessary to set up the procedure.

ACKNOWLEDGEMENT

The authors would like to acknowledge the excellent typing work and editorial assistance of Mrs. J. Ingvaldstad.

REFERENCES

1. L. F. Mattheiss, J. H. Wood, and A. C. Switendick, "A Procedure for Calculating Electronic Energy Bands Using Symmetrized Augmented Plane Waves", in Methods in Computational Physics Vol. 8, edited by B. Alder, S. Fernbach and M. Rotenberg, Academic Press, New York (1968).
2. R. N. Euwema and D. J. Stukel, "OPW Convergence of Some Tetrahedral Semiconductors", Phys. Rev., 132 (1970).
3. J. C. Slater, Quantum Theory of Molecules and Solids: Volume 2, McGraw-Hill Book Co., New York (1965).
4. A. W. Luehrmann, "Crystal Symmetries of Plane-wave-like Functions I. The Symmorphic Space Groups", Advances in Physics, 17, 1 (1968).
5. M. Tinkham, Group Theory and Quantum Mechanics, McGraw-Hill Book Co., New York (1964).
6. F. Herman, R. L. Kortum, C. D. Kuglin and J. P. Van Dyke, "Relativistic Band Structure of GeTe, SnTe, PbTe, PbSe, and PbS", Journal de Physique, 29, C4 (1968).
7. L. P. Bouckaert, R. Smoluchowski and E. P. Wigner, "Theory of Brillouin Zones and Symmetry Properties of Wave Functions in Crystals", Phys. Rev., 50, 58 (1936).

SOME NOTES ON A MODIFIED OPW METHOD

D.M. Gray and R.J. Karpien*

Watervliet Arsenal, Watervliet, New York

A brief description of a modification of the standard OPW method is given. This modification was used by Gray and Brown for Cu₃Au and is essentially the same as that of Deegan and Twoose. A comparison of standard OPW and the modified version as applied to three selected irreducible representations of Cu is discussed in detail. Results of a similar comparison for LaSn₃ are discussed briefly. For the two crystals chosen, standard OPW was found to be reasonably good for LaSn₃ but inadequate for Cu whereas MOPW was quite adequate for both.

INTRODUCTION

It is fairly well known that, in standard form, OPW does not work well for certain crystals,¹ particularly the transition elements where the 3d shell is being filled or Cu with a just-filled 3d shell. Orthogonalizing to the 3d function means orthogonalizing to a function which is not a good crystal eigenfunction whereas ignoring the 3d function generally produces very slow convergence.

To avoid such difficulties, Brown and Krumhansl² developed a method which treats all the atomic-like functions (defined below) allowed by symmetry for a given irreducible representation as expansion functions on the same footing as the plane waves. In this technique, which we refer to as the modified plane wave (MPW)

* Present address: Physics Dept., Rensselaer Polytechnic Institute, Troy, N.Y.

method, there is no orthogonalization. We call particular attention to the application of the MPW method to Cu by Butler et.al.³ as this paper is of particular importance in establishing the MPW method. Their results (for about 75 levels calculated and including points in k space of seven different types of symmetry) agree with Burdick's APW results⁴ generally to within about 0.005 Ry. Both Burdick and Butler et.al. used the Chodorow⁵ potential. Although developed historically in reverse order, one can think of OPW as starting with the MPW set of functions and then orthogonalizing the plane waves to all the atomic-like functions.

THE MOPW METHOD

We now describe a method which combines features of the MPW and OPW methods. The expression "atomic-like function" as used here means those functions, essentially atomic functions, whose angular parts are identical to the angular parts of the corresponding atomic function but whose radial parts have been obtained by re-solving the Schrödinger equation with the added restriction that the function go to zero at the radius of the muffin-tin sphere. The lower lying atomic-like functions will thus be identical to the corresponding atomic functions. The higher lying atomic-like functions will be squeezed inward somewhat with slightly higher energies than in the free atom case. In what we call the modified orthogonalized plane wave (MOPW) method, one starts with the MPW set, proceeds as in OPW but orthogonalizes the symmetrized plane waves only to those lower lying atomic-like functions which one "trusts" as being quite good crystal eigenfunctions as well. The plane waves are not orthogonalized to the higher lying atomic-like functions; these higher lying atomic-like functions are kept in the expansion explicitly as symmetrized Bloch sums on the same footing as the symmetrized OPW's. Our procedure is essentially the same as that of Deegan and Twose⁶ differing mainly in the manner of obtaining the atomic-like functions. A somewhat similar approach was used by Callaway⁷ in a calculation for iron. This idea of supplementing the orthogonalized plane waves with functions picked to represent the appropriate atomic function symmetries had already been suggested by Herring⁸ in his original OPW paper. Brown⁹ has shown that MOPW is mathematically equivalent to MPW if the atomic functions to which one orthogonalizes are exact eigenfunctions for the crystal. A schematic outline of our version of MOPW is given in an appendix.

COMPARISON OF MOPW AND OPW

We now compare MOPW and OPW for some actual cases. For this comparison we take the MPW results as being "correct" since, of the three methods (MPW, MOPW, OPW), MPW has the highest degree of

Table I. X_4' energy levels for Cu. See text for a description of the various methods of calculation. All energies are in Ry. For reference, Burdick's APW level⁴ is at -0.235 Ry.

Symmetrized Plane Waves	Square Mag.	MPW ^a	MOPW	OPW1	OPW2
1	1	-0.2151	-0.2151	-0.2250	
6	13	-0.2335	-0.2336	-0.2428	
12	22	-0.2345	-0.2346	-0.2436	0.2053
18	30	-0.2346	-0.2346		
24	38	-0.2347	-0.2346	-0.2435	-0.0752
38	54			-0.2439	-0.1702

^a The MPW level at 24 plane waves is from Butler et.al.³. The other MPW levels (and the fourth significant figure of the 24 plane wave level) are from Butler.¹⁹ We recalculated the 1 and 12 plane wave levels as a check and agree to four significant figures.

variational flexibility. (Although all the examples discussed here use the muffin-tin model, the main advantage of MOPW over OPW as pointed out here should, in principle, apply to non-muffin-tin models also.) We confine ourselves here to a fairly detailed discussion of three irreducible representations of Cu and to a brief description of similar results for LaSn₃¹⁰.

Copper

The same Chodorow⁵ potential used by Burdick⁴ and by Butler et.al.³ was used for our Cu calculations. For reference, the Cu atomic-like 2p, 3p, and 3d energies are as follows:

$$\begin{array}{ll} \text{Cu 3d} & -0.5 \text{ Ry}, \\ \text{Cu 3p} & -5.4 \text{ Ry}, \\ \text{Cu 2p} & -65.9 \text{ Ry}. \end{array}$$

The Cu Fermi level is at -0.38 Ry.⁴

X_4' , $\vec{k}=(100)^{11}$: This representation¹² allows p symmetry only (thus no Cu valence functions). Table I lists electron energy levels for four choices of calculation. In Table I through III, column 1 gives the number of symmetrized plane waves used in a particular truncation, column 2 lists the square of the magnitude¹¹ of the largest plane wave in that truncation and column 3 gives the MPW energy level. All energies are in Rydbergs. Retention of the fourth decimal place in the energies listed is merely to show comparison. In all three tables, only the levels of interest are shown. In Table I, MOPW orthogonalizes (orthogs) to 2p and "keeps"

$3p$ as an explicit expansion function. OPW1 orthogs to $2p$ and $3p$. OPW2 orthogs to $2p$ and ignores $3p$.

As expected, MOPW is virtually identical to MPW. This, plus the fact that the orthogonalized plane waves do contribute importantly to this level, indicates merely that the Cu $2p$ atomic function is an accurate eigenfunction for the crystal. Orthogonalizing to $3p$ (OPW1) produces a level which is consistently low by about 0.01 Ry. At -5.4 Ry, the atomic-like $3p$ function is probably a quite good eigenfunction for the crystal. Ignoring the $3p$ function (OPW2) produces quite slow convergence. Thus OPW (in the OPW1 scheme) did not do too badly here but this was a p-type representation and the major contributions to the level of interest come from the plane waves anyway.

Γ_{25}' , $k=(000)$: This representation allows d symmetry only. Since the only atomic-like function allowed here is the $3d$ function and since MOPW would keep that function, then, for this case, MPW and MOPW are identical by definition. Table II lists energy levels for three choices of calculation. OPW1 orthogs to $3d$. OPW2 ignores $3d$. As one would expect, orthogonalizing to the $3d$ function (OPW1) produces very poor results and ignoring the $3d$ (OPW2) gives extremely slow convergence.

Σ_4' , $k=(1/4 \ 1/4 \ 0)$: This representation allows p and d symmetry. Table III lists energies for five choices of calculation. MOPW orthogs to $2p$ and keeps $3p$ and $3d$. OPW1 orthogs to $2p$, $3p$ and $3d$. OPW2 orthogs to $2p$ and $3p$ and ignores $3d$. OPW3 orthogs to $2p$

Table II. Γ_{25}' energy levels for ${}^4\text{Cu}$. All energies are in Ry. For reference, Burdick's APW level is at -0.640 Ry.

Symmetrized Plane Waves	Square Mag.	MPW ^a and MOPW	OPW1	OPW2
1	3	-0.6220		
6	19	-0.6339		
12	27	-0.6411	1.575	0.644
24	44	-0.6453	1.574	0.493
38	59	-0.6465	1.572	0.370

^a The MPW level at 24 plane waves is from Butler et.al.³

The 38 plane wave MPW level was calculated by us.

The other MPW levels (and the fourth significant figure of the 24 plane wave level) are from Butler.¹⁹

and ignores $3p$ and $3d$.

Table III shows that orthogonalizing to the $3d$ function (OPW1) is unsuitable and that ignoring the $3d$ (OPW2 or OPW3) produces

Table III. Σ_4 ($1/4 \ 1/4 \ 0$) energy levels for Cu. All energies are in Ry. For reference, Burdick's APW level⁴ is at -0.600 Ry.

Symmetrized Plane Waves	Square Mag.	MPW	MOPW	OPW1	OPW2	OPW3
1	3.125	-0.4936 ^a	-0.4936			
6	9.125	-0.5769	-0.5769		1.188	
12	13.125	-0.5863	-0.5863	1.599	1.029	1.030
18	18.125	-0.5879	-0.5879			
26	21.125	-0.5888		1.592	0.921	0.924
36	27.125			1.591	0.848	0.850

^a This level is from Butler.¹⁹

extremely slow convergence. MOPW, however, gives results which are identical (to 4 figures) to MPW. The 3d function contributes very strongly to the level of interest in Σ_4 . Thus, from Tables II and III, whenever the Cu level desired has strong d character, standard OPW (either orthogonalizing to or ignoring the 3d function) gives quite poor results whereas MOPW is quite adequate.

Part of the OPW error here is due to forcing our higher lying atomic-like functions to go to zero inside the radius of the muffin-tin sphere. This raises their energies somewhat so that orthogonalizing to these functions is particularly bad. However, the energy shift associated with this truncating hardly accounts for the roughly 2 Ry. discrepancy between MPW and OPW in Tables II and III (see OPW1 of both tables). The error might not be as big with a non-muffin-tin model but the principle that the (in this case, 3d) atomic-like function or the atomic function is not a good crystal eigenfunction should still be valid.

LaSn₃

It should be appreciated that for Cu, where we orthogonalize to only one or two functions, MOPW has no big advantage over MPW but that for crystals like LaSn₃ with many atomic-like functions allowed by symmetry (47 for Σ_1 for example), orthogonalizing to all but the higher lying atomic-like functions means a considerable reduction in matrix size.

LaSn₃ has the Cu₃Au structure (simple cubic lattice).¹³ For a representation allowing d symmetry only, the important atomic-like function is the La 5d valence function. Our LaSn₃ potential was constructed from the atomic potentials for La and Sn of Herman and Skillman.¹⁴ Results (not shown) for LaSn₃ for some selected irreducible representations were similar to those for Cu with MOPW

again converging more reliably and more rapidly than OPW although here the discrepancy between OPW and MPW was considerably less than in the Cu case. Ignoring the La 5d function, for example, gave much better results than ignoring the Cu 3d function. This is very likely due to the presence of the lower lying La and Sn 3d and 4d functions in the LaSn_3 case.

SUMMARY

For LaSn_3 (with many lower lying atomic functions) and for some representations of Cu (e.g., X_4' with p symmetry only, Table I) standard OPW works fairly well if a judicious choice is made of whether to orthogonalize to or ignore a given function. For MOPW, no such choice need be made; one simply keeps all the higher lying and marginal atomic-like functions as explicit expansion functions at the expense of a slightly larger matrix. For those Cu points which allow d symmetry, standard OPW appears to be inadequate whereas MOPW is quite adequate (Tables II and III).

It could be argued that in going from OPW to MOPW, two additional types of integrals are introduced into the matrix formulation, namely atomic-atomic and atomic-plane wave integrals. However, once programmed, this is of minor importance and it seems to us that these integrals are needed in forming the orthogonalized plane waves anyway.

ACKNOWLEDGMENTS

The authors are particularly indebted to Dr. F. Kenneth Bloom, Jr. for the program used to calculate the muffin-tin-restricted atomic-like functions and also for the program to calculate the Cu energy levels. We thank Dr. Frank A. Butler for providing us with his unpublished MPW levels for various numbers of plane waves. Most of our computations were performed on an IBM 360-65 computer at Picatinny Arsenal while some were done on an IBM 360-44 at Watervliet Arsenal. We wish to thank the staffs of both facilities for their cooperation, in particular Mrs. M.J. Webster at Picatinny. We are pleased to acknowledge many fruitful discussions with Drs. Alma Gray and L.V. Meisel and Professor E. Brown.

APPENDIX

In our version of MOPW we approximate the true crystal wave function for the n^{th} energy level by

$$\psi_n(\vec{r}) = \sum_i c_{ni} \phi_i(\vec{r}),$$

where the ϕ_i are of two types: (1) higher lying atomic-like functions or cutoff functions (CO), (2) orthogonalized plane waves (OPW's). The major difference between our method and that of Deegan and Twose⁶ occurs in forming the CO functions. To obtain the radial part of their CO functions, they first integrate the radial Schrödinger equation outwards from $r=0$ for the correct ℓ value and some appropriate energy using the atomic potential. This function is then cut off at some $r=r_m$ and a tail is joined smoothly to it, the tail being chosen so as to vanish at one half the nearest neighbor distance. (See Appendix C of ref. 6.) In our method, the radial Schrödinger equation is solved for the correct ℓ value using the atomic potential. This is an iterative procedure with the energy being changed for each iteration until the function goes to zero at the radius of the muffin-tin sphere. For the functions so constructed which do not have zero slope at the muffin-tin radius, r_{mt} , a cubic function was fitted in the region near r_{mt} . We emphasize that this cubic fitting is quite minor and is applied to functions which already have zero value at r_{mt} .

All the ϕ_i are symmetrized functions and we use only those ϕ_i allowed by symmetry for a given irreducible representation. An unsymmetrized plane wave has the form $\exp[i(\vec{k}+\vec{K}_i)\cdot\vec{r}]$ which symmetrizes to $\sum_R D_{11}(R) \exp[i(\vec{k}+\vec{K}_i)\cdot R^{-1}\vec{r}]$ with \vec{r} ranging over the primitive cell. The D 's are the matrices for the representation in question and R is any element of the group associated with the symmetry point k . For Cu, an unsymmetrized atomic-like function has the form $[u_{nl}(r)/r] X_{lq}(\theta, \phi)$ where $u_{nl}(r)/r$ is the ordinary radial part (with u determined as discussed above) and the X_{lq} are cubic harmonics. For LaSn₃, one must define the position of an atom within the primitive cell and in symmetrizing one must be careful with operations which take a function from one site to another. (For details of symmetrizing with more than one atom per primitive cell see ref. 15.)

For simplicity, we now suppress the symmetrization notation and also consider only the central atom of any multi-atom basis. The following discussion is thus schematic only. The variational procedure leads to the matrix equation

$$\hat{H}\vec{c}_n = E_n \vec{s}\vec{c}_n \quad (1)$$

$$\text{with } H_{ij} = \int \phi_i^*(\vec{r}) \hat{H} \phi_j(\vec{r}) d\tau \quad (2)$$

$$S_{ij} = \int \phi_i^*(\vec{r}) \phi_j(\vec{r}) d\tau \quad (3)$$

Schematically, the matrices are of the form

$$H = \begin{pmatrix} (CO|H|CO) & (CO|H|OPW) \\ (OPW|H|CO) & (OPW|H|OPW) \end{pmatrix} \quad (4)$$

We define an orthogonalized plane wave as

$$\phi(\vec{k}_j) = PW(\vec{k}_j) - \sum_t a_{jt} c_t(\vec{r}) \quad (5)$$

with $PW(\vec{k}_j)$ being the ordinary (symmetrized) plane wave term and the $c_t(\vec{r})$ being the lower (symmetrized) atomic-like functions. We write \vec{k}_j for $\vec{k} + \vec{k}_j$. If the c_t satisfy

$$\langle c_b | c_t \rangle = \delta_{bt} \quad (6)$$

$$\text{then } a_{jt} = \langle c_t | PW(\vec{k}_j) \rangle \quad (7)$$

Instead of numerically evaluating all the $\phi(\vec{k}_i)$ above and integrating these new functions, we first express the OPW-OPW elements in terms of their constituent parts, i.e.,

$$\langle \phi(\vec{k}_i) | H | \phi(\vec{k}_j) \rangle = \langle PW(\vec{k}_i) | H | PW(\vec{k}_j) \rangle - A - B + D \quad (8)$$

$$\text{with } A = \sum_b a_{ib}^* \langle c_b | H | PW(\vec{k}_j) \rangle$$

$$B = \sum_t a_{jt} \langle PW(\vec{k}_i) | H | c_t \rangle$$

$$D = \sum_{b,t} a_{ib}^* a_{jt} \langle c_b | H | c_t \rangle$$

If each c_t satisfies

$$H c_t(\vec{r}) = E_t c_t(\vec{r}) \quad (9)$$

Then Eq. (8) reduces to

$$\langle \phi(\vec{k}_i) | H | \phi(\vec{k}_j) \rangle = \langle PW(\vec{k}_i) | H | PW(\vec{k}_j) \rangle - \sum_t a_{it}^* a_{jt} E_t \quad (10)$$

In our application, some of our "lower" atomic-like functions do not satisfy (9) exactly (due to the slight cubic fitting discussed above). We therefore used Eq. (8) and not Eq. (10). Calculation of the actual values indicated that taking $\langle c_b | H | c_t \rangle$ equal to zero for $b \neq t$ was still a good approximation. One then sums, as indicated in Eq. (8), over the MPW type integrals (for details of the evaluation of these integrals see refs. 15 and 16).

Similarly,

$$\langle \phi(\vec{k}_i) | \phi(\vec{k}_j) \rangle = \langle PW(\vec{k}_i) | PW(\vec{k}_j) \rangle - \sum_t a_{it}^* a_{jt} \quad (11)$$

$$\text{and } \langle CO_i | H | \phi(\vec{k}_j) \rangle = \langle CO_i | H | PW(\vec{k}_j) \rangle - \sum_t a_{jt} \langle CO_i | H | c_t \rangle \quad (12)$$

The CO functions could be explicitly orthogonalized to the c_t as was done by Deegan and Twose⁶. Our method of constructing the CO makes $\langle CO_i | H | c_t \rangle$ automatically equal to zero for $\ell(i) \neq \ell(t)$ and

actual calculation indicated that these terms are negligible for $n(i) \neq n(t)$ even when $\ell(i) = \ell(t)$.

$$\langle c_{0i} | \emptyset(\vec{k}_j) \rangle = \langle c_{0i} | P_W(\vec{k}_j) \rangle - \sum_t a_{jt} \langle c_{0i} | c_t \rangle \quad (13)$$

As in Eq. (12), we took the \sum_t above to be zero.

$$\langle c_{0i} | c_{0j} \rangle = M_i \int u_i(r) u_j(r) dr \Delta_{ij} \quad (14)$$

$$\langle c_{0i} | H | c_{0j} \rangle = M_i \int u_i(r) H_r u_j(r) dr \Delta_{ij} \quad (15)$$

$$\text{with } M_i = \int x_i^2(\theta, \phi) d\Omega, \quad (16)$$

$$\Delta_{ij} = \delta_{\ell(i), \ell(j)} \delta_{q(i), q(j)} \quad (17)$$

$$\text{and } H_r = -d^2/dr^2 + \ell(\ell+1)/r^2 + V(r) \quad (18)$$

Once the S and H matrix elements have been formed a unitary transformation which diagonalizes S is applied to Eq. (1) and then the transformed H matrix is diagonalized giving the eigenvalues and eigenfunctions. This diagonalization process is the major time consumer in the computer program. The process can undoubtedly be speeded up by writing $S = LL^\dagger$ where L is "lower triangular" thus avoiding diagonalizing $S^{1/2}$ and by replacing our present diagonalization routine (Jacobi method) with the QR method.¹⁸

References

1. John C. Slater, Quantum Theory of Molecules and Solids - Vol. 2, Symmetry and Energy Bands in Crystals (McGraw-Hill, Inc., New York, 1965), In particular, Sec. 8-4
2. E. Brown and J.A. Krumhansl, Phys. Rev. 109, 30 (1958).
3. F.A. Butler, F.K. Bloom, Jr., and E. Brown, Phys. Rev. 180, 744 (1969).
4. G.A. Burdick, Phys. Rev. 129, 138 (1963).
5. M.I. Chodorow, Ph.D. Thesis, MIT (1939), (unpublished).
6. R.A. Deegan and W.D. Tows, Phys. Rev. 164, 993 (1967).
7. J. Callaway, Phys. Rev. 99, 500 (1955). See also J. Callaway, Phys. Rev. 97, 933 (1955).
8. C. Herring, Phys. Rev. 57, 1169 (1940).
9. E. Brown, Phys. Rev. 126, 421 (1962).
10. LaSn₃ was picked for convenience: One of us (D.M.G.) is presently calculating the band structure of this crystal.
11. The factor $2\pi/a$ is omitted here. For Cu, a is taken as 6.83113 Bohr radii.
12. The symbols for the irreducible representations are those of L.P. Bouckaert, R. Smoluchowski, and E. Wigner, Phys. Rev. 50, 58 (1936).

13. R.J. Gambino, N.R. Stemple, and A.M. Toxen, *J. Phys. Chem. Solids* 29, 295 (1968).
14. F. Herman and S. Skillman, Atomic Structure Calculations (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1963).
15. D.M. Gray, Watervliet Arsenal Technical Report, WVT-7005 (1970).
16. D. Gray and E. Brown, *Phys. Rev.* 160, 567 (1967), Appendix.
17. D.G. Shankland, private communication.
18. See the paper in this report by R.A. Faulkner.
19. Frank A. Butler (private communication).

2. KKR METHODS IN THE BAND PROBLEM

RECENT DEVELOPMENTS IN KKR THEORY

A. R. Williams, S. M. Hu^{*}, and D. W. Jepsen

IBM Thomas J. Watson Research Center

Yorktown Heights, New York

ABSTRACT

The KKR or Green's function method has been generalized and sophisticated by many workers. We present here three steps in the method's continuing evolution. First, we generalize the method to accommodate greater deviations of the potential from muffin-tin form. Second, by clarifying the relation between KKR and the cellular method, we arrive at a new prescription for the wave function. Third, we describe our efforts (not yet completely successful) to qualitatively decrease the time required for KKR calculations by eliminating the necessity for determinant evaluation. Finally, we offer our opinion concerning the relative merits of various methods discussed at the conference.

INTRODUCTION

We wish to discuss three different aspects of band calculations using the KKR method. In Part I the application of KKR to loosely packed structures where the muffin-tin approximation is not realistic is discussed. A generalization of the inelastic Green's function formalism¹ which accommodates greater anisotropy in the interstitial crystal potential is presented along with numerical results for Ge and Si. In Part II a prescription for obtaining the wave function in KKR calculations is given. It is shown that the same expansion commonly used by the KKR and APW methods inside the muffin-tin sphere can be used throughout the

* IBM Components Division, East Fishkill, New York

atomic cell and that there are several advantages in doing this. In Part III we consider the problem of using the KKR method to find energies and wave functions on a high density mesh of points in \vec{k} -space. The energy and wave function are predicted by extrapolation from neighboring mesh points and then corrected by the secant method² as applied to both the energy and the wave function. While this scheme functions accurately most of the time, we have not as yet been able to make it sufficiently reliable for completely automated use.

I. KKR IN LOOSELY PACKED STRUCTURES

The KKR method is based on finding the combinations of \vec{k} and p^2 for which the functional Λ ,

$$\Lambda = \langle \Psi_{p\vec{k}} | V - V G_p^o V | \Psi_{p\vec{k}} \rangle \quad 1.1$$

is stationary. In equation 1.1 p^2 is the energy, \vec{k} the Bloch propagation vector, and V is the total crystal potential,

$$V(\vec{r}) = \sum_{\vec{R}} v(\vec{r} - \vec{R}) \quad 1.2$$

and $v(\vec{r})$ is the potential describing the interaction of the particle and any single ion. Using the Bloch symmetry of the wave function,

$$\Psi_{p\vec{k}}(\vec{r} + \vec{R}) = e^{+i\vec{k}\cdot\vec{R}} \Psi_{p\vec{k}}(\vec{r}) \quad 1.3$$

we can rewrite 1.1 as follows:⁽¹⁾

$$\langle \Psi_{p\vec{k}} | V - V G_p^o V - V G^E(p, \vec{k}) V - V G^I(p, \vec{k}) V | \Psi_{p\vec{k}} \rangle = \Lambda \quad 1.4$$

$G^E(E, \vec{k})$ is the elastic Green's function and describes propagation between scattering events with kinetic energy p^2 . $G^I(p, \vec{k})$ is the inelastic Green's function and describes interstitial propagation at kinetic energies other than p^2 . The significance of the muffin-tin approximation is that it ensures that all interstitial

propagation takes place with the same energy thereby making the G^I term in 1.4 unnecessary.

The variation of 1.4 is accomplished by expressing the wave function as a linear combination of trial orbitals,

$$\Psi_{\rho \vec{k}}(\vec{r}) = \sum_L \beta_L(\vec{k}) Y_L(\vec{r}) R_{\ell \rho}(r) \quad 1.5$$

and then varying the expansion coefficients $\{\beta_L(\vec{k})\}$ to obtain the following determinantal relation between p^2 and \vec{k} :

$$\text{Det} \left\{ \langle R_L | v - v G_\rho^o v - v G^E(p, \vec{k}) v - v G^I(p, \vec{k}) v | R_{L'} \rangle \right\} = 0 \quad 1.6$$

The index L is a composite for the angular momentum quantum numbers ℓ and m and $R_{\ell \rho}(r)$ is the solution of Schrodinger equation corresponding to the potential v and energy E . Using the fact that matrix elements of $G^E(p, \vec{k})$ in a representation consisting of spherical Bessel functions and Harmonics,

$$\langle \vec{r} | L \rho \rangle = Y_L(\hat{r}) j_\ell(r) \quad 1.7$$

are the KKR structure constants,

$$G^E(p, \vec{k}) = \sum_{LL'} | L \rho \rangle B_{LL'}(p, \vec{k}) \langle \rho L' | \quad 1.8$$

we may rewrite 1.6 as follows

$$\text{Det} \left\{ \frac{\langle R_L | v - v G_p^0 v | R_L \rangle}{\langle R_L | v | L_p \rangle \langle p L | v | R_L \rangle} \delta_{LL'} - B_{LL'}(p, \vec{k}) \right. \\ \left. - \frac{\langle R_L | v G^I(p, \vec{k}) v | R_{L'} \rangle}{\langle R_L | v | L_p \rangle \langle p L' | v | R_{L'} \rangle} \right\} = 0 \quad 1.9$$

In the muffin-tin approximation the final term in 1.9 can, as we have said, be dropped and the first term can be identified as the Lippman-Schwinger expression for $-p \cot(\delta_\ell(E))$,³ where $\delta_\ell(E)$ is the usual scattering phase shift. Another virtue of the muffin-tin approximation is the fact that the only approximation implied by 1.5 is the truncation of the L - sum. When the muffin-tin approximation is removed, as it must be in loosely packed structures, two complications arise. First we must include the G^I term in the secular equation; since the latter does not involve much computer time, this complication is not particularly serious. The second effect of ionic overlap is that the individual trial functions in the expansion 1.5 are no longer solutions of the crystal Schrodinger equation. This does not mean that the expansion 1.5 is worthless; we can still find the extrema of Λ in the space corresponding to 1.5. The resulting energies will, however, suffer from the lack of completeness of the basis set. Energies for Si were computed using the full Λ and the expansion 1.5 with high success.¹ When, however, the same technique was applied to Ge, we found that the inadequacies of the basis set produced greater errors (2 to 4 tenths of an ev.). To cope with this problem, we generalized the basis set 1.5 to include two radial functions for each L :

$$\Psi_{p\vec{k}}(\vec{r}) = \sum_L Y_L(\hat{r}) \left\{ \beta_L R_{\ell p}(r) + \beta'_L R'_{\ell p}(r) \right\} \quad 1.10$$

where $R'_{\ell p}(r)$ is the energy derivative of $R_{\ell p}(r)$. The rational

behind 1.10 is revealed by considering the complete generalization of 1.5:

$$\Psi_{pk}(\vec{r}) = \sum_L Y_L(\hat{r}) \int d\vec{q} \beta_{Lq}(\vec{k}) R_{Lq}(r) \quad 1.11$$

Our truncated basis set 1.10 is therefore spanning that portion of the complete set 1.11 lying near p in q -space. This Hilbert space argument is supported by the more intuitive observation that $R_{kp}(r)$ is the optimum generalization of 1.10 because it is essentially zero in the core region ($v(r)$ dominates p^2 in the radial Schrodinger equation for small r , thereby making $R_{kp}(r)$ almost energy independent) and is completely out of phase with $R_{kp}(r)$ outside the core where $R_{kp}(r)$ oscillates.

The doubling of our basis set, of course, doubles the size of our secular equation and therefore substantially increases computation time.

For a general \vec{k} -point and the zinc-blende structure the secular matrix is 100×100 . In comparison with KKR calculations for metals, which usually employ a 9×9 secular matrix, the dimension of 100 appears enormous. The dimension of 100 does however compare favorably with the size of OPW secular equations used to study the same materials, for the latter are often of dimension greater than 400.⁴ A summary of our calculated results appears in Table I, and a comparison of our results with those of pseudopotential and OPW calculations appears in Table II.

II. THE KKR WAVE FUNCTION

Almost all quantitative comparisons of band calculations with experiment involve not only the single particle energies but the corresponding wave functions as well. As the computation of energies has become more routine, interest has shifted increasingly to the wave functions. A pertinent example of this is our continuing interest in the interpretation of photoemission data. Our KKR calculations of single particle energies throughout the Brillouin zone have permitted us to predict/analyze the location in energy of structure in photoemission data; however, a detailed examination of the adequacy of one-electron theory in this context is impossible without inclusion of the momentum matrix elements whose calculation requires the one electron wave functions.

TABLE I

Calculated and Measured Band Parameters for Si and Ge

Band Param.	Si		Ge	
	CALC.	MEAS. ^a	CALC.	MEAS. ^a
$\Delta_1^m - \Gamma_{25}$,	1.1	1.1		
$\Gamma_2 - \Gamma_{25}$,	3.7	$3.7 \pm .2$	0.9 ^c	0.9
$\Gamma_{15} - \Gamma_{25}$,	3.1	$2.9 \pm .2$	3.0	3.0
$x_1 - x_4$	3.9	$3.75 \pm .2$	4.2	4.1
$L_1 - L_3$,	3.0	$3.2 \pm .1$	2.2	2.1
$L_3 - L_3$,	4.9	$5.3 \pm .4$	5.5	5.4
$m_{//}^d$	0.97	0.97	1.6	1.6
m_{\perp}^d	0.22	$0.19 \pm .1$	0.1	0.08
A ^d	-3.79	-4.28	-9.2	-13.0
B ^d	-0.81	-0.75	-5.3	-8.9
C ^d	3.6	$3.3 \pm .5$	6.2	10.3

a - A fine compilation of experimental data, its source and its interpretation in terms of the band parameters is given by G. Dresselhaus and M. S. Dresselhaus.⁵

b - Δ_1^m is the lowest energy in the conduction band.

c - The single parameter in the screened ion model⁽¹⁾ is fit to this measurement.

d - $m_{//}$ and m_{\perp} are the longitudinal and transverse effective masses of the conduction band. A, B, and C are the effective mass constants of the valence band.⁽⁵⁾

TABLE II

Comparison of KKR, OPW and Pseudopotential Results for Silicon

Band Param.	KKR ^a	OPW ^b	PSEUDO. ^c	EXPER. ^d
$\Delta_1^m - \Gamma_{25}^m$	1.1	1.1	0.8	1.1
$\Gamma_2^m - \Gamma_{25}^m$	3.7	3.8	3.8	$3.7 \pm .2$
$\Gamma_{15}^m - \Gamma_{25}^m$	3.1	2.8	3.4	$2.9 \pm .2$
$X_1^m - X_4^m$	3.9	4.05	4.0	$3.75 \pm .2$
$L_1^m - L_3^m$	3.0	3.2	3.1	$3.2 \pm .1$
$L_3^m - L_3^m$	4.9	5.0	5.2	$5.3 \pm .4$

a - Present work (using two radial functions per L). One adjustable parameter fit to experiment.

b - Herman, reference 6. Two adjustable parameters fit to experiment.

c - Cohen and Bergstresser, reference 7. Three adjustable parameters fit to experiment.

One electron wave functions are almost always expanded in one of four basic ways. One of the oldest but nonetheless most rapidly converging expansions is the LCAO. This expansion set consists of Bloch symmetrized atomic-like orbitals:

$$\Psi_{p\vec{k}}(\vec{r}) = \sum_{\mu} C_{\mu}(\vec{k}) \sum_{\vec{R}} e^{+i\vec{k}\cdot\vec{R}} \phi_{\mu}(\vec{r}-\vec{R})$$

where the $\{C_{\mu}(\vec{k})\}$ are the expansion coefficients found by minimizing the energy and $\phi_{\mu}(\vec{r})$ is the atomic-like orbital. A second expansion set consists of atomic-like orbitals and plane waves which may or may not be orthogonalized to them. A third set consists of a set of exact solutions of the Schrodinger equation in the core region matched to a set of plane waves in the interstitial region, so-called APW's. The fourth expansion set consists

of the same functions used by the APW in the core region; the difference being that these functions are used throughout the unit cell; we shall call this set the cellular expansion set.

There is a fundamental difference between the first three sets mentioned: the LCAO's, OPW's, and APW's and the cellular set. None of the first three are individually solutions to the Schrodinger equation in question but all three possess precise Bloch symmetry with the implication that they are all intrinsically \vec{k} -dependent. Just the opposite is true of the cellular set; they are individually solutions of the Schrodinger equation but lack Bloch periodicity. There are two important advantages of the cellular set which we mention here because they provide motivation for what is to follow. First, the cellular set is intrinsically two-dimensional reflecting the fact that each member of the set has the correct energy. Second, since the set does not depend on \vec{k} , matrix elements in this basis depend only on energy and can therefore be efficiently tabulated. This is an important advantage for calculations involving integration over \vec{k} and possible \vec{k}' of matrix elements. Using the cellular set, the matrix elements are simply \vec{k} -dependent combinations of tabulated quantities.

It will be noted that with each of the four expansion sets there is an associated method of band calculation resulting from a straightforward minimization of the Hamiltonian in each set. It will be further noted that these four methods do not include the KKR method. This results from the fact that KKR finds stationary values not of the Hamiltonian but of another functional (the inverse T-matrix for the entire crystal). The KKR method does not use the wave function in the same direct way that other methods do. In what follows we attempt to show that this fact provides a third and even greater motivation for use of the cellular set.

The KKR method requires an expansion of the wave function only in the region of space where the ionic potential is non-zero. We will here assume the ionic potential to be of muffin-tin form; the generalization to arbitrary potentials is analytically straightforward but involves considerable programming complexity. Inside the muffin-tin sphere the cellular expansion is used:

$$\Psi_{P\vec{k}}(\vec{r}) = \sum_L \beta_L(\vec{k}) Y_L(\hat{\vec{r}}) R_{\ell m}(\vec{r}) \quad 2.1$$

where L is a composite specifying both angular momentum quantum numbers ℓ and m , $\beta_L(\vec{k})$ is the expansion coefficient determined by the variational procedure, and $R_{\ell m}(\vec{r})$ is the solution of the radial

Schrodinger equation corresponding to energy p^2 :

$$R_{\ell p}(r) = j_\ell(pr) \left\{ 1 - p \int_r^{R_I} r_i^2 dr_i h_\ell^+(pr_i) v(r_i) R_{\ell p}(r_i) \right\}$$

$$-p h_\ell^+(r) \int_0^r r_i^2 dr_i j_\ell(pr_i) v(r_i) R_{\ell p}(r_i). \quad 2.2$$

The notation used here is that of Messiah⁸ and R_I is the radius of the muffin-tin sphere. We use the integral form of the radial Schrodinger equation in preference to the corresponding differential equation because it specifies the arbitrary normalization of $R_{\ell p}(r)$ in a particularly convenient way. Note, for example, that for $r > R_I$

$$R_{\ell p}(r) = j_\ell(pr) - p \langle p\ell | v | R_{\ell p} \rangle h_\ell^+(pr) \quad 2.3$$

where

$$\langle p\ell | v | R_{\ell p} \rangle = \int_0^{R_I} r^2 dr j_\ell(pr) v(r) R_{\ell p}(r). \quad 2.4$$

In order to elucidate the relationship between the KKR method and the cellular functions, we first review the KKR energy determination procedure with special attention paid to the role played by the wave function expansion 2.1. We begin with the Schrodinger equation in integral form.

$$\Psi_{p\vec{k}}(\vec{r}) = \int d^3 r_i G_p^*(\vec{r} - \vec{r}_i) \sum_{\vec{R}} v(\vec{r}_i - \vec{R}) \Psi_{p\vec{k}}(\vec{r}_i) \quad 2.5$$

Using the Bloch symmetry of the wave function:

$$\Psi_{\rho \vec{k}}(\vec{r} + \vec{R}) = e^{+i\vec{k} \cdot \vec{R}} \Psi_{\rho \vec{k}}(\vec{r}) \quad 2.6$$

we rewrite 2.5 as follows:

$$\begin{aligned} \Psi_{\rho \vec{k}}(\vec{r}) &= \int d^3 r_i G_p^\circ(\vec{r} - \vec{r}_i) V(\vec{r}_i) \Psi_{\rho \vec{k}}(\vec{r}_i) \\ &= \int d^3 r_i \sum_{\vec{R} \neq 0} e^{+i\vec{k} \cdot \vec{R}} G_p^\circ(\vec{r} - \vec{r}_i - \vec{R}) V(\vec{r}_i) \Psi_{\rho \vec{k}}(\vec{r}_i) \end{aligned} \quad 2.7$$

Using the spherical wave decomposition of the free particle Green's function:

$$\begin{aligned} G_p^\circ(\vec{r} - \vec{r}_i) &= \frac{-1}{4\pi} \frac{e^{+i\rho|\vec{r} - \vec{r}_i|}}{|\vec{r} - \vec{r}_i|} \\ &= -\rho \sum_L j_\ell(\rho r_i) h_\ell^+(\rho r_i) Y_L(\hat{\vec{r}}) Y_L(\hat{\vec{r}}) \end{aligned} \quad 2.8$$

where $r_>$ is the greater of r and r' and the spherical harmonics are the real variety⁹ we can, for $r < R_I$, rewrite 2.7 as follows:

$$\begin{aligned} \sum_L Y_L(\hat{\vec{r}}) j_\ell(\rho r) \beta_L(\vec{k}) &= \\ -\rho \int d^3 r' \sum_{L''} j_{\ell''}(\rho |r - r'|) Y_{L''}(\vec{r} - \vec{r}') \sum_{\vec{R} \neq 0} e^{+i\vec{k} \cdot \vec{R}} Y_{L''}(\hat{\vec{R}}) h_{\ell''}^+(\rho R) \Psi_{\rho \vec{k}}(\end{aligned} \quad 2.9$$

where we have substituted for $\psi_{\vec{r}}(r)$ in the left-hand side and used 2.2 to simplify. Using the first of the two expansion theorems (AI.3) derived in Appendix I to simplify the right-hand side, we obtain:

$$\beta_L(\vec{k}) = \sum_{L'} B_{LL'}(\rho, \vec{k}) \langle \rho \ell' | v | R_{\ell' p} \rangle \beta_{L'}(\vec{k}) \quad 2.10$$

where

$$B_{LL'}(\rho, \vec{k}) = -4\pi\rho \sum_{L''} C_{LL'L''} i^{L-L'-L''} \sum_{R \neq 0} e^{+i\vec{k} \cdot \vec{R}} Y_{L''}(\vec{R}) h_{\ell''}^+(\rho R) \quad 2.11$$

In passing from 2.9 to 2.10 the summation over L was eliminated by multiplying 2.10 by a spherical harmonic and integrating over angles. For the normalization of $R_{\ell p}(r)$ implied by 2.2 $\rho \ell | v | R_{\ell p}$ is related to the scattering phase shifts η_ℓ as follows:

$$\langle \rho \ell | v | R_{\ell p} \rangle = -\rho e^{+i\eta_\ell} \sin(\eta_\ell) \quad 2.12$$

If we make the following definition:

$$\tilde{B}_{LL'}(\rho, \vec{k}) = B_{LL'}(\rho, \vec{k}) + i\rho S_{LL'} \quad 2.13$$

then we arrive at the usual KKR equation⁹

$$\text{Det} \left\{ -\frac{1}{p} \cot(\eta_\ell) S_{LL'} - \tilde{\beta}_{LL'}(p, \vec{k}) \right\} = 0 \quad 2.14$$

where we recognize the $\tilde{\beta}_{LL'}$'s as the KKR structure constants.

We have said nothing as yet about the convergence of the expansion 2.1 or the size of the KKR secular equation 2.14. The convergence of 2.1 with increasing ℓ is similar to corresponding expansion of a plane wave of the same dimensionless argument and is the convergence with which the cellular method must contend. The convergence of the KKR equation (2.14) is quite another story, for we see that convergence of the summation over L' in 2.10 is governed not by $\beta_L(k)$ but by $\langle p\ell | v | R_{\ell p} \rangle \beta_L(k)$. As we can see from 2.12 $\langle p\ell | v | R_{\ell p} \rangle$ itself converges as $\sin(\eta_\ell)$ with increasing ℓ , so that the product $\langle p\ell | v | R_{\ell p} \rangle \beta_L(k)$ converges more rapidly than $\beta_L(k)$ alone. This is the fundamental advantage of KKR over the cellular method. A somewhat more intuitive interpretation of $\langle p\ell | v | R_{\ell p} \rangle \beta_L(k)$ is provided by 2.3, for we see that it is just the coefficient of the scattered wave outside the potential; this is the basis of the statement that KKR requires only an expansion of the scattered wave.

We have thus far shown that KKR requires only a minimal knowledge of the wave function in order to determine the energy; we now turn our attention to the complimentary part of the problem, namely, the determination of the wave function given the energy. For $r < R_I$ the cellular expansion of the wave function is given directly by 2.10 in terms of the non-negligible $\langle p\ell | y | R_{\ell p} \rangle \beta_L(k)$ produced by the energy determination. (The $B_{LL'}(p, k)$ matrix is not square; 2.14 uses only a square subblock of it.)

Let us now consider the wave function outside the muffin-tin sphere. For $r > R_I$ we can rewrite 2.7 as follows:

$$\begin{aligned} \Psi_{pk}^+(\vec{r}) = & -p \sum_L \beta_L(\vec{k}) \left\{ Y_L(\hat{r}) h_\ell^+ (pr) \langle p\ell | v | R_{\ell p} \rangle + \right. \\ & \left. \sum_{\vec{R} \neq 0} e^{+i\vec{k} \cdot \vec{R}} Y_L(\vec{r} - \vec{R}) h_\ell^+ (p|\vec{r} - \vec{R}|) \langle p\ell | v | R_{\ell p} \rangle \right\} \end{aligned} \quad 2.15$$

where we have substituted the cellular expansion (2.1) for $\psi_{\rho k}^+(r)$ inside the muffin-tin and performed the integration over r' . Using the second expansion theorem (AI.8) derived in Appendix I we rewrite 2.15 as follows:

$$\begin{aligned}\Psi_{\rho k}^+(r) &= \sum_L \beta_L(\vec{k}) Y_L(\hat{r}) \langle \rho \ell | \nabla | R_{\ell p} \rangle h_\ell^+(\rho r) \\ &+ \sum_L Y_L(\hat{r}) j_\ell(\rho r) \sum_{L'} \hat{B}_{LL'}(\rho, \vec{k}) \langle \rho \ell' | \nabla | R_{\ell' p} \rangle \beta_{L'}(\vec{k})\end{aligned}\quad 2.16$$

which can be further simplified by substituting 2.10:

$$\Psi_{\rho k}^+(\vec{r}) = \sum_L \beta_L(\vec{k}) Y_L(\hat{r}) \left\{ j_\ell(\rho r) - \rho \langle \rho \ell | \nabla | R_{\ell p} \rangle h_\ell^+(\rho r) \right\} \quad 2.17$$

Since $r > R_I$, the bracketed quantity is just $R_{\ell p}(r)$ (2.3). What we have shown, therefore, is that the cellular expansion 2.1 holds not only inside the muffin-tin sphere but throughout the atomic cell. The entire wave function is therefore given by 2.10 once the energy is found.

In summary then, we see that a KKR energy determination requires and generates the first few coefficients in the cellular expansion 2.1. Knowing the energy then permits the evaluation of the remaining coefficients in the expansion using equation 2.10. It is important to note that any number of coefficients can be determined without complicating the energy calculation in any way. Furthermore, the wave function expansion arrived at in this way has the important advantages of being two-dimensional and a \vec{k} -dependent linear combination of \vec{k} -independent functions.

Appendix AI

The purpose of this appendix is to establish two expansion theorems required for our study of the KKR wave function in Section II. The first of the two is well known and follows

directly from the following plane wave decomposition:⁸

$$e^{+i\hat{p} \cdot (\vec{r} - \vec{r}')} = 4\pi \sum_{L''} i^{\ell''} Y_{L''}(\hat{p}) j_{\ell''}(\rho |\vec{r} - \vec{r}'|) Y_{L''}(\vec{r} - \vec{r}') \text{AI.1}$$

$$= (4\pi)^2 \sum_{LL'} i^{\ell-\ell'} Y_L(\hat{p}) Y_{L'}(\hat{p}) j_{\ell}(\rho r) j_{\ell'}(\rho r') Y_L(\hat{r}) Y_{L'}(\hat{r}') \text{AI.2}$$

Integrating over \hat{p} we obtain

$$j_{\ell''}(\rho |\vec{r} - \vec{r}'|) Y_{L''}(\vec{r} - \vec{r}') = 4\pi \sum_{LL'} C_{LL'L''} i^{\ell-\ell'-\ell''} j_{\ell}(\rho r) j_{\ell'}(\rho r') \text{AI.3} \\ Y_L(\hat{r}) Y_{L'}(\hat{r}')$$

where

$$C_{LL'L''} \equiv \int d\hat{p} Y_L(\hat{p}) Y_{L'}(\hat{p}) Y_{L''}(\hat{p}). \text{AI.4}$$

The second expansion follows from a similar decomposition of $h_o^+(\rho |\vec{r} - \vec{R} - \vec{\epsilon}|)$ where $R > |\vec{r} - \vec{\epsilon}|$ and $\epsilon < |\vec{R} - \vec{r}|$

$$e^{+i\rho |\vec{r} - \vec{R} - \vec{\epsilon}|} = 4\pi \sum_{L'} j_{\ell'}(\rho |\vec{r} - \vec{\epsilon}|) h_{\ell'}^+(\rho R) Y_{L'}(\vec{r} - \vec{\epsilon}) Y_{L'}(\hat{R}) \text{AI.5}$$

$$= 4\pi \sum_{L''} j_{\ell''}(\rho \epsilon) Y_{L''}(\hat{\epsilon}) h_{\ell''}^+(\rho |\vec{R} - \vec{r}|) Y_{L''}(\vec{R} - \vec{r}) \text{AI.6}$$

Using AI.3 to substitute for $j\ell'(p|\vec{r} - \vec{\varepsilon}|) Y_L(\vec{r} - \vec{\varepsilon})$ in AI.5 we obtain

$$\sum_{L''} j\ell''(p\varepsilon) Y_{L''}(\hat{\epsilon}) h_{\ell''}^+(p|\vec{r} - \vec{R}|) Y_{L''}(\vec{r} - \vec{R}) = \text{AI.7}$$

$$4\pi \sum_L h_{\ell'}^+(pR) Y_L(\hat{R}) \sum_{L''} C_{LL''L'} i^{l-l'-l''} j\ell''(p\varepsilon) Y_{L''}(\hat{\epsilon}) j\ell(p\gamma) Y_L(\hat{r})$$

Comparing coefficients of $j\ell''(p\varepsilon) Y_{L''}(\hat{\epsilon})$ we obtain:

$$h_{\ell''}^+(p|\vec{r} - \vec{R}|) Y_{L''}(\vec{r} - \vec{R}) =$$

$$4\pi \sum_{L''} C_{LL''L'} i^{l-l'-l''} j\ell(p\gamma) h_{\ell'}^+(pR) Y_L(\hat{r}) Y_{L'}(\hat{R}) \text{ AI.8}$$

where, in order to emphasize the similarity of AI.3 and AI.8, we have rearranged the indices on the C in AI.8.

III. PREDICTOR CORRECTOR ALGORITHM FOR KKR

A growing body of experience suggests that in order to hold the fractional error in the Brillouin zone integrations required to describe reflectivity and photoemission to less than a few percent band energies are required at 10^4 to 10^6 independent positions in the Brillouin zone. In this section we address ourselves to the problem of obtaining accurate band energies at these densities.

Recent calculations by Mueller¹⁰ and Janak¹¹ suggest that what we might call mathematical interpolation is adequate for generating energies at 10^4 to 10^6 \vec{k} 's given energies at $\sim 3 \cdot 10^3$ \vec{k} 's. By mathematical interpolation we mean any scheme not based on a

physical model, e.g., quadratic interpolation.

Due to its speed, mathematical interpolation is preferred whenever it is adequate. The problem, therefore, is to obtain energies at $3 \cdot 10^3$ k's. The most common solution to this problem consists of computing energies at ~ 100 k's using APW, OPW or KKR techniques and then fitting what we can call physical interpolation schemes to these energies. Examples of what we refer to as physical interpolation schemes are the empirical LCAO, the $k \cdot p$, pseudopotential and the combined interpolation scheme of Mueller. These methods have proven capable of duplicating APW or KKR energies to .01 Ryd. The only disadvantage of these techniques is that the parameters specifying the various schemes are related in very complicated nonlinear ways to the input "first principle" energies. We would like to propose in what follows that the KKR scheme can be made fast enough to obviate the need for these physical interpolation schemes. A supplementary motivation for reducing to time required for KKR calculations is that band calculations corresponding to large and/or complicated unit cells constitute a very promising approach to the alloy problem and other disordered systems.

Any attempt to speed up a KKR calculation must, of course, begin with a knowledge of where and how the time is being spent. For calculations at low energies ($l_{\max} < 2$ or 3) and for simple unit cells (one or two atoms per unit cell), computer time is divided about evenly between structure constant evaluation and determinant evaluation.⁹ At higher energies and for larger unit cells a greater and greater fraction of the time is spent locating the solutions of secular equation; conversely a smaller and smaller fraction of the time is devoted to the construction of the secular matrix. The time T required for root finding is given approximately by

$$T \propto N_d [N_a (l_{\max} + 1)^2]^3 \quad 3.1$$

where N_d is the number of determinant evaluations required to locate a solution; N_a is the number of atoms per unit cell and l_{\max} is the largest value of angular momentum considered to be scattered by the atomic potential.

Frequently N_d is between 10 and 15 so that a considerable timesaving can be realized by accurately predicting the band

energies $E_n(\vec{k})$, thereby reducing the scope of the energy search. (We are assuming here that E is being scanned for fixed \vec{k} .) Accurate prediction can be obtained by extrapolation from solutions already found at neighboring \vec{k} -points. We are presently using quadratic extrapolation along lines in k -space. Since the secular matrix must be constructed for each determinant evaluation, the reduction of N_d by a factor of five, say, also reduces the total computation time by almost the same amount.

As important as N_d is to the computation time, 3.1 clearly shows that at least for large systems the greatest potential time-saving lies in the cubic term. The KKR scheme, in large part because the basis set is used to expand only the scattered wave and not the entire wave function, requires a relatively small secular equation. If, however, the same techniques used to predict $E_n(\vec{k})$ can be used to predict the coefficient vector $\beta_L(\vec{k})$ appearing in 1.10, we might hope to replace the cube in 3.1 by a square. This change would clearly be of profound significance for very large systems.

The cubic term in 3.1 arises if we correct our predicted energies and coefficient vector using any method involving determinant evaluation or linear equation solver employing Gaussian elimination. To be more concrete, let us describe a working but nonetheless cubic process. Given a predicted energy and β -vector ($\vec{\beta}$) and the KKR secular matrix evaluated at the predicted energy (call it M), we can then use the inverse power method which is defined by the following iteration

$$\vec{\beta}^{n+1} = M^{-1} \vec{\beta}^n \quad 3.2$$

to correct $\vec{\beta}$. Since this is a quadratic process once M is triangularly decomposed and since the iteration usually converges in 2 or 3 steps, the cubic part of this procedure is the triangularization of M . Since M is hermitian, the latter takes the same amount of time and is essentially identical to computing the determinant of M . We must, of course, also correct our predicted energy. This can be done by noting that $\langle \beta^n | \beta^n \rangle / \langle \beta^n | \beta^{n+1} \rangle$ is the relevant eigenvalue (the one we want to force to zero by choice of energy) of M . Thus, if we carry out the inverse iteration at a different but nearby energy, we can interpolate (the secant method) to find an improved energy and β -vector. This super-iteration, if you will, also converges in 2 or 3 steps usually, but since the underlying process is Gaussian elimination, we have by comparison with more common techniques succeeded only in reducing N_b by a factor of 3 or so.

A factor of 3 is still, however, an important time reduction, so before going on to our efforts to replace the cubic process let us point out some of the advantages of this particular cubic process. If we are going to use any procedure at thousands of points in the Brillouin zone, it must be completely automated. In particular it must work with at most a minor slowdown in the vicinity of the free electron singularities of M . In the vicinity of such a singularity M has eigenvalues which become arbitrarily large and discontinuously change sign. These eigenvalues and the corresponding eigenvectors are, of course, irrelevant to the problem inasmuch as we are interested only in the smallest eigenvalue and its eigenvector. Inverse iteration has the virtue of quickly eliminating the effects of large eigenvalues, so that aside from limits on the numerical accuracy of the small eigenvalues and the corresponding vectors imposed by the presence of very large eigenvalues, inverse iteration offers a complete solution to the free-electron singularity problem.

We should also mention that loss of significance near singularities of M caused us to abandon an attempt to use the Newton-Raphson analogue of the secant method to correct the predicted energy. The Newton-Raphson method, defined by the following:

$$E = E_p - \langle \beta | M(E_p) | \beta \rangle / \langle \beta | M'(E_p) | \beta \rangle \quad 3.3$$

where E_p is the predicted energy and E the corrected energy, requires the evaluation of the energy derivative ($M'(E_p)$) of M . The loss of significance near singularities has yet to seriously corrupt our evaluations of M ; it is, of course, a much more serious problem in the evaluation of M' and did on occasion cause the Newton-Raphson iteration to break down. An occasional breakdown is, as we said, unacceptable in a completely automated procedure.

Let us now describe our efforts to eliminate the N^3 (where $N = N_a (\ell_{\max}+1)^2$) portions of the calculation altogether.

Nesbet¹² has suggested that a variation on the usual relaxation technique provides just what we are looking for. Relaxation in this context consists of correcting β by iteratively sweeping through the following sequence of operations:

$$\lambda = \langle \beta | M | \beta \rangle / \langle \beta | \beta \rangle \quad 3.4$$

$$\Delta_i = (\lambda \beta_i - \sum_j M_{ij} \beta_j) / (M_{ii} - \lambda) \quad 3.5$$

$$\beta_i = \beta_i + \Delta_i \quad 3.6$$

Problems arise, of course, when $M_{ii} - \lambda$ becomes small. Nesbet suggests restarting the sweep through the i index whenever Δ_i exceeds a tolerance. We suggest the following route around such an obstacle. When the Δ_i given by 3.5 exceeds a tolerance such as .1 (assume $\langle \beta | \beta \rangle \sim 1$) construct and solve exactly for the eigenvalues and vectors of M in the two-dimensional basis consisting of the present approximation to β and the i th trial function. The 2×2 matrix is the following:

$$\begin{aligned} & \langle \beta | M | \beta \rangle - \lambda \langle \beta | \beta \rangle \quad \langle \beta | M | i \rangle - \lambda \langle \beta | i \rangle \\ & \langle i | M | \beta \rangle - \lambda \langle i | \beta \rangle \quad M_{ii} - \lambda \end{aligned} \quad 3.7$$

Of the two solutions which emerge we choose the one whose eigenvector has the largest projection on β . This procedure has worked without fail in all our testing. Since, however, convergence is not guaranteed, we strongly suggest monitoring this part of the iteration.

Our numerical experiments with this iteration technique have been confined to the KKR equation implied for the Chodorow potential for Cu and $l_{\max} = 4$. Since the corresponding secular equation is 25×25 and since the coefficient of N^3 for triangularization is only 1/6, the effect of replacing inverse iteration with relaxation in this case is not profound. It does improve our speed by an additional factor of 2, and we expect the effect to be much greater when the technique is applied to our semiconductor programs described in Part I (the secular equation in that case is 100×100).

Conclusions

We shall conclude now with a series of comparative remarks concerning the various methods of band calculation considered during this conference. There is probably little disagreement with the contention that the APW and OPW methods have been more productive than other methods for tightly and loosely packed materials, respectively. It is also probable, however, that the two most important criteria upon which methods will be judged in the future are (1) the dimension of the secular equation required and (2) whether or not the band structure is implicit (implied by the determinant of a matrix equalling zero) or explicit (requiring a single diagonalization to obtain all the energies given k or vice versa). Both the APW and KKR methods are implicit, so that the only important distinction between the two is the fact that the KKR secular equation is much smaller. Connolly¹³ has suggested that the relative simplicity of APW matrix elements compensates for the larger secular equation. We suggest that this statement is valid only in an increasingly small number of cases. First, the time required to construct the KKR matrix will be steadily reduced by more sophisticated programming;⁹ the time required for determinant evaluation, on the other hand, is not likely to diminish. Second, the statement can only be true for low energies (small ℓ_{\max}) and simple unit cells, for as either ℓ_{\max} or the complexity of the basis is increased, the fraction of time spent on matrix construction in the KKR method rapidly decreases.

In loosely packed structures the OPW, LCAO and KKR methods have been used with success. Here again the most frequently used method (OPW) is characterized by relative simplicity but very large secular equations. Both the OPW and LCAO methods are explicit while the KKR method is implicit. The KKR and LCAO methods both employ much smaller secular equations than the OPW method. The size of the LCAO secular equation is very small for the light elements giving the method a double advantage for low atomic number (z); since, however, it is roughly proportional to z , the advantage of a small secular equation obtains only for the upper part of the periodic table.

Two recent developments offer evidence that the KKR method may soon prove to be significantly superior for all materials. First, the work of Jepsen and Marcus¹⁴ and Kambé¹⁵ shows that KKR can be made explicit and that the size of the secular equation can be reduced substantially for all but the simplest cases ($\ell_{\max} \leq 2$, one atom per unit cell and close packing). The second development may be relevant only to the computation of the wave function and matrix elements. Kleinman and Shurteff have recently made a very detailed study of the MOPW¹⁶ method. In this method wave functions for d-electrons in solids are constructed by combining atomic-like

d-orbitals which are arbitrarily accurate near the nucleus with plane waves which are intended to correct the d-orbital in the outer portion of the unit cell. It is shown¹⁶ that the plane waves so introduced corrupt the wave function near the nucleus. To avoid this difficulty, one must use a set of plane waves large enough (thousands) to ensure the existence of linear combinations capable of providing the required correction in the outer portion of the cell and possessing negligible amplitude in the core region. It is difficult to see why this phenomenon will not create difficulties with the LCAO method - in other words, why the tails of orbitals centered on other sites will not corrupt the core portion of an orbital centered at a given site. Thus on the basis of existing experience and recent developments, we expect to see the productivity of the LCAO and particularly the KKR method rise significantly as time goes on.

References

1. A. R. Williams: Phys. Rev. 1, 3417 (1970).
2. A. Ralston: A First Course in Numerical Analysis (McGraw-Hill Inc., New York, 1965).
3. L. I. Schiff: Quantum Mechanics, 2nd Edition (McGraw-Hill Inc., New York, 1955).
4. R. N. Euwema, D. J. Stukel, and T. C. Collins: (This Proceedings).
5. G. Dresselhaus and M. S. Dresselhaus: Phys. Rev. 160 No. 3, 649 (1967).
6. F. Herman, R. L. Kortum, C. D. Kuglin, and R. A. Short: Quantum Theory of Atoms, Molecules, and the Solid State: A Tribute to J. C. Slater, edited by Per-Olov Löwdin (Academic Press Inc., New York, 1966).
7. M. L. Cohen and T. K. Bergstresser: Phys. Rev. 141, 789 (1966).
8. A. Messiah: Quantum Mechanics, Vol. I, (Interscience Publishers Inc., New York 1961).
9. H. L. Davis: (This Proceedings).
10. F. M. Mueller, J. W. Garland, M. H. Cohen, K. H. Bennemann: (To be published).
11. J. F. Janak: (This Proceedings).
12. R. K. Nesbet: J. Chem., Phys. 43 No. 1, 311 (1965).
13. J. W. D. Connolly: (This Proceedings).
14. D. J. Jepsen and P. M. Marcus: (This Proceedings).
15. K. Kambe: (This Proceedings).
16. L. Kleinman and R. Shurtliff: Phys. Rev. 188 No. 3, 1111 (1969).

COMMENTS ON THE KKR WAVEFUNCTIONS; EXTENSION OF THE SPHERICAL
WAVE EXPANSION BEYOND THE MUFFIN TINS

O. Krogh Andersen*

Physics Department, University of Pennsylvania
Philadelphia, Pennsylvania

The solution of Schrödinger's equation for an assembly of non-overlapping muffin-tin potentials may be constructed in terms of Linear Combinations of - suitably chosen - Muffin Tin Orbitals (LCMTO). For the special case of a perfect crystal this leads to the well-known KKR equations.^{1,2,3} We hope to demonstrate that this approach, which bears some resemblance to the original work of Korringa¹, gives a better insight into the properties of the wavefunctions than does the more popular Green's-function formalism introduced by Kohn and Rostoker.^{2,3,4}

We shall define $\psi_{lm,E}$ as that solution of Schrödinger's equation for an isolated, muffin-tin potential, V , which is regular at the muffin-tin center:

$$\psi_{lm,E}(r) = i^l \sum_{lm}(\hat{r}) \begin{cases} u_{l,E}(r); & \text{for } r \leq s \\ t_{l,E}(r); & \text{for } s \leq r \end{cases} \quad (1)$$

where:

$$t_{l,E}(r) = \lambda l \left[n_l(\lambda r) - \cot \delta_{l,E} f_l(\lambda r) \right] \quad (1a)$$

$u_{l,E}$ is obtained by integration of the radial Schrödinger equation

*Present address: Fysisk Lab. I, H. C. Ørsted Institute,
University of Copenhagen, Denmark

and $\delta_{\ell,E}$ is the phase shift, so that ψ and its first derivative are continuous at the muffin-tin sphere, $r = S$. Further $\kappa = \sqrt{E}$ or $i\sqrt{-E}$ for $E > 0$ or $E < 0$ respectively. In the continuum ($E > 0$) ψ decreases as $1/r$ at infinite and it may be δ -function normalized. For $E < 0$ ψ generally diverges as $\exp(r\sqrt{-E})/r$; only if $\cot\delta = i$, is ψ a bound state with a tail decreasing like $\exp(-r\sqrt{-E})/r$.

The Muffin Tin Orbital (MTO), ϕ , is obtained by adding on to ψ a solution of the wave equation (homogeneous Schrödinger equation) which is regular at the center. Thereby the MTO is made normalizable at all energies:

$$\phi_{\ell m, E}(r) = i^\ell Y_{\ell m}(\hat{r}) \begin{cases} U_{\ell, E}(r) + \lambda (\cot \delta_{\ell, E} - i) j_\ell(\lambda r) \\ \lambda [N_\ell(\lambda r) - i] j_\ell(\lambda r) \end{cases} \quad (2)$$

Due to the wave equation term inside the muffin-tin, the MTO is, however, not a wavefunction of the muffin-tin potential.

The wavefunction for an assembly of non-overlapping, possibly different, muffin-tin potentials, $\sum_Q V^Q(r-Q)$, may now be written as an LCMTO:

$$\Phi = \sum_Q \sum_{\ell m} C_{\ell m}^Q \phi_{\ell m, E}^Q(r-Q) \quad (3)$$

For $Q' \neq Q$, only the tail of $\phi^Q(r-Q)$ overlaps with the potential $V^{Q'}(r-Q')$. This tail, centered at Q' , is a solution of the wave equation and is regular at Q' , therefore its expansion around Q' contains only spherical Bessel rather than spherical Neumann functions. It now follows that the coefficients $C_{\ell m}^Q$ are determined by the condition that inside any particular muffin tin (center Q') the sum of the tails from all other muffin-tins cancel the wave equation-term of the central MTO:

$$\sum_{\ell m} \left\{ C_{\ell m}^{Q'} i^\ell Y_{\ell m}(\hat{r}-\hat{Q}') \lambda (\cot \delta_{\ell, E}^{Q'} - i) j_\ell(\lambda(r-Q')) + \sum_{Q \neq Q'} C_{\ell m}^Q \phi_{\ell m, E}^Q(r-Q) \right\} = 0 \quad (4)$$

or:

$$\sum_{\ell m Q} (\ell' m' Q' | \ell m Q) C_{\ell m}^Q = 0 \quad (5)$$

where $(\ell' m' Q' | \ell m Q)$ is the coefficient in the expansion of $\phi_{\ell m}^Q(\underline{r}-\underline{Q})$ in $i^\ell Y_{\ell m}(\underline{r}-\underline{Q}') j_{\ell'}(\kappa |\underline{r}-\underline{Q}'|)$. For finite systems there is further the condition that the energy must be negative.

The expansion around \underline{Q}' of the tail centered at \underline{Q} is:⁵

$$\begin{aligned} i^\ell \sum_{\ell m} Y_{\ell m}(\hat{\gamma}) h_L(\underline{r} \hat{\gamma}) = & \sum_{\ell m} \sum_{\ell' m'} 4\pi C_{\ell m \ell' m'} i^\ell Y_{\ell m}(\hat{\gamma}) i^{-\ell'} Y_{\ell' m'}^*(\hat{Q}-\hat{Q}') \times \\ & \begin{cases} j_\ell(\underline{r} \hat{\gamma}) h_{\ell'}(\underline{r} |\underline{Q}-\underline{Q}'|) ; & \text{for } \hat{\gamma}' < |\underline{Q}-\underline{Q}'| \\ h_\ell(\underline{r} \hat{\gamma}') j_{\ell'}(\underline{r} |\underline{Q}-\underline{Q}'|) ; & \text{for } \hat{\gamma}' > |\underline{Q}-\underline{Q}'| \end{cases} \end{aligned} \quad (6)$$

Here $\underline{\rho}' = \underline{r}-\underline{Q}'$ and $\underline{\rho} = \underline{r}-\underline{Q}$, h_L is any linear combination of spherical Bessel and Neumann functions, specifically $\kappa(n_L - ij_L)$, and $C_{\ell M, \ell' m, \ell' m'}$ are the Gaunt coefficients. For the special case of $h_L = j_L$ this expansion has previously been given by Kohn and Rostoker² (KR.A2.12). We are now able to express:

$$\begin{aligned} (\ell' m' Q' | \ell m Q) = & \delta_{\ell \ell'} (\cot \delta_{\ell, \varepsilon}^Q - i) \delta_{m m'} \delta_{Q Q'} \\ & + (1 - \delta_{Q Q'}) \sum_L 4\pi C_{\ell m \ell' m' L M} i^{-\ell} Y_{L M}^*(\hat{Q}-\hat{Q}') \left[\delta_{\ell \ell'} (\underline{r} |\underline{Q}-\underline{Q}'|) - i j_{\ell'} (\underline{r} |\underline{Q}-\underline{Q}'|) \right] \end{aligned}$$

In the case of a perfect crystal $\underline{Q} = \underline{q} + \underline{R}$ and $C_{\ell m}^{Q+R} = e^{i \underline{k} \cdot \underline{R}} C_{\ell m}^Q$, where \underline{R} is a lattice translation and \underline{q} the position of an atom in the unit cell. Contraction on \underline{R} in (5) and (7) yields the KKR equation. With only one atom per unit cell this matrix

reduces to that of Kohn and Rostoker² (K.R.:3.17, A3.3, A2.14, and A2.22) apart from factors of i^l , which, following Ham and Segal³, we have included in the definition of the wavefunction. For $E > 0$ the lattice-summation converges in the generalized sense, as the Fourier-transform of the tail exists, and the summation may be performed by the standard Ewald technique.

From the above development it is obvious that if $C_{\ell'm'}^{Q'}$ is a solution of (5), then inside the muffin-tin at \underline{Q}' the wavefunction is:

$$\underline{\Phi} = \sum_{\ell'm'} C_{\ell'm'}^{Q'} \Psi_{\ell'm'}^{Q'} (\underline{r} - \underline{Q}') \quad (8)$$

which should be compared to (3). Moreover since the upper part of (6) is valid not only inside the muffin-tin at \underline{Q}' , but inside a sphere touching the center of the nearest neighbor, the spherical wave expansion (8) holds for the interstitial region, beyond the muffin-tin, but still inside the nearest neighbor sphere. For fcc and bcc structures with one atom per unit cell (8) is valid throughout the Wigner-Seitz cell. For more open structures, structures with more than one atom per unit cell, and molecules it is useful to further continue the spherical wave expansion. Using the lower part of (5) we obtain:

$$\begin{aligned} \underline{\Phi} = & \sum_{\ell'm'} i^\ell Y_{\ell'm'}(\underline{r} - \underline{Q}') \times \\ & [n_\ell(\underline{r} | \underline{r} - \underline{Q}|) \sum_{\ell'm} \sum_Q J_{\ell'm'}^{Q'Q} C_{\ell'm}^Q - i_\ell(\underline{r} | \underline{r} - \underline{Q}'|) \sum_{\ell'm} \sum_Q N_{\ell'm'}^{Q'Q} C_{\ell'm}^Q] \end{aligned}$$

where:

$$\begin{aligned} J_{\ell'm'l'm'}^{Q'Q} = & \delta_{\ell\ell'} \delta_{mm'} \delta_{QQ'} + \\ & (1 - \delta_{QQ'}) \sum_L 4\pi C_{\ell'm'l'm'L}^{-1} Y_{\ell'm'}^{*(Q-Q')} \times \\ & j_L(\underline{r} | \underline{Q} - \underline{Q}'|). \end{aligned}$$

$$N_{lm'l'm}^{Q'Q} = \chi \cot \delta_{l,E}^Q \delta_{ll'} \delta_{mm'} \delta_{QQ'} + \\ (1 - \delta_{QQ'}) \sum_L 4\pi C_{lm'l'm'Lm} i^{-L} \hat{Y}_{Lm}^*(\underline{Q} - \underline{Q}') \chi n_l(\chi |Q - Q'|) \quad (9)$$

The usefulness of this formula comes from the fact that the summation, \sum_L , includes only centers inside a sphere of radius $\rho' = |\underline{r} - \underline{Q}'|^Q$ centered at \underline{Q}' . Finally it may be noted that, when in practice C_{lm}^Q are determined from a matrix truncated in l_m , the representation (9) may be poor if more than a few Q 's are included in the summation.

Further developments and applications of these ideas including the treatment of nonmuffin-tin perturbation will be published elsewhere.⁵

References

1. J. Korringa, Physica 13, 392 (1947).
2. W. Kohn and N. Rostoker, Phys. Rev. 94, 1111 (1954).
3. F. S. Ham and B. Segall, Phys. Rev. 124, 1786 (1961).
4. A. R. Williams, this Proceedings.
K. H. Johnson, this Proceedings.
5. O. K. Andersen, to be published.

EFFICIENT NUMERICAL TECHNIQUES FOR THE CALCULATION OF KKR
STRUCTURE CONSTANTS*

Harold L. Davis

Solid State Division
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37830

ABSTRACT

Simple computational techniques are described which enable the inherent rapid convergence (low order matrix) of the KKR method to be fully exploited for the performance of efficient band structural calculations. These techniques are for the computer evaluation of the KKR structure constants, which are the basic potential independent functions necessary when using this first principles method. By using both analytical and computational partitioning of standard formulas for the structure constants, the present approach enables them to be evaluated by expending computer times which are of the same order as the times required to calculate the KKR determinants from their matrix elements.

INTRODUCTION

As may be seen by inspecting various articles in this volume, many methods have been proposed for the calculation of electronic band structures from periodic potentials. Each of these methods presently being actively employed for calculations has its distinct advantages and disadvantages, depending upon the major components of the orbital angular momentum for the band electrons of interest, upon topology and other details of the approximate potential necessary to insure that the resulting calculations have realistic applicability to the solid of interest, etc. As an example, and germane

* Research sponsored by the U. S. Atomic Energy Commission under contract with Union Carbide Corporation.

to this article, for metals where "d- or f-bands" are of prime importance, the work of many investigators has shown the Korringa-Kohn-Rostoker^{1,2} (KKR) method to be of both continuing applicability and outstanding practicability because of its rapid convergence. This rapid convergence results in the matrix, or determinant, which must be manipulated in KKR calculations being inherently of low order, even for points of lowest symmetry in the Brillouin zone. The importance of low matrix order is well illustrated by the fact that the computer time required for evaluation of the determinant of a matrix is approximately proportional to the cube of the order.

The price paid when taking advantage of the low order of the KKR matrix is the necessity of dealing with fair complexity in the individual matrix elements. Such complexity is traced to the mathematical form of the KKR "structure constants", which are the basic building blocks of the potential-independent part of a KKR matrix element. However, in performing KKR calculations two techniques are useful in obtaining the necessary structure constants. First, since by introduction of dimensionless units the KKR structure constants are cast in a form independent of lattice constant, they may be worthwhilely tabulated as functions of dimensionless energy, E , and dimensionless wave number, \vec{k} , for the various individual lattice symmetries. For those \vec{k} points and E regions where the tabulations apply, E vs. \vec{k} determinations may then proceed by doing only reasonably simple numerical calculations. In fact, a limited tabulation of the structure constants for the bcc and fcc lattices by Segall and Ham³ has been of value in this connection; also, some KKR calculations have proceeded using structure constant tabulations stored on computer cards or magnetic tape. However, the tabulation approach to performing KKR calculations has the liability that one's interest is intrinsically restricted to those \vec{k} points where tabulation exists. Also, the possibility exists that the final calculational accuracy can suffer unless a reasonably fine E -mesh is used for the tabulation, which could greatly increase the size of the tabulation.

An alternative procedure to tabulation, and the approach we favor, is the direct computation of the KKR structure constants as needed during the course of calculating band structural information from a potential. This procedure is, in fact, almost a necessity if one desires the freedom of performing accurate KKR calculations throughout the Brillouin zone; and, once implemented, it allows the precision calculation of many band structural details having physical interest.⁴⁻⁹ However, provided one exercises reasonable care in the procedures used in the computer programming of the structure constants, only a modest price in increase of computer running time is paid when adopting the approach of direct computation of structure constants as needed. In fact, as we shall attempt to document here, this approach requires expenditures of computer time of only the same order as the time required to evaluate the determinant of the KKR matrix given its matrix elements. This time is then

reasonably short because of the inherent low order of the KKR matrix, which means the KKR method can indeed be an exceedingly efficient mechanism for obtaining band theoretic information directly from a given potential. It is thus the purpose of this article to describe certain simple computational techniques which we have found to be of value when constructing and implementing computer codes for performing KKR calculations by the direct evaluation of structure constants as needed. Our motivation in presenting these techniques is that we feel that they might be of value to other workers who wish to implement the KKR method, and, perhaps, our experiences can ease their labor in performing the necessary computational details.

Herein, to illustrate the computational techniques, we will restrict the discussion to a description of nonrelativistic KKR calculations for cubic lattices having one atom per unit cell, and to solids where the muffin-tin approximation to the potential is valid. Such illustrative restrictions, however, do not drastically limit the application of the techniques, since the same structure constants discussed here are also required for many other cases not satisfying these restrictions. For relativistic KKR calculations, several authors¹⁰⁻¹² have shown that the required matrix elements are expressible in terms of the same structure constants required in non-relativistic calculations. When the solid of interest has more than one atom per unit cell it has been shown by Faulkner,¹³ and independently by Johnson,¹⁴ that the required structure constants can be expressed as linear combinations of structure constants for the Bravais lattice; thus, e.g., NaCl-type structure constants are obtainable from those for the fcc lattice. Also, as recently advocated by Segall,¹⁵ using experience gained from an early calculation of Ham and Segall¹⁶ on the Mathieu potential, a fruitful method of dealing with non-muffin-tin potentials appears to be a perturbation treatment based on a muffin-tin calculation.

THE KKR METHOD

Since the derivation of the KKR method is amply covered by the original papers^{1,2,16} and a recent review,¹⁷ only an outline of the KKR equations required in the present discussion will be included here. Also, throughout this article exclusive use will be made of the dimensionless units described by Segall and Ham,¹⁷ with the zero of energy fixed to coincide with the constant value of the potential between the muffin-tin spheres.

It is useful to expand the wave functions within a muffin-tin sphere as

$$\psi_{\vec{K}}(\vec{r}) = \sum_{\ell=0}^{\ell_{\max}} \sum_{m=-\ell}^{\ell} i^{\ell} C_{\ell m}^{\vec{K}} R_{\ell}^E(r) Y_{\ell m}(\theta, \phi). \quad (1)$$

The $C_{\ell m}^{\vec{k}}$ are expansion coefficients determinable by the KKR techniques, the $Y_{\ell m}(\theta, \phi)$ are real spherical harmonics, and the radial functions $R_{\ell}^E(r)$ are calculable using the radial equation containing the potential $V(r)$. Use of this expansion leads to consideration of the determinant

$$\text{Det } |A_{\ell m, \ell' m'} + \sqrt{E} \delta_{\ell \ell}, \delta_{mm}, \cot \eta_{\ell}(E)|. \quad (2)$$

It is the zeros of this determinant which implicitly generate the dispersion relations, E vs. \vec{k} , for the various bands of the muffin-tin potential. The order of the determinant, being governed by the truncation ℓ value in the expansion (1), is $(\ell_{\max} + 1)^2$. Since it is well established that highly accurate results are obtained with $\ell_{\max} = 2$, or at most = 3, at least for solids whose composite elements are in the first half of the periodic table, the resulting low order of the determinant is one of the major practical advantages of the KKR method.

The $\cot \eta_{\ell}(E)$ in (2) are cotangents of phase shifts for a particle of energy E scattering off a single potential, $V(r)$, isolated in free space, and obtainable from solutions of the radial equation or used in parameterization schemes to fit data.^{17,18} Irregardless, our interest here is the potential-independent parts of the KKR matrix. These parts are the $A_{\ell m, \ell' m'}$ which are conveniently expressed as linear combinations of quantities, $D_{LM}(\vec{k}, E)$, called "structure constants",

$$A_{\ell m, \ell' m'} = 4\pi \sum_L \sum_M C_{\ell m, \ell' m'}^{LM} D_{LM}(\vec{k}, E), \quad (3)$$

with the $C_{\ell m, \ell' m'}^{LM}$ being Gaunt factors

$$C_{\ell m, \ell' m'}^{LM} = \int d\Omega Y_{LM}(\theta, \phi) Y_{\ell m}(\theta, \phi) Y_{\ell' m'}(\theta, \phi). \quad (4)$$

For a given \vec{k} and E pair, due to properties of the $C_{\ell m, \ell' m'}^{LM}$, only $(2\ell_{\max} + 1)^2$ distinct structure constants need be evaluated since the largest L value entering is $2\ell_{\max}$.

The $D_{LM}(\vec{k}, E)$ are most readily evaluated in practice by expressing them as the sum of three parts,

$$D_{LM}(\vec{k}, E) = D_{LM}^{(1)} + D_{LM}^{(2)} + \delta_{L0} D_{00}^{(3)}, \quad (5)$$

by use of an Ewald procedure. Explicit expressions for these three parts, as given by Ham and Segall,¹⁶ are

$$D_{LM}^{(1)} = -4\pi \exp(E/\eta) \tau^{-1} (\sqrt{E})^{-L} \sum_n k_n^L [k_n^2 - E]^{-1} \exp(-k_n^2/\eta) Y_{LM}(\vec{k}_n), \quad (6)$$

$$\begin{aligned} D_{LM}^{(2)} = & (-2)^{L+1} \pi^{-1/2} (\sqrt{E})^{-L} \sum_s' [i^L \exp(i\vec{k} \cdot \vec{r}_s)] \times \\ & \times Y_{LM}(\vec{r}_s) r_s^L \int_{\sqrt{\eta}/2}^{\infty} \xi^{2L} \exp[-\xi^2 r_s^2 + (E/4\xi^2)] d\xi, \end{aligned} \quad (7)$$

$$D_{00}^{(3)} = -\sqrt{\eta} (2\pi)^{-1} \sum_{j=0}^{\infty} (E/\eta)^j [j!(2j-1)]^{-1}. \quad (8)$$

In (6), τ is the unit cell volume, $\vec{k}_n = \vec{k} + \vec{k}_n$, and $k_n = |\vec{k}_n|$, with the \vec{k}_n being the reciprocal lattice vectors. The prime on the summation in (7) indicates the term $\vec{r}_s = 0$ is omitted, with the \vec{r}_s being the spatial lattice vectors. The value of expression (5) in practical calculations is the arbitrariness of the Ewald parameter, η , since the result (5) is independent of the value of $\eta > 0$ provided the sums in (6) through (8) are carried to convergence. But, of course, η must be held fixed in (6) through (8) during a given evaluation of (5). This means that η may be judiciously chosen in order to have the calculation of a structure constant set, for all allowed L and M, proceed in the most efficient manner. Furthermore, the exact choice of η may vary, from calculation to calculation, depending upon the explicit means being used to find the zeros of the determinant (2).

BASIC COMPUTATIONAL PHILOSOPHY AND GENERAL COMMENTS

Numerous analytical and numerical procedures may be adopted for the evaluation of the functions $D_{LM}(\vec{k}, E)$ on a high speed digital computer--so many different procedures that it is difficult to convince oneself that a given set of procedures would be the most efficient, i.e., the set which would use the least computer time. Thus, we will here only describe those techniques which we have found to be most useful during our experiences with the KKR method, without attempting justification of ultimate theoretical efficiency.

In order to find the zeros of the KKR determinant (2), two obvious searching patterns present themselves. First, constant- \vec{k} searches, or fix \vec{k} in the determinant and vary E to find the determinantal zeros. Second, constant-E searches, or fix E and vary \vec{k} to find zeros. Both of these searching possibilities have possible applications, so any versatile computer program for the evaluation of the KKR structure constants should enable efficient implementation

of either possibility. That is, it is desirable to have a set of independent subprograms whose order of execution can be interchanged at will depending upon the type of search being performed during a given calculation. Before writing such subprograms, however, it is helpful to perform a series of test calculations, whose results indicate the number of terms which must be retained in the sums (6) through (8) for a given η and a desired accuracy. Such test results are invaluable for making programming decisions regarding storage size (dimensions) for FORTRAN arrays representing intermediate calculations.

Following the above approach, the computer evaluation of a set of structure constants is then best performed by subprograms which perform four separate computation steps. Assuming ℓ_{\max} and η fixed, these steps are:

- (a). The computation and storage, ready for future use, of all quantities independent of E and \vec{k} . This procedure need be executed only once per computer run.
- (b). The computation and storage of all quantities dependent upon \vec{k} .
- (c). The computation and storage of all quantities dependent upon E .
- (d). The use of the results stored in steps (a) through (c) to compute all the $D_{LM}(\vec{k}, E)$ for a given \vec{k} and E .

As mentioned before, programming care should be exercised to enable steps (b) and (c) to be interchanged depending upon the type of search. With such an allowance, the performance of constant- \vec{k} searches require only steps (c) and (d) to be continuously re-executed as E is varied to find the zeros of the determinant (2), while for the performance of constant- E searches only steps (b) and (d) need be re-executed as \vec{k} is varied. More detailed suggestions concerning specific steps of the calculation are given below.

COMPUTATION OF SPHERICAL HARMONICS AND $D_{00}^{(3)}$

The real spherical harmonics occurring in (6) and (7) are

$$Y_{LM}(\theta, \phi) = q_L^J P_L^J (\cos \theta) \left\{ \begin{array}{l} \cos(J\phi) \\ \sin(J\phi) \end{array} \right\}. \quad (9)$$

Here, and throughout this article, $J = |M|$. The upper factor in the bracket is taken when $M \geq 0$, while the lower factor is for $M < 0$. The factors q_L^J are

$$q_L^J = \begin{cases} [(2L + 1)/4\pi]^{1/2} & \text{if } J = 0, \\ \{[(2L + 1)(L - J)!]/[2\pi(L + J)!]\}^{1/2} & \text{if } J \neq 0. \end{cases} \quad (10)$$

The Legendre polynomials are easily evaluated by use of the direct relations $P_0^0(x) = 1$, $P_1^0(x) = x$, $P_1^1(x) = (1-x^2)^{1/2}$, $P_2^1(x) = 3(1-x^2)^{1/2}x$, and standard recursion relations

$$P_n^m(x) = [(2n-1)/(n-m)]x P_{n-1}^m(x) - [(n+m-1)/(n-m)]P_{n-2}^m(x), \quad (11)$$

$$P_n^{m+1}(x) = 2m[x/(1-x^2)^{1/2}]P_n^m(x) - [(n+m)(n-m+1)]P_{n-1}^{m-1}(x). \quad (12)$$

The quantities q_L^J given by (10), and the factors involving n and m in (11) and (12), are suited to be included in step (a) of the calculations, as are the Legendre polynomials for the special cases $x = 0, \pm 1$. Once these are available the spherical harmonics are evaluated in either step (a) of the total calculation or step (b), depending upon whether the interest is in the spherical harmonics of (7) or (6). In these evaluations, the $\cos(J\phi)$ and $\sin(J\phi)$ factors, for all $J \leq 2\ell_{\max}$, are conveniently and efficiently obtainable from just $\cos \phi$ and $\sin \phi$ by using trigonometric identities for angle addition.

For the evaluation of $D_{00}^{(3)}$, the sum (8) is rewritten

$$D_{00}^{(3)} = \sum_{n=0}^{\infty} \alpha_n E^n, \quad (13)$$

with the α_n being recursively evaluated in step (a) of the calculations and the sum performed in step (c).

COMPUTATION OF THE $D_{LM}^{(1)}$

Use of dimensionless units allows the \vec{k}_n to be expressed

$$\vec{k}_n = m_n^x \hat{e}_x + m_n^y \hat{e}_y + m_n^z \hat{e}_z, \quad (14)$$

where the m 's are always integers for cubic lattices. With this notation, the exponential factors in the sum (6) may be expressed as

$$\exp(-k_n^2/\eta) = \exp[-(\vec{k}_n)^2/\eta] \exp[-(\vec{k})^2/\eta] \prod_{x,y,z} [\exp(-2k_x/\eta)]^{m_n^x}. \quad (15)$$

Thus, in step (b) of the total calculation, regardless of the number of \bar{k}_n 's which must be included in the evaluation of $D_{LM}^{(1)}$, it is necessary to compute only four exponentials; all others follow by simple, and fast, multiplicative arithmetic. This follows from (15), since the factors $\exp[-(\bar{k}_n)^2/\eta]$ are logically included in step (a) of the total calculation. The other factors present in Eq. (6) are evaluated in steps (a), (b), or (c) depending on their \bar{k} and E dependence, while the summation contained in (6) is performed in step (d).

COMPUTATION OF THE $D_{LM}^{(2)}$

Perhaps the most critical feature, with respect to the use of computer time, in the calculation of $D_{LM}^{(2)}$ is the technique one uses to evaluate the integrals in Eq. (7). Because the integrals contain the energy, their calculation or some fraction thereof must be included in step (c) of the total calculation. Thus, it is desirable to recast the integrals into a form where as much as possible of their evaluation is performed in step (a); also, it is necessary to evaluate those integrals for only unique r_s^2 . To expedite such an approach, notice

$$\begin{aligned} \int_{\sqrt{\eta}/2}^{\infty} \xi^{2L} \exp[-\xi^2 r^2 + E/(4\xi^2)] d\xi &= (1/2r^{2L+1}) \times \\ &\times \sum_{m=0}^{\infty} \{ [E(r/2)^2]^m / m! \} \Gamma(L - m + \frac{1}{2}, r^2 \eta/4), \end{aligned} \quad (16)$$

where $\Gamma(a,x)$ is an incomplete gamma function.¹⁹ We prefer the use of the convergence series (16) for these integrals over the asymptotic series given by Morse,²⁰ since (16) enables much more of the integrals' computation to be executed in step (a) of the calculations. To specifically evaluate the $\Gamma(a,x)$'s on a computer, we have found recursive relations to be unreliable, even in double precision arithmetic. This is true at least for the range of η 's we find desirable to use in the structure constant calculations. Instead, we use¹⁹

$$\Gamma(a,x) = e^{-x} x^a f(a,x), \quad (17)$$

where $f(a,x)$ is the continued fraction

$$f(a,x) = [\frac{1}{x+} \frac{1-a}{1+} \frac{1}{x+} \frac{2-a}{1+} \dots]. \quad (18)$$

Another technique we feel is useful in evaluating the $D_{LM}^{(2)}$ is to first group the sum over the spatial lattice vectors into sums over shells, and then analytically perform partial summations over

portions of the shell. That is, for cubic lattices a general spatial lattice vector is expressible,

$$\vec{r}_s = \pi(n_s^x \vec{e}_x + n_s^y \vec{e}_y + n_s^z \vec{e}_z), \quad (19)$$

where the n 's are integers. Then $D_{LM}^{(2)}$ may be written

$$D_{LM}^{(2)} = \sum''_s Q_L^J(\vec{r}_s) U_L^J(\vec{r}_s, \vec{k}) I_L(E, r_s) \sum_{i=1}^4 a_i^J(\vec{r}_s, \vec{k}) V_i^M(\vec{r}_s), \quad (20)$$

with the double prime on the s summation meaning the vector $\vec{r}_s = 0$ is excluded and the sum is to include only those \vec{r}_s vectors (19) where both $n_s^x \geq n_s^y \geq 0$ and $n_s^z \geq 0$.

To obtain (20) use has been made of the fact¹⁶ that for $E < 0$ $i^L D_{LM}(\vec{k}, E)$ must be real, and for $E > 0$, $D_{LM}(\vec{k}, E)$ must be real. So

$$\text{Real}[i^L \exp(i\vec{k} \cdot \vec{r}_s)] = \begin{cases} (-1)^{L/2} \cos(\vec{k} \cdot \vec{r}_s) & L \text{ even,} \\ (-1)^{(L+1)/2} \sin(\vec{k} \cdot \vec{r}_s) & L \text{ odd.} \end{cases} \quad (21)$$

In addition, to derive expressions for the quantities in (20), besides using (16), (17), (19), and (21), liberal use has been made of trigonometric identities, the equality $P_L^J(-x) = (-1)^{L+J} P_J^L(x)$, and the equality $P_J^L(0) = 0$ if $(L+J)$ odd. Expressions for the quantities $a_i^J(\vec{r}_s, \vec{k})$ and $V_i^M(\vec{r}_s)$ then result as are given by Tables I, II, and III.

TABLE I. Values of the functions $a_i^J(\vec{r}_s, \vec{k})$ used in Eq. (20).

	J even	J odd
i = 1	$\cos(\pi n_s^x k_x) \cos(\pi n_s^y k_y)$	$\sin(\pi n_s^x k_x) \cos(\pi n_s^y k_y)$
i = 2	$-\sin(\pi n_s^x k_x) \sin(\pi n_s^y k_y)$	$\cos(\pi n_s^x k_x) \sin(\pi n_s^y k_y)$
i = 3	$\cos(\pi n_s^y k_x) \cos(\pi n_s^x k_y)$	$\sin(\pi n_s^y k_x) \cos(\pi n_s^x k_y)$
i = 4	$-\sin(\pi n_s^y k_x) \sin(\pi n_s^x k_y)$	$\cos(\pi n_s^y k_x) \sin(\pi n_s^x k_y)$

TABLE II. Values of the functions $V_i^M(\vec{r}_S)$ used in Eq. (20) for J even. Here, $\phi = \arctan(n_y/n_x)$, $\beta = \arctan(n_x^x/n_y^y)$, and $\lambda_J = (-1)^J/2$. The first line of a tabular entry is for $M = 0$, the second for $M > 0$, and the last for $M < 0$.

	$0 \neq n_s^x \neq n_s^y \neq 0$	$n_s^x = n_s^y \neq 0$ or $n_s^x \neq n_s^y = 0$	$n_s^x = n_s^y = 0$
$i = 1$	$\begin{Bmatrix} 1 \\ \cos(J\phi) \\ \sin(J\phi) \end{Bmatrix} + \lambda_J \begin{Bmatrix} 1 \\ \cos(J\beta) \\ \sin(J\beta) \end{Bmatrix}$	$\begin{Bmatrix} 1 \\ \cos(J\phi) \\ \sin(J\phi) \end{Bmatrix}$	$\begin{Bmatrix} 1/2 \\ 1/2 \\ 0 \end{Bmatrix}$
$i = 2$	$\begin{Bmatrix} 1 \\ \cos(J\phi) \\ \sin(J\phi) \end{Bmatrix} - \lambda_J \begin{Bmatrix} 1 \\ \cos(J\beta) \\ \sin(J\beta) \end{Bmatrix}$	$\begin{Bmatrix} 1 \\ \cos(J\phi) \\ \sin(J\phi) \end{Bmatrix}$	$\begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}$
$i = 3$	$\begin{Bmatrix} 1 \\ \cos(J\beta) \\ \sin(J\beta) \end{Bmatrix} + \lambda_J \begin{Bmatrix} 1 \\ \cos(J\phi) \\ \sin(J\phi) \end{Bmatrix}$	$\lambda_J \begin{Bmatrix} 1 \\ \cos(J\phi) \\ \sin(J\phi) \end{Bmatrix}$	$\begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}$
$i = 4$	$\begin{Bmatrix} 1 \\ \cos(J\beta) \\ \sin(J\beta) \end{Bmatrix} - \lambda_J \begin{Bmatrix} 1 \\ \cos(J\phi) \\ \sin(J\phi) \end{Bmatrix}$	$-\lambda_J \begin{Bmatrix} 1 \\ \cos(J\phi) \\ \sin(J\phi) \end{Bmatrix}$	$\begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}$

TABLE III. Values of the functions $V_i^M(\vec{r}_S)$ used in Eq. (20) for J odd. The angles ϕ and β are as defined in the caption of Table II, and $\mu_J = (-1)^{(J-1)/2}$. The first line of a tabular entry is for $M > 0$, while the second is for $M < 0$.

	$0 \neq n_s^x \neq n_s^y \neq 0$	$n_s^x = n_s^y \neq 0$ or $n_s^x \neq n_s^y = 0$	$n_s^x = n_s^y = 0$
$i = 1$	$\begin{Bmatrix} \cos(J\phi) \\ \sin(J\phi) \end{Bmatrix} + \mu_J \begin{Bmatrix} \sin(J\beta) \\ -\cos(J\beta) \end{Bmatrix}$	$\begin{Bmatrix} \cos(J\phi) \\ \sin(J\phi) \end{Bmatrix}$	$\begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$
$i = 2$	$\begin{Bmatrix} \cos(J\phi) \\ \sin(J\phi) \end{Bmatrix} - \mu_J \begin{Bmatrix} \sin(J\beta) \\ -\cos(J\beta) \end{Bmatrix}$	$\begin{Bmatrix} \cos(J\phi) \\ \sin(J\phi) \end{Bmatrix}$	$\begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$
$i = 3$	$\begin{Bmatrix} \cos(J\beta) \\ \sin(J\beta) \end{Bmatrix} + \mu_J \begin{Bmatrix} \sin(J\phi) \\ -\cos(J\phi) \end{Bmatrix}$	$\mu_J \begin{Bmatrix} \sin(J\phi) \\ -\cos(J\phi) \end{Bmatrix}$	$\begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$
$i = 4$	$\begin{Bmatrix} \cos(J\beta) \\ \sin(J\beta) \end{Bmatrix} - \mu_J \begin{Bmatrix} \sin(J\phi) \\ -\cos(J\phi) \end{Bmatrix}$	$-\mu_J \begin{Bmatrix} \sin(J\phi) \\ -\cos(J\phi) \end{Bmatrix}$	$\begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$

Also,

$$I_L(E, r_s) = E^{-L/2} \sum_{j=0}^{\infty} E^j g_L^j(r_s), \quad (22)$$

with

$$g_L^j(r_s) = [1/\eta^j j!] f(L - j + \frac{1}{2}, r_s^2 \eta/4). \quad (23)$$

The other quantities are

$$U_L^J(\vec{r}_s, \vec{k}) = \begin{cases} 1 & \text{if } n_s^z = 0, \\ \cos(\pi n_s^z k_z) & \text{if } n_s^z \neq 0 \text{ and } L + J \text{ even,} \\ \sin(\pi n_s^z k_z) & \text{if } n_s^z \neq 0 \text{ and } L + J \text{ odd,} \end{cases} \quad (24)$$

$$Q_L^J(\vec{r}_s) = A_L^J r_s^L P_L^J(\cos \theta_s) \exp(-\eta r_s^2/4) b_L^J(\vec{r}_s), \quad (25)$$

with

$$A_L^J = (-1)^{L+1} 2^{-L} \pi^{-1/2} \eta^{L+1/2} q_L^J, \quad (26)$$

and the functions $b_L^J(\vec{r}_s)$ are given in Table IV.

TABLE IV. Values of the functions $b_L^J(\vec{r}_s)$ used in Eq. (25).

	$n_s^z = 0$	$n_s^z \neq 0$
L even	$(-1)^{L/2}$	$2(-1)^J (-1)^{L/2}$
L odd	$(-1)^{(L+1)/2}$	$2(-1)^{(L+1)/2}$

When one first compares Eq. (7) with Eqs. (20) through (26) and Tables I through IV, it may appear that undue complication has been introduced into the evaluation of the $D_{LM}^{(2)}$. However, some form of detailed partitioning would be introduced in any computer programming for the evaluation of the $D_{LM}^{(2)}$. The explicit form of

Eqs. (20) through (26) is just the partitioning procedure which we find, presently, to be the most efficient in use of available computer time. Also, with $D_{LM}^{(2)}$ written as in (20), the $Q_L^J(\vec{r}_s)$ and the $V_i^M(\vec{r}_s)$ are most logically computed during step (a) of the total calculation, while the $U_L^J(\vec{r}_s, \vec{k})$ and $a_i^J(\vec{r}_s, \vec{k})$ are computed during step (b), as is the i summation in (20). The $I_L^J(E, r_s)$ are computed during step (c), using the $g_L^J(r_s)$ of (23) which are computed and stored during step (a). All trigonometric functions, occurring in the $U_L^J(\vec{r}_s, \vec{k})$ and the $a_i^J(\vec{r}_s, \vec{k})$, are most efficiently computed from just $\cos(\pi k_x)$, $\cos(\pi k_y)$, $\cos(\pi k_z)$, $\sin(\pi k_x)$, $\sin(\pi k_y)$, $\sin(\pi k_z)$, using trigonometric identities for addition of arguments. Finally, the summations of (20) are performed in step (d), as are all the evaluations (5) necessary for a given input value of ℓ_{\max} .

CHOICE OF THE EWALD PARAMETER

It is highly desirable to use the full versatility of the Ewald procedure and choose η in order to minimize the computer time required to evaluate a set of KKR structure constants. The particular choice of η for such minimization could easily depend upon whether the structure constants are being used in band calculations using constant- k searches or using constant- E searches; that is, whether it is necessary to continuously re-execute steps (c) and (d) or to continuously re-execute steps (b) and (d) of the total calculation. Since the computer time required for the evaluation of $D_{OO}^{(3)}$ is insignificant compared with the other time required in the calculation of the KKR structure constants, no further mention will be made here concerning its evaluation.

The computer time required for evaluation of $D_{LM}^{(1)}$ to some desired accuracy is approximately directly proportional to the number of \vec{k}_n 's required to be included in the sum (6). Since the exponential factors in $D_{LM}^{(1)}$ are easily, and quickly, evaluated using Eq. (15), the execution time required for $D_{LM}^{(1)}$ is, to first order, dominated by the evaluation of the Legendre polynomials. The number of Legendre polynomials required for a given \vec{k} point is [<# of \vec{k}_n 's included in sum (6)] $\times (\ell_{\max} + 1)(2\ell_{\max} + 1)$. To reduce this number it is important to choose η in order that the number of \vec{k}_n 's required is not unwieldy. Since the $D_{LM}^{(1)}$ convergence is primarily governed by the exponential factors, this means picking η small compared with those k^2 which are to be neglected. Thus, the smaller η , the fewer computations in steps (b) and (d) to obtain $D_{LM}^{(1)}$ to the desired accuracy; however, as η decreases, the terms required in the $D_{LM}^{(2)}$ sum (20) increase, which increases the computations necessary for the $D_{LM}^{(2)}$ evaluations in steps (b), (c), and (d) of the total calculation. At the same time, it is always desirable to retain enough terms in the $D_{LM}^{(1)}$ sum (6) to ascertain that the final structure constants easily retain their inherent behavior in the vicinity of free-electron singularities, which fall in the ranges of the \vec{k} and the E of interest.

From the above discussion it should be clear that to compute the KKR structure constants to some desired accuracy a balance should be achieved between the number of terms necessarily retained in the sum (6) and the terms necessarily retained in the sum (20). This balance is achieved by the choice of η ; and, ideally, perhaps two optimum choices exist, one for constant- \vec{k} searches and one for constant-E searches. To expedite such choices, test calculations of the type illustrated by Table V are of importance. There values

TABLE V. Deviation of the KKR structure constants as functions of η and arbitrarily terminated lattice sums. Here, m^2 and n^2 are the maximum values of $(\vec{k}_n)^2$ and $(\vec{r}_s)^2/\pi^2$ included in a calculation; rms is defined by Eq. (27), while i_6 , i_7 , and i_{20} are, respectively, the number of terms included in the sums (6), (7), and (20). These results are for the fcc lattice, $\ell_{\max} = 3$, $\vec{k} = (.3,.4,.5)$ and $E = 0.656$.

η	m^2	n^2	i_6	i_7	i_{20}	rms
0.20	4	24	15	248	28	0.0001249
0.20	4	26	15	320	34	0.0000422
0.20	4	30	15	368	37	0.0000067
0.25	4	20	15	200	24	0.0000611
0.25	4	22	15	224	26	0.0000248
0.25	4	30	15	368	37	0.0000002
0.30	4	16	15	140	17	0.0000789
0.30	4	18	15	176	21	0.0000185
0.30	8	24	27	248	28	0.0000006
0.35	4	14	15	134	15	0.0000572
0.35	8	18	27	176	21	0.0000020
0.35	8	22	27	224	26	0.0000001
0.40	8	14	27	134	15	0.0000145
0.40	8	16	27	140	17	0.0000015
0.40	11	18	51	176	21	0.0000005
0.45	11	12	51	86	12	0.0000382
0.45	11	14	51	134	15	0.0000025
0.45	12	16	59	140	17	0.0000003
0.50	11	10	51	78	11	0.0000335
0.50	11	12	51	86	12	0.0000177
0.50	12	14	59	134	15	0.0000006
0.55	12	10	59	78	11	0.0000125
0.55	12	12	59	86	12	0.0000023
0.55	12	14	59	134	15	0.0000001
0.60	12	10	59	78	11	0.0000038
0.60	12	12	59	86	12	0.0000006
0.60	12	14	59	134	15	0.0000000
0.65	12	8	59	54	8	0.0000424
0.65	12	10	59	78	11	0.0000012
0.65	16	12	65	86	12	0.0000002

of the rms deviation of calculated sets of structure constants are presented as functions of η and the number of terms arbitrarily retained in the sums (6) and (20). Since the structure constants for a given \vec{k} and E are not always of the same order of magnitude as L and M vary, for illustrative purposes in Table V we have used the deviation

$$\text{rms} = \{(2\ell_{\max} + 1)^{-2} \sum_L \sum_M [(D_{LM}^t - D_{LM}^c)/D_{LM}^c]^2\}^{1/2}. \quad (27)$$

Here, the D_{LM}^t are the values of $D_{LM}(\vec{k}, E)$ calculated using the parameters of Table V, while the D_{LM}^c are the "converged" values given in Table VI. We also have included Table VI in order to present accurate structure constant values for a general \vec{k} point, since they may be of help to others for program debugging. Of course, a rms entry of Table V for a given η , i_6 , and i_{20} will vary as \vec{k} and E vary; however, a study of results like those of Table V for numerous \vec{k} and E pairs verifies that the rms behavior of Table V is quantitatively representative for the fcc lattice.

The results of Table V nicely illustrate the balance, as η varies, between the number of terms retained in (6) and the number

TABLE VI. Converged values of the KKR structure constants for the fcc lattice, $\ell_{\max} = 3$, $\vec{k} = (.3, .4, .5)$, and $E = 0.656$. The tabulated values are estimated to be accurate to within ± 1 in the last quoted figure, and were calculated using $\eta = 0.50$, 113 \vec{k}_n 's in the sum (6), and 37 reduced r_s 's in the sum (20).

L	M	$D_{LM}(\vec{k}, E)$	L	M	$D_{LM}(\vec{k}, E)$	L	M	$D_{LM}(\vec{k}, E)$
0	0	0.3637069	4	-3	-0.5645986	5	4	-1.5614385
1	-1	0.4729255	4	-2	0.0665038	5	5	0.0352270
1	0	0.5585931	4	-1	0.4518756	6	-6	4.3426328
1	1	0.3855073	4	0	-0.2631646	6	-5	-2.2561097
2	-2	0.1383943	4	1	0.4007033	6	-4	-0.3471268
2	-1	0.3854948	4	2	-0.1747519	6	-3	4.5963681
2	0	0.2294027	4	3	0.1291246	6	-2	-2.3159161
2	1	0.2425004	4	4	0.2032239	6	-1	3.5271685
2	2	-0.1231771	5	-5	0.4958597	6	0	-1.1349680
3	-3	0.3342447	5	-4	-0.0768520	6	1	3.0624581
3	-2	0.7787625	5	-3	0.8811268	6	2	0.1585912
3	-1	0.3760202	5	-2	0.2049815	6	3	-4.1491049
3	0	-0.3433054	5	-1	-0.9030668	6	4	0.9372909
3	1	0.3140171	5	0	0.6678767	6	5	-1.4263739
3	2	-0.0644856	5	1	-0.6314083	6	6	0.3302557
3	3	-0.3585324	5	2	-0.4744635			
4	-4	-0.1350521	5	3	-0.8026519			

of terms retained in (20) in order to achieve some desired overall accuracy in a set of KKR structure constants. To use that information to pick the most favorable value(s) for η , it is also important to possess information concerning computer execution times for the various calculational steps. Such timing information is illustrated by the entries of Table VII. Although some minor variations

TABLE VII. Typical IBM System/360, Model 91, execution times (in 10^{-2} sec) for the steps (a), (b), (c), and (d) in the calculation of KKR structure constants for the fcc lattice. The times are tabulated as functions of η , l_{\max} , and i_6 and i_{20} , which are the number of terms retained in the sums (6) and (20). For comparison, given the structure constants, it requires approximately 0.14×10^{-2} sec for $l_{\max} = 2$, or 0.44×10^{-2} sec for $l_{\max} = 3$, to calculate the quantities defined by Eq. (3) and then evaluate the KKR determinant (2).

η	i_6	i_{20}	$l_{\max} = 2$				$l_{\max} = 3$			
			(a)	(b)	(c)	(d)	(a)	(b)	(c)	(d)
0.20	15	28	21.0	1.43	0.15	0.12	25.0	2.55	0.20	0.24
0.20	15	34	23.3	1.58	0.15	0.14	28.1	2.83	0.22	0.25
0.20	15	37	24.5	1.68	0.17	0.15	29.8	2.97	0.22	0.25
0.25	15	24	17.0	1.28	0.13	0.11	20.7	2.33	0.16	0.21
0.25	15	26	18.0	1.38	0.13	0.13	22.0	2.43	0.17	0.25
0.25	15	37	22.5	1.65	0.13	0.13	27.9	3.04	0.20	0.27
0.30	15	17	13.3	1.10	0.12	0.10	16.0	1.97	0.13	0.22
0.30	15	21	14.8	1.22	0.11	0.10	17.8	2.13	0.15	0.23
0.30	27	28	18.6	1.92	0.13	0.17	22.5	3.37	0.15	0.33
0.35	15	15	12.0	1.05	0.10	0.12	14.1	1.85	0.12	0.22
0.35	27	21	14.8	1.68	0.12	0.15	18.0	3.02	0.15	0.29
0.35	27	26	16.6	1.85	0.13	0.17	20.3	3.27	0.16	0.28
0.40	27	15	12.2	1.52	0.10	0.17	14.4	2.65	0.13	0.28
0.40	27	17	12.7	1.59	0.10	0.15	15.5	2.78	0.14	0.30
0.40	51	21	18.7	2.67	0.11	0.24	21.0	4.76	0.14	0.48
0.45	51	12	15.0	2.39	0.10	0.25	16.8	4.21	0.12	0.49
0.45	51	15	15.9	2.48	0.09	0.25	18.1	4.40	0.13	0.47
0.45	59	17	19.0	2.90	0.09	0.28	21.3	5.08	0.14	0.52
0.50	51	11	14.0	2.37	0.08	0.23	15.7	4.16	0.09	0.49
0.50	51	12	14.8	2.41	0.08	0.25	16.8	4.31	0.12	0.47
0.50	59	15	17.8	2.76	0.09	0.28	19.7	5.01	0.12	0.52
0.55	59	11	15.8	2.68	0.08	0.25	17.5	4.83	0.10	0.52
0.55	59	12	16.5	2.73	0.09	0.26	18.1	4.85	0.14	0.53
0.55	59	15	17.5	2.82	0.10	0.29	20.5	5.00	0.11	0.53
0.60	59	11	16.0	2.72	0.09	0.27	17.6	4.78	0.10	0.53
0.60	59	12	16.8	2.77	0.08	0.27	18.0	4.85	0.12	0.50
0.60	59	15	17.0	2.80	0.10	0.25	19.2	4.95	0.12	0.52
0.65	59	8	14.7	2.60	0.07	0.26	15.7	4.61	0.09	0.48
0.65	59	11	15.5	2.67	0.08	0.28	17.2	4.77	0.10	0.50
0.65	65	12	17.9	2.95	0.08	0.29	19.9	5.22	0.10	0.55

may exist in the execution times, depending on E or \vec{k} , results of the type given in Table VII are almost independent of the E or \vec{k} pair used in a given execution.

DISCUSSION

In our opinion, the most striking feature of the results illustrated by Tables V and VII is the absolute shortness of the time entries for steps (c) and (d) of the calculations, even for $k_{\max} = 3$ which is seldom necessary when using the KKR method. It is also seen that the times for execution of steps (c) and (d) are both of the same order as the time necessary to evaluate the KKR determinant from its matrix elements, which verifies our introductory statement that the structure constants are able to be evaluated in times of this order. These short times also mean that constant- \vec{k} searches are very efficiently performed, and band calculations for several thousand \vec{k} points in 1/48-th of the Brillouin zone are possible using relatively modest amounts of computer time. Although the execution times for step (b) are about an order of magnitude longer than those for step (c), the results of Tables V and VII also illustrate that constant- E searches are also completely practicable along 500 to 1000 \vec{k} directions in 1/48-th of the Brillouin zone.

The precise choices of η , i_6 , and i_{20} of Tables V and VII to be used in band structural calculations are somewhat a matter of personal preference, depending upon the accuracy desired, upon the ranges of \vec{k} and E of interest, etc. However, for most of our actual band structural calculations dealing with the fcc lattice, we find $\eta = 0.30$, $i_6 = 15$, and $i_{20} = 21$ to be quite efficient and also provide more than sufficient accuracy. Although in the interest of brevity we have not presented the results of test calculations pertinent to the bcc lattice, the corresponding values we usually use for those cases are $\eta = 0.25$, $i_6 = 43$, and $i_{20} = 20$, with $m^2 = 6$ and $n^2 = 27$.

REFERENCES

1. J. Korringa, *Physica* 13, 392 (1947).
2. W. Kohn and N. Rostoker, *Phys. Rev.* 94, 1111 (1954).
3. B. Segall and F. S. Ham, General Electric Res. Lab. Rept. No. 61-RL-(2876G), unpublished.
4. J. S. Faulkner, H. L. Davis, and H. W. Joy, *Phys. Rev.* 161, 656 (1967).
5. H. L. Davis, J. S. Faulkner, and H. W. Joy, *Phys. Rev.* 167, 601 (1968).

6. H. L. Davis, Physics Letters 28A, 85 (1968).
7. R. W. Williams and H. L. Davis, Physics Letters 28A, 412 (1968).
8. H. L. Davis, to be published in the proceedings of the International Colloquium on the Physics of Solids under Pressure, held at Grenoble, France, September 8-10, 1969.
9. R. W. Williams and H. L. Davis, to be published in the proceedings of the Electronic Density of States Symposium, held at Washington, D.C., November 3-6, 1969.
10. Y. Onodera and Okazaki, J. Phys. Soc. Japan 21, 1273 (1966).
11. S. Takada, Progr. Theoret. Physics (Kyoto) 36, 224 (1966).
12. J. Treusch, Phys. Status Solidi 19, 603 (1967).
13. J. S. Faulkner, Physics Letters 31A, 227 (1970).
14. K. H. Johnson, article elsewhere this volume.
15. B. Segall, article elsewhere this volume.
16. F. S. Ham and B. Segall, Phys. Rev. 124, 1786 (1961).
17. B. Segall and F. S. Ham, Methods Comput. Phys. 8, 251 (1968).
18. B. R. Cooper, E. L. Kreiger, and B. Segall, article elsewhere this volume.
19. P. J. Davis, in "Handbook of Mathematical Physics," edited by M. Abramowitz and I. A. Stegun, National Bureau of Standards AMS 55, Washington, D.C., 1964, p. 253.
20. P. M. Morse, Proc. Natl. Acad. Sci. U.S. 42, 276 (1956).

CALCULATIONS WITH "NON-MUFFIN TIN" POTENTIALS

BY THE GREEN'S FUNCTION METHOD

Benjamin Segall

Case Western Reserve University

ABSTRACT

A discussion is given of a procedure within the framework of the Green's function (or KKR) method for calculating the perturbations of the energy bands due to deviations from a "muffin-tin" potential. The essential aspect of the approach is the determination of accurate wave functions in practical forms in the regions inside and outside the "muffin-tin" spheres. The forms of the perturbing potential in these regions are also discussed. This procedure was suggested by F. S. Ham and the author who applied it (in low order perturbation theory) to the case of the Mathieu potential.

I. INTRODUCTION

In the vast majority of band calculations of metals to date, the crystal potential has been approximated by a potential of the well-known "muffin-tin" form.¹ This form of the potential is essential to the simple, or standard, application of the Green's function method (GFM) - often referred to as the KKR method - the method with which we will be concerned here. While the "muffin-tin" potential provides a fairly good approximation to the potential for metals, particularly for crystal structures with high symmetry and one atom per unit cell, it would be expected to be inaccurate for highly covalent semiconductors. The covalent materials have characteristically directed bonds and only a relatively small fraction of the cell volume lies inside the muffin-tin spheres. In these crystals a generalization of the standard approach would clearly be necessary. Furthermore, following the success of the band calculations for metals over the past decade there has been an

increasing interest in improving the accuracy of these calculations by attempting to take into account the relatively small shifts in the bands due to deviations of the crystal potential from the approximate muffin-tin form.²

Although it has been stated on a number of occasions that the GFM cannot handle such general potentials, a definite procedure to do so was discussed and applied by F. S. Ham and the author³ almost a decade ago. As no other published work with this method has appeared and since many who work in the field may not be familiar with it, I thought it would be appropriate to discuss it at this conference. I also wish to note the existence of an alternative approach for coping with the more general potentials by the GFM recently proposed by A. R. Williams⁴ which differs significantly from the one to be presented here.

II. WAVE FUNCTIONS

Basically, the method is straightforward: it consists essentially of obtaining reasonably accurate wave functions for the muffin-tin potential throughout the unit cell in a convenient form and using them to set up the (secular) perturbation problem. As is well known, the wave function inside the sphere (i.e. $r < r_i$) is expanded in spherical harmonics (or lattice harmonics for \vec{k} -values of special symmetry)

$$\psi(\vec{r}) = \sum_{\ell,m} i^\ell C_{\ell m} R_\ell(r) Y_{\ell m}(\vec{r}), \quad (1)$$

where the $Y_{\ell m}$ are the usual spherical harmonics and $R_\ell(r)$ the solutions of the radial Schrödinger equation at the energy eigenvalue. The wave function $\psi(\vec{r})$ satisfies the integral equation

$$\psi(\vec{r}) = \int_{\tau} G(\vec{r}, \vec{r}') V(\vec{r}') \psi(\vec{r}') d\tau', \quad (2)$$

where $G(\vec{r}, \vec{r}')$ is the Green's function which satisfies the Bloch condition. For a muffin-tin potential, only the region $r < r_i$ contributes to the integral (since V can be taken to be zero for $r > r_i$). Eq.(2), or the equivalent variational principle, for \vec{r} inside the sphere leads to a system of linear homogeneous equations for the $C_{\ell m}$ which have non-trivial solutions only when the determinant of the matrix of the coefficients of the $C_{\ell m}$ vanishes. This determinantal equation yields the energy eigenvalue for the given \vec{k} (or vice versa). At this energy, the system of homogeneous equations yields the $C_{\ell m}$, which by (1) gives the ψ for $r < r_i$. This is all quite standard.

Since the integral in (2) is restricted to the region inside the sphere, the integral equation for $r > r_i$ as it stands directly

provides the wave function in that region in terms of the already determined function for $r < r_i$. However, considerable simplification can be effected by eliminating the $V(\vec{r})$ through the Schroedinger equation and by using Green's theorem and the properties of the Green's function to obtain ψ as a surface integral over the sphere

$$\psi(\vec{r}) = \int_{r'=r_i} \hat{n} \cdot [G(\vec{r}, \vec{r}') \vec{\nabla}' \psi(\vec{r}')] - \psi(\vec{r}') \vec{\nabla}' G(\vec{r}, \vec{r}')] dS'. \quad (3)$$

Since $\vec{r} \neq \vec{r}'$ ($r > r_i \geq r'$) the $G(\vec{r}, \vec{r}')$ is nonsingular in the domain of integration and its expansion in spherical harmonics in the angles of \vec{r}' is

$$G(\vec{r}, \vec{r}') = \sum_{LM} (-i)^L D_{LM}(\vec{r}) j_L(\kappa r') Y_{LM}(\vec{r}'), \quad (4)$$

where as usual $\kappa = \sqrt{E}$ or $i\sqrt{-E}$ for $E > 0$ or $E < 0$. With (4) and (1) we readily obtain from (3)

$$\psi(\vec{r}) = r_i^2 \sum_{\ell m} C_{\ell m} D_{\ell m}(\vec{r}) [j_\ell(\kappa r_i), R_\ell(r_i)] \quad (5)$$

where the bracket of two functions in general denotes

$$[F, G] = [F(r) dG(r)/dr - G(r) dF(r)/dr]_{r=r_i},$$

which implies that the bracket in (5) is a quantity proportional to $\tan \eta_\ell$, η_ℓ being the phase shift for angular momentum ℓ .

In (5) we have ψ in terms of the wave function inside (i.e. the $C_{\ell m}$), some known quantities and the function $D_{LM}(\vec{r})$. By procedures completely similar to those used in obtaining the "structure constants" we find

$$D_{LM}(\vec{r}) = -\frac{4\pi}{\tau} \frac{1}{j_L(\kappa s)} \sum_n \frac{\exp[i(\vec{k}_n + \vec{k}) \cdot \vec{r}]}{(\vec{k}_n + \vec{k})^2 - E} \\ \times j_L(|\vec{k}_n + \vec{k}| s) Y_{LM}(\vec{k}_n + \vec{k}). \quad (6)$$

The expression (6) is formally identical to the "off-diagonal" structure constants for a lattice with a basis.⁵ To the extent that the spherical harmonic expansion in (1) has converged, the $\psi(\vec{r})$ for $r > r_i$ given by (5) and (6) is exact.

Two features of (6) are noteworthy. First is the fact that (6) and thus ψ for $r > r_i$ is an expansion in plane waves. This is

just what one would desire. For it is to be expected that the wave function in the region exterior to the spheres will be rather smooth and as a consequence that it could be represented by a few plane waves. Secondly we note that while the $\mathcal{D}_{LM}(\vec{r})$ are independent of the arbitrary parameter s appearing in (6), the expansion coefficients of the plane waves are dependent on them. The expansion is thus not unique and, in fact, there is an infinite number of different expansions. The lack of uniqueness results from the fact that we are expanding a given function $[\psi(\vec{r})]$ only over a part of the unit cell in terms of a set of functions (the plane waves) which is complete over the whole cell and thus over complete in the region outside the spheres. While at first sight this non-uniqueness might appear to be a considerable inconvenience, it can, in fact, be utilized to advantage. What is desired, of course, is an approximation in a small number of plane waves

$$\psi(\vec{r}) \approx \phi_N(\vec{r}) = \sum_{n=1}^N F_n \exp[i(\vec{k}_n + \vec{k}) \cdot \vec{r}]. \quad (7)$$

We thus require the "best" approximation to the exact $\psi(r)$ in the sense that

$$\int_{r > r_i} |\psi(\vec{r}) - \phi_N(\vec{r})|^2 d\tau \quad (8)$$

be a minimum with respect to variations of the F_n . Substituting (5) and (6) for ψ and (7) for ϕ_N we obtain the following equations for the F_n

$$\sum_{n=1}^N F_n \mu_{mn} = - \frac{4\pi r_i^2}{\tau} \sum_{\ell m} \frac{C_{\ell m}}{j_\ell(\kappa s)} [j_\ell(\kappa \vec{r}_i), R_\ell(r_i)] \\ \times \sum_{n=1}^{\infty} \frac{\mu_{mn}}{(\vec{k}_n + \vec{k})^2 - E} j_\ell(|\vec{k}_n + \vec{k}| s) Y_{\ell m}(\vec{k}_n + \vec{k}),$$

for $1 \leq m \leq N$, where

$$\mu_{mn} = \int_{r > r_i} \exp[-i(\vec{k}_m - \vec{k}_n) \cdot \vec{r}] d\tau \\ = \tau \delta_{mn} - (4\pi r_i^2 / |\vec{k}_m - \vec{k}_n|) j_1(|\vec{k}_m - \vec{k}_n| r_i). \quad (9)$$

This set of inhomogeneous equations determines the F_n 's in terms of the $C_{\ell m}$. The reciprocal space sums on the right-hand side of (9) are similar to those in the structure constants aside from the μ_{mn} factor (which, incidentally, improves the convergence).

Aside from the references already cited^{2,3} the only other application of these procedures was by H. L. Davis,⁶ who calculated ψ for $r > r_i$ for Cu directly by (5) and (6) and also by (7) and (9). He found good agreement between both results.

III. POTENTIAL

It is convenient, particularly for $r < r_i$, to expand the potential about the nuclei in the lattice harmonics belonging to the symmetric representation. For the case of cubic crystals with one atom per cell we have

$$V(r) = V_0(r) + V_4(r)L_4(\vec{r}) + V_6(r)L_6(\vec{r}) + \dots, \quad (10)$$

where $V_0(r)$ is the muffin-tin potential V_{mt} for $r < r_i$. For $r > r_i$, V_{mt} is taken to be the constant average of the potential in that region.

The deviation of $V(\vec{r})$ from $V_{mt}(\vec{r})$, $\delta V = V(\vec{r}) - V_{mt}(\vec{r})$, also can be expanded in a Fourier series throughout the cell

$$\delta V(\vec{r}) = \sum_n \delta V_n \exp(i\vec{k}_n \cdot \vec{r}). \quad (11)$$

But, just as for $\psi(\vec{r})$, a plane wave expansion restricted to $r > r_i$ is not unique. Again we choose the "best" approximation to $V(\vec{r})$ in terms of the N plane waves

$$\delta V(\vec{r}) \approx W_N(\vec{r}) = \sum_{n=1}^N W_n \exp(i\vec{k}_n \cdot \vec{r}), \quad (12)$$

which minimize

$$\int_{r > r_i} |\delta V(\vec{r}) - W_N(\vec{r})|^2 d\tau. \quad (13)$$

The W_n are found to satisfy

$$\begin{aligned} \sum_{n=1}^N W_n \mu_{mn} &= \tau \delta V_m - 4\pi \sum_{\ell=4,6,\dots} (-)^{\ell/2} L_\ell(\vec{k}_m) \\ &\times \int_0^{r_i} r^2 V_\ell(r) j_\ell(k_m r) dr \end{aligned} \quad (14)$$

for the cubic crystals with one atom per cell. The generalizations to other cases are straightforward.

With the $\psi(\vec{r})$ and $V(\vec{r})$ in the form (1) and (10) for $r < r_i$ and in the form (7) and (12) for $r > r_i$, matrix elements can readily be

evaluated. It is noteworthy that with the representation (7) the contribution to the matrix elements from $r < r_i$ are simply the products of coefficients times the μ_{mn} . In other forms of ψ in the exterior region, three-dimensional integrations over the complicated exterior regions are required. Having the matrix elements one can set up the standard secular perturbation problem to evaluate the perturbations from the eigenvalues for the muffin-tin potential and to determine the perturbed wave functions if desired. While we will not elaborate on implementation of the secular problem, we will suggest that it is practical. In the problems that one would be concerned with, the perturbations of the eigenvalues will typically be under 1 eV. It would thus suffice to include in the approximate basis set only those states with the same symmetry within about 1 Ry of the level being studied. This should involve only a small number of states, particularly at symmetry points.

To our knowledge, the only attempt to test the proposed perturbation approach was in the work of Ham and the author³ where the three-dimensional Mathieu potential,

$$V(\vec{r}) = U_1 + U_2 [\cos(2\pi x/a) + \cos(2\pi y/a) + \cos(2\pi z/a)], \quad (15)$$

with $U_2 = -\frac{1}{2}(2\pi/a)^2$

was studied. We note that the Mathieu potential is the only non-trivial three-dimensional potential for which exact solutions are available. In addition, as it can be approximated only quite poorly by a potential of the muffin-tin form, it presents a severe test of the Green's function method. In Table I results of the perturbation calculation for the lowest Γ_1 state are presented in the dimensionless energy parameter ϵ defined by $E = (2\pi/a)^2 \epsilon$. It is seen that the muffin-tin result differs from the exact result by about 4%. Since

Table I. Calculated perturbation of the lowest Γ_1 level for the three dimensional Mathieu potential discussed in the text.

$$\underline{\epsilon(\Gamma_1) = (a/2\pi)^2 E(\Gamma_1)}$$

Exact	- 0.843
Muffin-tin value	- 0.810
Difference	- 0.033
$\epsilon[\Gamma_1(2)] - \epsilon[\Gamma_1(1)]$	1.207
Perturbations	
$\Delta\epsilon^{(1)}$	- 0.025
$\Delta\epsilon^{(2)} \text{ (Eq. (16))}$	- 0.01

we expect the contribution to the shifts from the region $r > r_i$ to be considerably larger than from $r < r_i$, we considered only the former. The normalized wave function (7) is given by $a^{2/3} F_n = 0.302, 0.048$ and 0.007 for the reciprocal vectors $\vec{K}_n = 0, (2\pi/a, 0, 0)$ and $(2\pi/a, 2\pi/a, 0)$ respectively. The first-order shift $\Delta\epsilon(1)$ calculated from these functions was about 75% of the difference from the exact result. Inasmuch as the first-order perturbation is expected to be small by virtue of the manner in which the constant part of V_{mt} is chosen (i.e. $\int \delta V d^3r = 0$) we estimated the second-order term. Taking the separation of the first two Γ_1 levels as the average excitation energy, $\langle E_n - E_o \rangle$, and summing the excited states (by closure) we obtain for the estimate of $\Delta\epsilon(2)$

$$(2\pi/a)^2 \Delta\epsilon(2) = \{ \langle 0 | (\delta V)^2 | 0 \rangle - (\langle 0 | \delta V | 0 \rangle)^2 \} \langle E_o - E_n \rangle^{-1} \quad (16)$$

Of course, the secular perturbation procedure outlined above would be more satisfactory than this cruder (but quicker) estimate. It is evident, however, that $\Delta\epsilon(1) + \Delta\epsilon(2)$ account for almost all of the difference from the exact result. We note that some perturbation calculations were carried out for Al⁽²⁾, but the shifts were generally small and there were no exact results available for comparison.

In conclusion, I would like to suggest that while the above description of the approach might appear a little involved, it is practical. The additional programming required over and above that for the usual operation of GFM is modest. A more extended application of these techniques would, of course, be helpful in assessing their practicability. Drs. G. Juras and C. B. Sommers and I are presently in the process of carrying out such tests.

References

1. In this approximate form the potential is spherically symmetric inside non-overlapping spheres centered on the nuclei and constant elsewhere.
2. An early estimation of these perturbations in Al were given in B. Segall, Phys. Rev. 124, 1797 (1961).
3. F. S. Ham and B. Segall, Phys. Rev. 124, 1786 (1961).
4. A. R. Williams, Phys. Rev. (to be published).
5. B. Segall, Phys. Rev. 105, 108 (1957).
6. H. L. Davis, unpublished.

PHASE SHIFT PARAMETRIZATION: BAND STRUCTURE OF SILVER

B. R. Cooper and E. L. Kreiger
General Electric Research and Development Center
and

B. Segall
Case Western Reserve University

Abstract

We discuss a band parametrization scheme (within the KKR framework) specifying the phase shifts η_0 , η_1 , and η_2 as functions of energy. Such an approach is particularly useful for the noble and transition metals where both d-band and free-electron-like effects are important. The $\eta_\ell(E)$ for a family of elements are expected to have characteristic energy dependences, with each $\eta_\ell(E)$ being specified over a substantial energy range by a few parameters. First, we show the existence of such characteristic behavior for the noble metals. We then use our phase shift parametrization scheme in a semi-empirical way to find the band structure of Ag. To do this, we use a first principles calculation as a guide, and adjust the parameters specifying the $\eta_\ell(E)$ to fit available Fermi surface, optical and photoemission data.

I. Introduction and Development of the Phase Shift Parametrization.

We have developed an energy band parametrization scheme within the framework of the Green's function (KKR) method.² In this method, the energy eigenvalues are the solutions of the determinental equation

$$\det \{B_{\ell j; \ell' j}, t \alpha \eta_\ell, +\kappa \delta_{\ell \ell}, \delta_{jj}, \} = 0 \quad (1)$$

where the phase shifts η_ℓ are given by

$$\tan \eta_\ell = \frac{[R_\ell(r), j_\ell(\kappa r)]_{r=r_i}}{[R_\ell(r), n_\ell(\kappa r)]_{r=r_i}} \quad (2)$$

with

$$[R_\ell(r), j_\ell(\kappa r)]_{r=r_i} = R_\ell(r_i) \left. \frac{dj_\ell(\kappa r)}{dr} \right|_{r=r_i} - \left. \frac{dR_\ell(r)}{dr} \right|_{r=r_i} j_\ell(\kappa r_i) \quad (3)$$

Here r_i denotes the radius of the inscribed sphere of a muffin-tin potential; $R_\ell(r)$ is the radial wave function; and $j_\ell(\kappa r)$ and $n_\ell(\kappa r)$ are the spherical Bessel and Neumann functions. The energy is given in dimensionless form by ϵ , where

$$E = \left(\frac{2\pi}{a} \right)^2 \epsilon + V_0 \quad (4)$$

So the zero of ϵ corresponds to the constant potential between spheres of the muffin-tin potential. For positive ϵ the quantity κ is proportional to the square root of ϵ ,

$$\kappa = \frac{2\pi}{a} \sqrt{\epsilon} \quad (5)$$

and is appropriately modified² for negative ϵ .

Our parametrization scheme is based on two facts. First is that in the KKR method the effects of the crystal potential in determining the energy eigenvalues (which satisfy the determinental equation, Eq. (1)) are entirely embodied in the phase shifts, η_ℓ . The effects of the potential are completely separated from crystal structure effects given by the (ϵ and k -dependent) coefficients $B_{\ell,j;\ell',j'}$ in (1). The phase shifts are, of course, functions of the energy. Second, because of the very rapid convergence of the KKR method, it suffices - except perhaps for very heavy metals - to include only the $\ell = 0, 1$, and 2 angular momentum components in order to achieve high accuracy in the energy eigenvalues for varying wave vector. Thus we need include only three phase shifts in a nonrelativistic calculation. Such a parametrization scheme is particularly promising for the noble and transition metals where both d-band and free-electron-like effects are important. Use of such a scheme depends on the energy-dependent phase shifts for

a family of elements having characteristic energy dependences so each $\tan \eta_\ell$ can be specified with a few parameters. As discussed below, we have used our scheme in a semiempirical way to find the band structure of Ag. To do this, we use a first principles calculation as a guide, and adjust the parameters specifying the $\tan \eta_\ell(\epsilon)$ to fit Fermi surface, optical, and photoemission data. One of the great virtues of this parametrization scheme is the flexibility one has in incorporating both the results of first principles calculations and of experiment.

The key step in such a procedure is to arrive at functional forms for the energy dependences of the $\tan \eta_\ell$ for a family of elements, so that each $\tan \eta_\ell$ can be specified over a wide range of energy in terms of a few parameters. Actually, one is more restricted in the choice of these forms than one might expect at first thought. We arrived at the forms used in the following way. First, one can get an idea of the small ϵ behavior by taking the leading terms in the series expansion of the Bessel functions in Eq. (2). For the noble and transition metals there is further guidance in that we expect a pole in $\tan \eta_2$, corresponding to a central location in the d-bands, the so-called d-band resonance. Also, the behavior of certain of the structure coefficients for the face centered cubic lattice dictates certain restrictions in the higher order terms in ϵ in order to have physically reasonable behavior for the $\tan \eta_\ell$. To complete the specification of the functional form of the $\tan \eta_\ell$, we adopted the empirical procedure of examining the results of first principles calculations for two different potentials for copper and two different potentials for silver. We then used as a rule of thumb the criterion of selecting the simplest form for the $\tan \eta_\ell$ that would fit all of the first principles energies across a range of a Rydberg or more, with a maximum error of 0.01 Ry for any single energy. Doing this brought us to the following forms.

$$\tan \eta_0 = s_1 |\epsilon|^{1/2} (1-s_2 \epsilon)/(1+s_3 \epsilon) + s_4 (1-\sigma(\epsilon)) |\epsilon|^{3/2} \quad (6)$$

$$\tan \eta_1 = p_1 |\epsilon|^{1/2} (1-p_2 \epsilon) \epsilon \quad (7)$$

$$\tan \eta_2 = d_2 |\epsilon|^{5/2}/(\epsilon-d_1) + \sigma(\epsilon-d_1) [d_3/(\epsilon-d_1)]^{1/2} + d_4 \epsilon^2 \quad (8)$$

where $\sigma(x) = 0$ for $x < 0$ and =1 for $x > 0$.

(A similar parametrization applies to a free-electron-like metal such as Al except, of course, that there is no resonant character for $\tan \eta_2$). Actually, for the simple forms shown here we did not quite achieve the goal of no departure from the first principles calculation greater than 0.01 Ry. Typically, for a few energies at the top of the band structure, roughly a Rydberg above Γ_1 , the error somewhat exceeded this. However, the root-mean-square deviation was very much smaller than 0.01 Ry.

Thus for the noble metals, the phase shift parametrization calls for 10 parameters: 4 for $\tan \eta_0$, 2 for $\tan \eta_1$, and 4 for $\tan \eta_2$.

For a typical case, the "Chodorow" potential for Cu, Fig. 1 shows a comparison between the first principles³ evaluation of the phase shifts and the fitting by the phase shift parametrization. The parametrized fitting almost exactly coincides with the first principles phase shifts except for the few places indicated by a dotted curve. A comparison of the

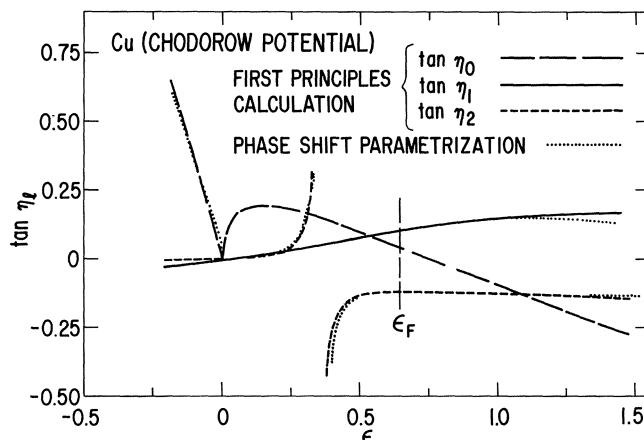


Fig. 1. Comparison of $\tan \eta_i$'s for Cu Chodorow potential from first principles and phase shift parametrization calculations. (In regions where the phase shift parametrization results are not shown, the results of the two calculations essentially coincide). The vertical dashed line labeled ϵ_F indicates the Fermi energy.

symmetry point energy eigenvalues calculated directly³ from the "Chodorow" potential and from the parametrization scheme is made in Table 1. These are typical of the results for the other potentials investigated and for the non-symmetry points.

II. Application to Silver.

The early band calculations for Ag done by Segall⁴ showed that the principal difficulty in obtaining satisfactory results was in properly positioning the d-bands with respect to the more free-electron like conduction bands. There have been suggestions for shifting the d-bands upward, from the value found for Segall's calculation (using a potential determined from Hartree-Fock free ion wave functions) in some fairly crude way in order to get agreement with experimental optical or photoemission data. The phase shift parametrization technique offers a particularly elegant and physically reasonable way of making this sort of correction to a first principles calculation by use of empirical data. Indeed, using the phase shift parametrization technique, one can blend a good deal of experimental information into the framework of a first principles calculation. Table 2 gives an outline of the way we have done this for Ag using Segall's calculations (for a potential determined from Hartree-Fock free ion wave functions) as a starting point. (In Table 2, "Hartree-Fock" or "H.F." is used to denote the results of Segall's calculation).

Since there are a substantial number of parameters, the values of which are to be determined, there is a correspondingly substantial number of steps and pieces of data to be used in the fitting. If one were willing to accept a band structure with less overall accuracy, one could use simpler forms involving fewer parameters for the $\tan \eta_\ell$, and hence need fewer data and steps in the process of evaluating the parameters.

We will present a detailed discussion of the parametrization procedure in a future publication. Table 2 is meant to show the combination of types of experimental data and first principles calculational results used in the procedure. The labels η_ℓ at the end of most lines indicates the phase shift about which that step provides information. Steps 1 and 2 set the Fermi energy relative to the conduction bands. Step 3 then sets the d-band energies relative to the conduction bands. Step 4 feeds

Table 1. Energies for Cu Chodorow Potential from First Principles³ and Phase Shift Parametrization Calculations.

State	$\stackrel{\rightarrow}{E}(k)$ (Ry) first principles	$\stackrel{\rightarrow}{E}(k)$ (Ry) phase shift param.
Γ_1	-1.043	-1.043
Γ_{12}	-0.584	-0.577
Γ_{25}	-0.644	-0.638
L_1	-0.778	-0.787
L_3	-0.648	-0.644
L_3'	-0.539	-0.534
L_2'	-0.422	-0.421
L_1	-0.081	-0.076
X_1	-0.781	-0.791
X_3	-0.745	-0.754
X_2	-0.541	-0.538
X_5	-0.526	-0.525
X_4'	-0.224	-0.218
X_1	0.169	0.173

Table 2. Phase Shift Parametrization Scheme for Silver.

1. $E(L_2')$ from Segall's Hartree-Fock⁴, η_1
2. Set E_F , optical⁵ & photoemission,⁶
 $E_F = E(L_2') + 0.3 \text{ eV}$
3. $E(L_{3\text{upper}}) = E_F - 3.97 \text{ eV}$, optical,⁵ η_2
4. $E(X_3) = E(L_{3\text{upper}}) - 3.5 \text{ eV}$, photoemission,⁷ η_2
5. $E(\Gamma_{25'}) - E(X_3)$ from Ag H.F.⁴ $\rightarrow E(\Gamma_{25'})$, η_2
6. $E(X_4)$, optical⁵, $\eta_1 \rightarrow \eta_1$ completely determined

7. Belly radii⁸ $\rightarrow \eta_0$, η_2 at $E_F \rightarrow \eta_2$ completely determined
8. $E(\Gamma_1)$ from Ag H.F.⁴, η_0
9. $E(L_{1\text{upper}})$, photoemission,⁶ η_0
10. $E(W_{1\text{lower}})$, relative position in d-bands,⁴
 $\eta_0 \rightarrow \eta_0$ completely determined
11. CHECKS, neck radius⁸, $E(X_4) - E(X_1)$ optical⁹,
self-consistency of E_F .

in experimental information about the width of the d-band complex. Step 5 uses some of the first principles calculation results to give some detail of the d-band structure. Step 6 uses some of the optical data to evaluate $\tan \eta_1$ at a second energy. Since $\tan \eta_1$ depends on only two parameters, this means that the energy-dependent behavior of $\tan \eta_1$ is completely determined at this point in the procedure.¹ In step 7, the belly radii anisotropy is used to determine $\tan \eta_0$ and $\tan \eta_2$ at the Fermi energy. This then gives $\tan \eta_0$ at four energies, so the four parameters in $\tan \eta_2$ are completely determined. In steps 8, 9, and 10 a combination of experimental and first principles calculational results give the further information necessary to fully determine $\tan \eta_0$. Finally, as indicated in line 11, as checks on the physical reasonableness of our results, we have used comparison with the experimental values of the neck radius and of the $E(X_4) - E(X_1)$ energy difference. In both cases the agreement is within one or two per cent. Also we find that the Fermi energy for the resulting band structure agrees with that used in the course of the calculation within the accuracy of our computational method.

Fig. 2 shows the resulting phase-shift behavior, and Fig. 3 shows the band structure that results from

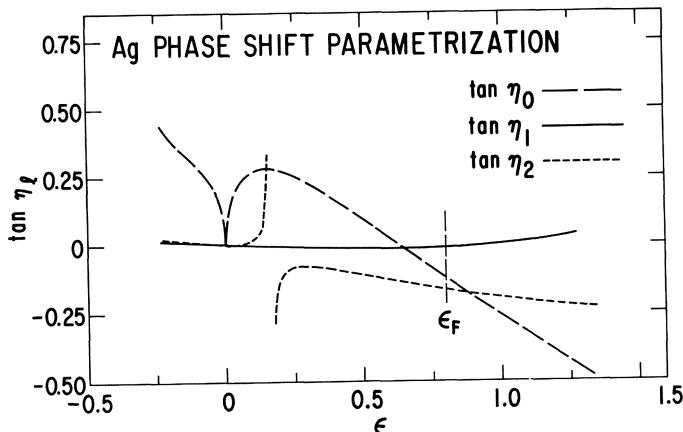


Fig. 2. Values of $\tan \eta_l$ for Ag obtained by the phase shift parametrization scheme outlined in Table 2. The vertical dashed line labeled ϵ_F indicates the Fermi energy.

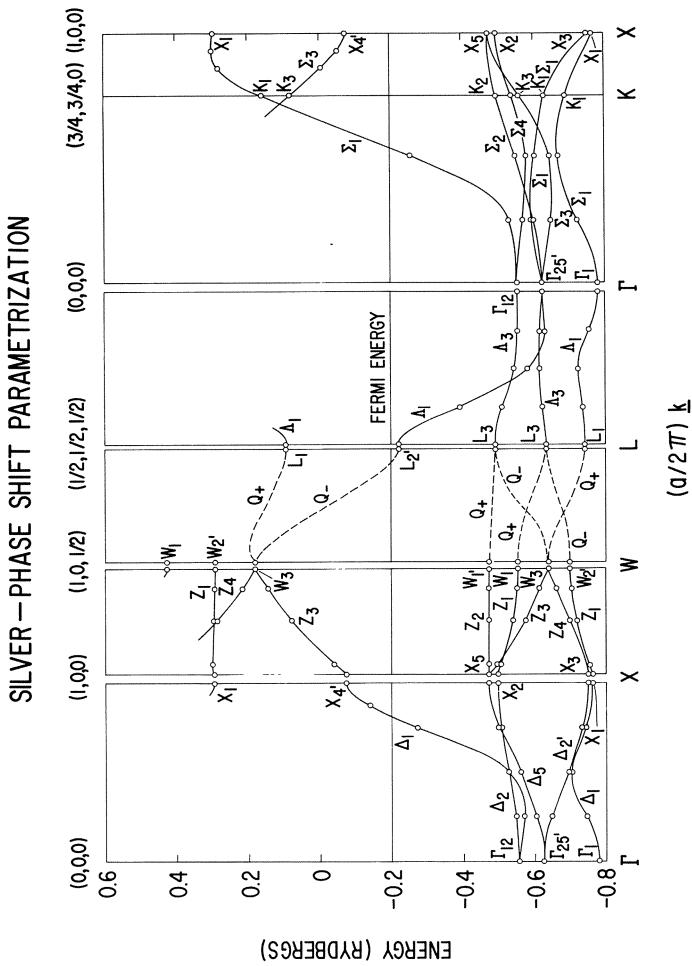


Fig. 3. The calculated energy bands for Ag along the various symmetry axes in the Brillouin zone and on the zone surface. These have been obtained by the phase shift parametrization scheme outlined in Table 2 and were calculated using the phase shifts of Fig. 2.

use of the parametrized phase shifts. Among other things the present band structure has shifted the energy difference between the top of the d-bands and the conduction electrons at the Fermi energy to the experimentally indicated value. Also, the band structure of Fig. 3 is in good agreement with many of the most important features of Snow's¹⁰ self-consistent calculation using 5/6 of the Slater exchange term. However, since we have fed in some of the experimental data, our calculation has a greater built-in agreement with some experimental features.

In conclusion, we wish to note that the knowledge about the phase shifts associated with the various atomic species, as gained in the application of the proposed scheme, will be of importance in itself. For it seems probable that these quantities will play an essential role in future realistic treatments of a number of important problems in the solid state such as itinerant electron magnetism and the electronic structure of disordered systems.

References

1. B. R. Cooper, E. L. Kreiger, and B. Segall, Physics Letters 30A, 333 (1969); The idea of developing such a scheme had its genesis in a suggestion made by B. Segall and F. S. Ham in Methods in Computational Physics, vol. 8 Energy Bands of Solids, edited by B. Alder, S. Fernbach, and M. Rotenberg (Academic Press, New York, 1968) Chapter 7.
2. For a review of the KKR method see the article by Segall and Ham cited above. We adopt the notation of that article.
3. B. Segall, Phys. Rev. 125, 109 (1962).
4. B. Segall, General Electric Research Laboratory Report No. 61-RL-2785 G, 1961 (unpublished).
5. B. R. Cooper, H. Ehrenreich, and H. R. Philipp, Phys. Rev. 138, A494 (1965).
6. C. N. Berglund and W. E. Spicer, Phys. Rev. 136, A1044 (1964).
7. D. E. Eastman and J. K. Cashion, Phys. Rev. Letters 24, 310 (1970).

8. D. J. Roaf, Phil. Trans. Roy. Soc. (London) 255, 135 (1962).
9. H. Ehrenreich and H. R. Philipp, Phys. Rev. 128, 1622 (1962).
10. E. C. Snow, Phys. Rev. 172, 708 (1968).

BAND STRUCTURE CALCULATIONS FOR SEMICONDUCTORS AND
INSULATORS USING THE KKR METHOD

H. Overhof

Institut für Theoretische Physik (II) der
Universität Marburg, Germany

We have used the KKR method to calculate energy bands for some twenty semiconductors and insulators of various crystal structure with good success. In doing this we always used the simple muffin-tin model with spherical symmetric potentials in nonoverlapping spheres and constant elsewhere. The constant potential value V_0 was taken as a fitting parameter to adjust the fundamental band gap.

No attempt was made to calculate matrix elements. But it seemed often instructive to have eigenfunctions. It is straightforward to calculate and normalize eigenvectors in the muffin-tin spheres. The contribution of the constant potential region to the normalization integral can be taken easily from the V_0 -dependence of the eigenvalues by a first order perturbation calculation.

In the course of our applications of the KKR method we developed improvements concerning the calculation of the structure constants and the inclusion of relativistic effects, on which I report here.

Calculation of the Structure Constants

In order to set up the KKR determinant one must calculate the energy and \vec{k} -dependent structure constants. These can be written [1,6] as

$$D_{Ln}^{ii'}(\vec{k}, E) = D_{Ln}^{ii'}(1) + D_{Ln}^{ii'}(2) + \delta_{L0} \delta_{ii'} D_{00}^{(3)}(1)$$

with

$$D_{L\eta}^{ii''}(1) = -\frac{4\pi i^L e^{\eta}}{\Omega |E|^{L/2}} \sum_{\vec{k}_n} \frac{e^{i\vec{k}_n(\vec{a}_j - \vec{a}_{j'})}}{\vec{k}_n^L - E} k_n^L Y_{L\eta}(\vec{k}_n) e^{-\frac{k_n^2}{\ell^2}} \quad (2a)$$

$$D_{L\eta}^{ii''}(2) = -\frac{2^{L+1}}{\sqrt{\pi} |E|^{L/2}} \sum_{\substack{\vec{k} \neq 0 \\ \vec{s}}} e^{i\vec{k}\cdot\vec{R}_s} |\vec{R}_s - \vec{a}_j - \vec{a}_{j'}|^L Y_{L\eta}(\vec{R}_s - \vec{a}_j - \vec{a}_{j'}) \cdot \int_{\vec{R}/2}^{\infty} e^{-(\vec{R}_s - \vec{a}_j - \vec{a}_{j'})^2 \xi^2 + \frac{E}{4\xi^2}} \xi^{2L} d\xi \quad (2b)$$

$$D_{oo}(3) = -\frac{\sqrt{\rho}}{2\pi} \sum_{n=0}^{\infty} \frac{(E/\eta)^n}{n!(2n-1)} \quad (2c)$$

Here we have used the following notation

\vec{a}_j = position vector of the atom j in the unit cell

η = Ewald parameter

$\vec{k}_n = \vec{k} + \vec{R}_n$, \vec{R}_n = reciprocal lattice vector

\vec{R}_s = lattice vector in real space

Ω = volume of the unit cell

The last term, $D(3)$, is very quickly convergent and can be evaluated easily. The convergence of the other two terms depends strongly on the choice of the Ewald parameter η . There are two ways to proceed: either η is chosen small and, therefore, $D(1)$ is quickly convergent but $D(2)$ must be considered or η is chosen large enough to fulfill the condition

$$D(2) < 10^{-5} D(1) \quad (3)$$

thus $D(2)$ can be neglected. The first method is that of Davis, while we used the second way. Both methods will need the same amount of computer time. Since the lattice structure effects the magnitude of $D(2)$ via the lattice vectors \vec{R}_s , η will be different for different lattices. Especially, for $j \neq j'$ the $D(2)$ contribution depends on the nearest neighbour distance, which will be smaller than the smallest \vec{R}_s and, therefore, η must be larger for $j \neq j'$ than for $j=j'$ in order to satisfy (3). The calculation of $D(1)$ then requires a summation over some thousand points in the reciprocal lattice. Since the calculation according to (2a) has to be repeated for each energy, it would take most of the

computer time.

We, therefore, use the following rearrangement of the sum. Expanding the denominator into a power series

$$\begin{aligned}\frac{1}{k_n^2 - E} &= \frac{1}{k_n^2} + \frac{E}{k_n^2(k_n^2 - E)} \\ &= \frac{1}{k_n^2} + \frac{E}{k_n^4} + \frac{E^2}{k_n^6} + \cdots + \frac{E^m}{k_n^{2m}(k_n^2 - E)}\end{aligned}\quad (4a)$$

which can be approximated by

$$\frac{1}{k_n^2 - E} = \frac{1}{k_n^2} + \frac{E}{k_n^4} + \frac{E^2}{k_n^6} + \cdots + \frac{E^m}{k_n^{2m+2}} \quad (4b)$$

if $k_n^2 > k_0^2$ is large compared to $|E|$, we can split the sum in a first part, for which $k_n^2 \leq k_0^2$ and, therefore, the expansion (4b) is not correct, and into a second part, for which $k_n^2 > k_0^2$ and (4b) holds. The first part is evaluated according to (2a), but now the summation runs only over a small sphere in the reciprocal lattice, whereas for the second part the expansion (4b) is used. The result then reads

$$\begin{aligned}D_{nn}^{ii'}(1) &= -\frac{4\pi i^L e^{E/2}}{\Omega |E| k_n^L} \cdot \left[\sum_{k_n^2 \leq k_0^2} \frac{e^{ik_n(\vec{a}_i - \vec{a}_{i'})}}{k_n^2 - E} k_n^L Y_{nn}(k_n) e^{-\frac{k_n^2}{2}} \right. \\ &\quad \left. + \sum_{k_n^2 > k_0^2} \frac{e^{ik_n(\vec{a}_i - \vec{a}_{i'})}}{k_n^2} k_n^L Y_{nn}(k_n) e^{-\frac{k_n^2}{2}} \right] \quad (2d)\end{aligned}$$

$$+ E \sum_{k_n^2 > k_0^2} \frac{e^{ik_n(\vec{a}_i - \vec{a}_{i'})}}{k_n^4} k_n^L Y_{nn}(k_n) e^{-\frac{k_n^2}{2}}$$

$$+ E^m \sum_{k_n^2 > k_0^2} \frac{e^{ik_n(\vec{a}_i - \vec{a}_{i'})}}{k_n^{2m+2}} k_n^L Y_{nn}(k_n) e^{-\frac{k_n^2}{2}}$$

Now only the first sum is energy dependent via the denominator. The other sums, which are generated by (4b) and rearranged, are energy independent except for prefactors. In our calculations we have used $m=10$. The condition, that (4b) gives an accuracy of 4 decimal characters for $D(1)$ is

$$K_o^2 \geq 4|E| \quad (5)$$

i.e., for the usual valence and conduction band energies, $K_o^2 \approx 20$.

Having fixed K_o , we need only a single computation of the energy independent sums in (2d), which include up to 5000 reciprocal lattice vectors. This takes between 1 and 3 minutes for each \vec{k} -point on a Telefunken TR4 computer, which can be compared to an IBM 7040. Using this sums, the energy dependent parts and the determinants are calculated in 2 to 5 minutes for a given k and 100 energies.

We have done a lot of work to test the above procedure and got the result, that the relative error of the structure constants is smaller than 10^{-5} . This accuracy is by far sufficient.

Relativistic Corrections

As has been shown by Treusch [2], relativistic corrections of the Pauli equation can be taken into account without any additional approximation by expanding the Ansatz functions inside the atomic spheres into Pauli central field spinors. The way proposed by Onodera and Okazaki [3] and by Takada [4], who solve the Dirac problem, is somewhat different but should end in the same result, since the energies are small compared to the electron mass. The expansion into Pauli central field spinors leads to the radial equation

$$\left[-\frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} + \frac{\ell(\ell+1)}{r^2} + V(r) - E - \frac{\alpha^2(1+\lambda)}{4} \frac{1}{r} \frac{dV}{dr} - \frac{\alpha^2}{4} (E-V)^2 + \frac{\alpha^2}{8} \Delta V(r) \right] R_\lambda(r) = 0 \quad (6)$$

with spin-orbit, mass-velocity, and Darwin terms besides the nonrelativistic equation

$$\lambda = \begin{cases} -\ell-1 & j = \ell + \frac{1}{2} \\ \ell & j = \ell - \frac{1}{2} \end{cases}$$

$$\alpha = \frac{1}{137} = \text{Sommerfeld's constant}$$

It is well known, that this equation has no regular solution for a Coulomb-like potential, because the spin-orbit and Darwin terms exhibit a r^{-3} and $\delta(r)$ singularity, respectively, in the origin. These singularities must not be cut off, since they give the main contribution to relativistic effects. If one tries to solve the problem by an parabolic or linear approximation of the potential in the nuclear region, which avoids these difficulties, the solutions depend strongly on the assumed radius of the nucleus. It turns out, that the Pauli equation is not appropriate in this inner region.

We, therefore, shall solve the Dirac equation in the vicinity of the origin. The solution has a rapid convergent series expansion for the atomic or ionic potentials. These behave as

$$V = -\frac{2Z}{r} + \text{const. } r \rightarrow 0$$

the constant being of the order of some hundred ryd. This is an effect of the screening of the nuclear potential by the s-electrons. It is taken into account by shifting the energy scale and solving the Dirac equation for a bare ionic potential.

Since we are only interested in solutions with energies small compared to the electron mass, the Dirac bispinor can be split into small and large components of different order of magnitude, the latter being equal to Pauli central field spinors. The large components at a small radius (typical value $r=0.01$ a.u.) are now taken as starting values for an integration of the Pauli equation (6). This procedure has the advantage, that we can use the Noumerov technique for the numerical integration and that the singularities of the solutions (s and $p^{1/2}$) and potentials need not be treated by the numerical integration.

We have checked this procedure taking bare ionic potentials for Z up to 82, comparing the solutions to those of the exact Dirac equation. The agreement between the logarithmic derivatives of the solutions at a typical muffin-tin radius was found to be excellent with relative errors smaller than 10^{-3} . Calculated spin-orbit splitting coincide with experiments to within .01 eV.

Applications and Convergence

As mentioned above, we have always used the simple muffin-tin model in our applications of the KKR method. This model is not suitable to describe purely covalent bound crystals as e.g. silicon or germanium. As can be seen already from our results for III-V compounds, the muffin-tin model can be expected to be appropriate for more ionic bound crystals. This is emphasized by our calculations for II-VI and I-VII compounds, which are in agreement with experiment to a few tenth of an eV. The table gives a survey on solids calculated so far in Marburg.

wurtzite	ZnO,ZnS,CdS,CdSe [10]
diamond	C,Si,Ge, α -Sn [5]
zincblende	ZnS,ZnSe,ZnTe,CdTe [6,7,8,9] GaAs,GaP,GaSb,AlP,InP,InAs [5]
rocksalt	PbS,PbSe,PbTe [11],KCl,KBr,KI [12] CsI [13]
CsCl	CsI [13]
trigonal	Se,Te [1,14,15]
fcc	Ne,Ar,Kr,Xe [16]

The use of different potentials, nonrelativistic atomic potentials of Herman and Skillman [17] and relativistic atomic and ionic potentials of Liberman [18] did not change the results markedly. We always calculated the energy bands of full groups of materials (e.g. lead salts of potassium halides) in order to find tendencies in the electronic properties which easily and confirm an interpretation of the results.

It is an advantage of this method, that no explicit use is made of the core levels as e.g. in the OPW method. Thus it is possible to calculate also energy bands which are related to atomic d-levels. With respect to this point the superiority of the KKR method was demonstrated in the case of ZnO, where the Zn 3d-bands are closely below the Oxygen 2p-bands, mixing strongly with each other.

The KKR method solves accurately the muffin-tin problem, provided the convergence in the angular momentum quantum number 1 or λ is reached. Generally this is the case for $1 \leq 2$. Convergence calculations with $1 \leq 3$ gave energy differences smaller than .02 eV, with the exception of the higher conduction bands in the lead salts, where f-bands occur. In this case it is necessary to take into account all the wave functions up to $1=3$. By this the computer time required is doubled.

References

- 1 J.Treusch Dissertation, Marburg 1955
- 2 J.Treusch phys.stat.sol. 19,603,1966
- 3 Y.Onodera and M.Okazaki J.Phys.Soc.Japan
21,1273,1966
- 4 S.Takada Progr.Teor.Phys.,Osaka 36,224,1966
- 5 K.Mertens Dissertation, Marburg 1967
- 6 P.Eckelt Dissertation, Marburg 1967
- 7 P.Eckelt Sol.State Comm. 6,489,1968
- 8 P.Eckelt phys.stat.sol. 23,307,1967
- 9 J.Treusch,P.Eckelt, and O.Madelung II-VI Semi-conducting Compounds,1967 Int. Conf.,p.588,Benjamin N.Y.,1967
- 10 U.Rössler Phys.Rev. 184,733,1969
- 11 H.Overhof and U.Rössler phys.stat.sol.37,691, 1970
- 12 H.Overhof to be published
- 13 U.Rössler phys.stat.sol.34,207,1969
- 14 J.Treusch and R.Sandrock phys.stat.sol.16, 487,1966
- 15 B.Kramer and P.Thomas phys.stat.sol.26,151,1968
- 16 U.Rössler to be published
- 17 F.Herman and S.Skillman Atomic Structure Calculations,Prentice Hall Inc., Englewood Cliffs,N.J. 1963
- 18 D.Liberman private communication

APPROXIMATE KKR BAND-STRUCTURE SCHEMES FOR TRANSITION METALS

N. W. Dalton

IBM Corporation, Monterey & Cottle Roads, San Jose,
California 95114

INTRODUCTION

In recent years there has been much progress in the solution of the energy band problem i.e., the determination of the eigenvalues and eigenvectors of the Schroedinger equation for a given periodic potential. The two principal reasons for this are (i) the progress made in computer technology giving rise to bigger (i.e. larger core size) and faster computers required for the numerical solution of the Schroedinger equation, and (ii) improved methods for solving the latter equation. The main purpose of this paper is to review and critically compare some of the methods which have been developed during the past year or two for solving the energy band problem for transition metals. It will become apparent from the theory that simple metals (i.e., non-transition metals), and in particular pseudopotential theory (Harrison (1966)), can be treated as an approximation to the theory developed in section 3 of this paper.

There are two standard first principle schemes for calculating the electronic band structure for transition metals i.e., the KKR (Korringa (1947), Kohn and Rostoker (1954)) method and the augmented plane wave (APW) method (Slater (1937)). These schemes are difficult to use in practice and require the use of high speed electronic computers. Furthermore, the formalism is involved and hides many of the physical features of the underlying band-structure which it contains. For example, if the KKR or APW method is used to calculate the band-structure of a transition metal, such as nickel or copper, the results can be explained in terms of a narrow tight-binding 3d-band crossing and hybridising with an s-p conduction. Although the latter is evident from the final computed

results, it is masked completely by the formalism of the APW or KKR schemes. It would therefore be desirable to have a version of these schemes which exhibits the physics of the band-structure in a more transparent way. Furthermore, finding the roots of the determinantal equations in the KKR or APW schemes is very tedious and time-consuming on the computer. A scheme of calculation which avoids or considerably simplifies this part of the calculation is also desirable.

The work of Hodges e.a. (1966) and Mueller (1967) was the first step towards an efficient method of calculating transition metal band-structures. They proposed 'model Hamiltonians' \bar{H}_M of the form

$$\bar{H}_M(k) = \begin{bmatrix} \bar{v} & | & \bar{h} \\ - & - & - \\ \bar{h}^* & | & \bar{T} \end{bmatrix} \quad (1)$$

such that the electronic band-structure (i.e. ϵ vs k) of a transition metal is given by the roots of the determinantal equation

$$\det | \bar{H}_M(k) - \epsilon \bar{T} | = 0. \quad (2)$$

For face-centered cubic (f.c.c.) metals, they proposed \bar{v} to be a 4x4 matrix, similar in form to the pseudo potential matrix, to describe the s-p conduction bands in the metal. For transition metals in the 1st series (i.e., Cr, Mn, Fe, Co, Ni, Cu) \bar{T} is a 5x5 (i.e., d-state) matrix and is assumed to give rise to the narrow tight-binding d-bands. The matrix elements of \bar{T} are similar to those given by the LCAO method (Slater and Koster (1954)). Finally the mixing or hybridisation between the conduction and narrow d-bands is expressed by the matrix \bar{h} . By expressing the matrix elements of $\bar{H}_M(k)$ in terms of ten to fifteen adjustable parameters and determining these by fitting to first principles calculations at symmetry points, they demonstrated that the band-structure of a transition metal could be well-represented by a matrix of the form $H_M(k)$. Since the KKR, APW, and model Hamiltonians schemes (i.e., equations (1) and (2)) must all be equivalent to the extent that they all produce the same band-structure (for a given metallic potential) it is natural to seek the relationship between them. In an important paper (Heine (1967)) Heine was the first to show how to derive a model Hamiltonian of the form (1) as an approximation to the exact KKR theory. Furthermore, he showed the origin of the adjustable parameters used by Hodges e.a. (1966) and Mueller (1967) and how they were related to the scattering properties of the electrons. The work of Heine (1967) was then subsequently put on a more formal basis by Hubbard (1967) and Jacobs (1968). Besides providing a justification for

the success of the model Hamiltonians, these authors showed how it was possible to derive various approximate schemes of band-structure calculation. It is with this aspect of their work with which this paper is concerned. In particular, we shall examine in detail the computational schemes developed by Hubbard and Dalton (1968), Pettifor (1969), Hum and Wong (1969) and Hubbard (1969).

The plan of this paper is as follows. In section 2, we summarize some of the principal results of the exact KKR theory and discuss the representation of the energy-dependent phase-shifts which enter the theory. This is followed (section 3) by a discussion of the derivation of various approximate KKR schemes of band-structure calculation. Numerical results obtained from four approximate schemes in current use are presented and compared in section 4. Finally, we briefly summarize the contents and conclusions of the paper and indicate possible future lines of research in section 5.

KKR THEORY AND PARTIAL-WAVE PHASE-SHIFTS

Although the KKR theory of band-structure calculation may be developed for an arbitrary crystal potential it assumes a relatively simple form for potentials which can be written as the sum of non-overlapping (muffin-MT) potentials of the form

$$\begin{aligned} V_{MT}(r) &= V(r), \quad r < r_i \\ &= V_0 \quad \quad \quad r > r_i \end{aligned} \quad (3)$$

where $V(r)$ is a spherically symmetric potential, r_i is the muffin radius (usually equal to half the nearest-neighbour interatomic distance), and V_0 is a constant potential in the interstitial region between the muffin spheres centered on the lattice sites of the crystal. For metals in the first transition series, V_0 is approximately one rydberg (negative). The wave-functions corresponding to such a potential may be expanded in terms of spherical harmonics inside the muffin spheres i.e.,

$$\psi(r, \varepsilon) = \sum_L \frac{C_L}{\kappa j'_L(\kappa r_i) - L j_L(\kappa r_i)} R_\ell(r) Y_L(r) \quad (r < r_i) \quad (4)$$

where the energy $\varepsilon = \kappa^2$, L stands for the pair of quantum numbers (ℓ, m) , $j'_\ell(x)$ is the spherical Bessel function of order ℓ (Schiff (1955)), $j'_\ell(x) = \frac{d}{dx} j_\ell(x)$, $Y_L(r)$ is a spherical harmonic i.e.,

$$Y_L(r) = \left[\frac{2\ell+1}{4\pi} \frac{(\ell - |m|)!}{(\ell + |m|)!} \right]^{1/2} P_\ell^{|m|} (\cos \theta) e^{im\phi}, \quad (5)$$

$R_\ell(r)$ is the regular solution of the radial wave-equation for energy ϵ

$$\left[-\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d}{dr} \right) + V(r) + \frac{\ell(\ell+1)}{r^2} - \epsilon \right] R_\ell(r, \epsilon) = 0, \quad (6)$$

normalised to $R_\ell(r_i) = 1$, and $L_\ell = \frac{d}{dr} \log R_\ell(r)|_{r=r_i}$.

The coefficients C_L appearing in eqn. (4) are given by the solutions of the simultaneous equations (Kohn and Rostoker (1954))

$$\sum_{L'} A_{LL'}(k, \epsilon) C_{L'} = 0 \quad (7)$$

where

$$A_{LL'}(k, \epsilon) = \delta_{LL'} + \frac{\tan \eta_\ell(\epsilon)}{\kappa} A_{LL'}(k, \epsilon). \quad (8)$$

The quantities $A_{LL'}(k, \epsilon)$ do not depend upon the potential $V(r)$ but only on the symmetry of the crystal lattice i.e.,

$$A_{LL'}(k, \epsilon) = - \sum_n \frac{V_L^*(k_n, r_\ell) V_{L'}(k_n, r_{\ell'}),}{k_n^2 - \epsilon} - \kappa \frac{n_\ell(k r_{\ell'})}{j_\ell(k r_{\ell'})} \delta_{LL'}, \quad (9)$$

where $n_\ell(x)$ is the spherical Neuman function (Schiff (1955)), $k_n = k + K_n$ (K_n is a reciprocal lattice vector), $r_\ell, r_{\ell'}$ are any two radii in the interval 0 to r_i , $r_{\ell'}$ is the larger of $r_\ell, r_{\ell'}$,

$$V_L(k_n, r_\ell) = \frac{4\pi}{\sqrt{\tau}} i^{-\ell} \frac{j_\ell(k_n r_\ell)}{j_\ell(k r_\ell)} Y_L(k_n), \quad (10)$$

τ is the volume of the unit cell (direct lattice), and $i^2 = -1$. The introduction of a different radius r_ℓ for each angular momentum ℓ is in fact a slight generalization of the original KKR theory (where $r_\ell = r_i$ for all ℓ) which is useful for deriving certain approximate KKR schemes (Hubbard (1969), Pettifor (1969)). Finally $\eta_\ell(\epsilon)$ denotes the phase-shift for the ℓ -th partial wave for the potential $V(r)$ and is related to the logarithmic derivative of $R_\ell(r, \epsilon)$, L_ℓ , (cf. equation (6)) by the equation

$$\tan \eta_{\ell}(\varepsilon) = \frac{\kappa j_{\ell}'(\kappa r_i) - L_{\ell} j_{\ell}(\kappa r_i)}{\kappa n_{\ell}'(\kappa r_i) - L_{\ell} n_{\ell}(\kappa r_i)} . \quad (11)$$

Since equations (7) are homogeneous in the C_L coefficients they will possess a non-zero solution only if the determinantal equation

$$\det | \delta_{LL'} + \frac{\tan \eta_{\ell}}{\kappa} A_{LL'}(\underline{k}, \varepsilon) | = 0 \quad (12)$$

is satisfied. The roots of this equation give the energy levels ε at the Bloch vector \underline{k} of an electron moving in the crystal potential. In view of the labelling of the elements of the determinant by the quantum numbers L, L' equations (7) and (12) are often referred to as the angular momentum (AM) representation of the KKR theory. However it has been shown by Ziman (1965) and Slater (1966) that it is possible to derive completely analogous equations labelled by reciprocal lattice vectors $\underline{K}_n \underline{K}'_n$ rather than quantum numbers L, L' i.e.,

$$\sum_{n'} B_{nn'}(\underline{k}, \varepsilon) D_{n'} = 0 \quad (13)$$

and

$$\det | (k_n^2 - \varepsilon) \delta_{nn'} + \Gamma_{nn'}(\underline{k}, \varepsilon) | = 0 . \quad (14)$$

In the latter two equations

$$B_{nn'}(\underline{k}, \varepsilon) = (k_n^2 - \varepsilon) \delta_{nn'} + \Gamma_{nn'}(\underline{k}, \varepsilon), \quad (15)$$

$$\Gamma_{nn'}(\underline{k}, \varepsilon) = \frac{4\pi}{\tau} \sum_{\ell} [(2\ell+1) r_{\ell}^2 \gamma_{\ell}(r_{\ell}, \varepsilon) j_{\ell}(\kappa_n r_{\ell}) j_{\ell}(\kappa_{n'} r_{\ell}) P_{\ell}(\cos \theta_{nn'})], \quad (16)$$

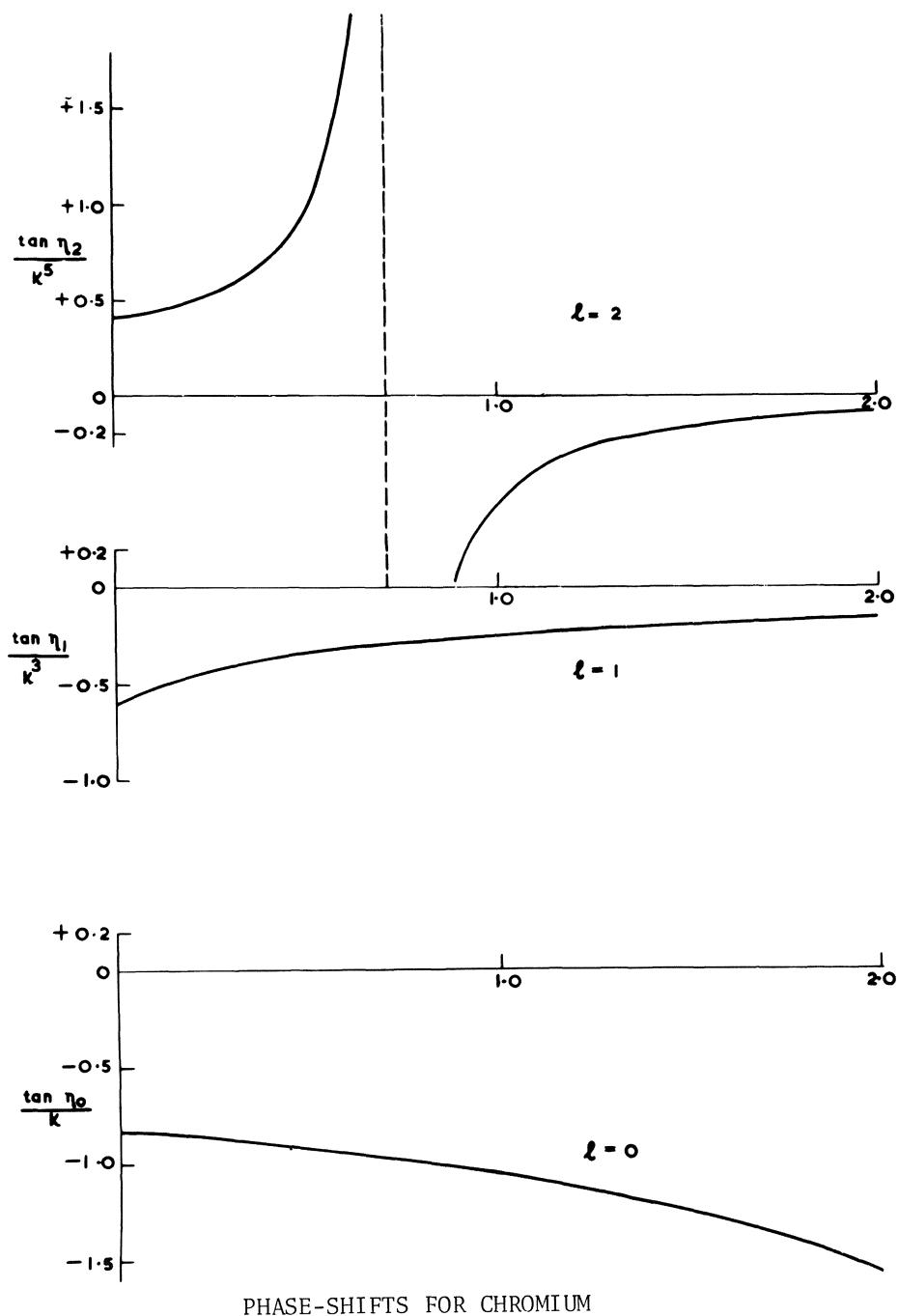
$$\gamma_{\ell}(r_{\ell}, \varepsilon) = \frac{-\tan \eta_{\ell}}{\kappa r_{\ell}^2 j_{\ell}(\kappa r_{\ell}) [j_{\ell}(\kappa r_{\ell}) - \tan \eta_{\ell} \cdot n_{\ell}(\kappa r_{\ell})]}, \quad (17)$$

$P_{\ell}(x)$ is the Legendre polynomial of order ℓ and $\theta_{nn'}$ is the angle between \underline{K}_n and $\underline{K}_{n'}$.

Equations (16) and (17) are in the form considered by Hubbard (1969) and represent a slight generalization of the equations derived by Ziman (1965) and Slater (1966). In view of the labeling of the elements of the determinant (c.f. (14)) by reciprocal lattice vectors, we refer to equations (13) and (14) as the reciprocal lattice (RL) representation of the KKR theory. Mathematically the RL and AM representations are equivalent. However, one of the main advantages of the AM representation is the separation of the potential (through the phase-shifts) from the lattice (i.e., structure constants) (c.f. (9)). On the other hand in the RL representation it is much easier to compare the KKR theory with the APW method (Slater (1937)) or pseudopotential theory (e.g., Harrison (1966)). Both forms have been used as a starting point for deriving approximate KKR schemes of band-structure calculation for transition metals. These are discussed in the following section. However, first we consider the energy behavior of the phase-shifts $\tan n_\ell(\epsilon)$, or the quantities $\gamma_\ell(r, \epsilon)$, since these contain the information regarding the crystal potential.

To calculate the phase-shifts $n_\ell(\epsilon)$ or $\tan n_\ell(\epsilon)$ for a given energy it is necessary to integrate the radial wave-equation (6) out to the radius r_i (for a given ϵ and $V(r)$) to obtain $R_\ell(r_i, \epsilon)$, and hence L_ℓ^i , and then use equation (11) to determine $\tan n_\ell(\epsilon)$. This is readily accomplished numerically by using an adaptation of the SCHEQ subroutine of the Herman-Shillman (1963) atomic structure program. Typical results for a metal in the first transition series are shown in fig. 1. (The actual potential used for this figure was taken from Asano and Yamashita (1967).) The behavior characteristic of a transition metal is the divergence in $\tan n_2(\epsilon)$ at some energy in the region 0 to 1 rydberg. (Unless otherwise stated energies are assumed to be measured relative to the constant potential between the muffin tin spheres i.e., V_0). It is important to specify the energy range of interest (Dalton and Hubbard (1968)) since $\tan n_0$, and $\tan n_1$ both exhibit divergences at energies above 1 rydberg (typically at 5-10 ryds.), and also at negative energies. However, the divergence of physical interest corresponds to a phase-shift increasing through an odd multiple of $\frac{\pi}{2}$, as occurs in $\tan n_2(\epsilon)$, and is connected with the appearance of a 'resonance' state (e.g., the 3d state in transition metals) (Newton (1966)). Although the divergences in $\tan n_0$ and $\tan n_1$ are not connected with any observable physical effects, it is important to be aware of such divergences particularly if one is attempting to parameterise the energy dependence of the phase-shifts (e.g., Cooper e.a. (1970)).

On the basis of fig. 1 and the above remarks it is tempting to express the phase-shifts in the following manner,



$$\tan \eta_\lambda(\varepsilon) \approx \tan \eta'_\lambda(\varepsilon) + \frac{\Gamma f(\varepsilon)}{\varepsilon_0 - \varepsilon} \delta_{\lambda}, \quad \varepsilon_L \leq \varepsilon \leq \varepsilon_u \quad (18)$$

where $\tan \eta'_\lambda(\varepsilon)$ is a relatively slowly varying function of energy, and the second term on the r.h.s. of equation (18) represents the divergent behavior in the $\lambda=\lambda$ component of the phase-shift. Since equation (18) is only meant to be an approximate representation for $\tan \eta'_\lambda(\varepsilon)$ in some energy region we have introduced ε_L and ε_u to denote the lower and upper bounds respectively of the energy range of interest. In practice $\varepsilon_L \approx 0.0$ ryds and $\varepsilon_u \approx 1-2$ ryds. The divergent behavior is represented by a pole at energy ε_0 and width Γ modified by a slowly varying function $f(\varepsilon)$. The choice of $f(\varepsilon)$ is clearly arbitrary since once chosen, equation (18) can be regarded as a definition of $\tan \eta'_\lambda(\varepsilon)$. If we adopt a function $f(\varepsilon)$ such that $f(\varepsilon_0) = 1.0$ then the parameters Γ and ε_0 are determined from the exact phase-shifts by the equations

$$\begin{aligned} \tan \eta_\lambda(\varepsilon_0) &= \infty, \\ \Gamma &= \lim_{\varepsilon \rightarrow \varepsilon_0} (\varepsilon_0 - \varepsilon) \tan \eta_\lambda(\varepsilon). \end{aligned} \quad (19)$$

Numerical values for these parameters for metals in the first-transition series (i.e., $\lambda=2$) are recorded in Table I. The smallness of Γ shows that, for the metals considered, the resonance is fairly narrow. The decrease in ε_0 in moving along the transition series from chromium to copper merely reflects the lowering of the 3d state energy as the 3d shell is filled.

An exact representation for $\tan \eta_\lambda(\varepsilon)$ has been obtained by Hubbard (1967) which illustrates more clearly the meaning of the terms in equation (18). Using an adaptation of the Kapur-Peierls (Kapur and Peierls (1938)) theory of resonances Hubbard derived the following results for $\tan \eta'_\lambda(\varepsilon)$ and the resonant term in equation (18) i.e.,

$$\tan \eta_\lambda(\varepsilon) = \tan \eta'_\lambda(\varepsilon) + R_\lambda(\varepsilon), \quad (20)$$

where

$$\begin{aligned} \tan \eta'_\lambda(\varepsilon) &= -k \int_0^{r_i} r^2 dr \int_0^{r_i} r'^2 dr' j_\lambda(kr) (r|U_\lambda|r') j_\lambda(kr') \\ &\quad - \sum_{\alpha}^{(c)} \frac{W(\varepsilon, r_i) \chi_{\lambda\alpha}(\varepsilon, r_i)}{E_{\lambda\alpha}(\varepsilon) - \varepsilon}, \end{aligned} \quad (21)$$

TABLE I Resonance Parameters for the $\ell=2$ partial wave phase-shift

METAL	Source of Potential (Muffin-Tin)	Constant Potential V_0 (ryds.)	Muffin-Tin Radius r_i (a.u.)	Resonance Energy ε_0 (ryds.)	Resonance Width Γ (ryds.)
Chromium (b.c.c.)	ASANO e.a. (1967)	-1.0216	2.3564	0.714	0.0806
Iron (b.c.c.)	WOOD (1962)	-0.8160	2.387	0.649	0.0608
Iron (f.c.c.)	WOOD (1962)	-0.7760	2.483	0.608	0.0546
Nickel (f.c.c.)	WAKOH (1965)	-1.0001	2.3511	0.535	0.0297
Nickel (f.c.c.)	HANUS (1962)	-0.4755	2.3530	0.420	0.0202
Copper (f.c.c.)	WAKOH (1965)	-0.9086	2.4151	0.342	0.0099
Copper (f.c.c.)	BURDICK (1963)	-0.9390	2.440	0.306	0.0082

$$(r|U_\ell|r') = V(r)\delta(r-r') - \sum^{(c)} V(r) \chi_{\ell\alpha}(r) \chi_{\ell\alpha}^*(r'), \quad (22)$$

and

$$W(\varepsilon, r_i) = r_i^2 j_\ell(kr_i) \frac{d}{dr} \log \left| \frac{n_\ell(kr_i)}{j_\ell(kr_i)} \right|_{r=r_i} \quad (23)$$

In equations (21) - (23) the wave-functions $\chi_{\ell\alpha}(r', \varepsilon)$ and the energies $E_{\ell\alpha}(\varepsilon)$ are the eigenstates and eigenvalues respectively of the radial part of the Schrödinger equation (6) subject to certain energy (ε) - dependent boundary conditions (see Hubbard (1967)). The introduction of the non-local pseudopotential discussed by Austin e.a. (1962) is made possible by a classification of the eigenstates $\chi_{\ell\alpha}(r, \varepsilon)$ into two sets, the 'core' solutions (c) which have large negative eigenvalues $E_{\ell\alpha}(\varepsilon)$ and do not differ significantly from the atomic bound-states, and the 'residual' solutions (R). The second term on the r.h.s. of equation (21) is a boundary effect term which in practice is negligible owing to the smallness of $\chi_{\ell\alpha}(\varepsilon, r_i) / [E_{\ell\alpha}(\varepsilon) - \varepsilon]$ for core solutions. We thus see that $\tan n'_\ell(\varepsilon)$ in equation (18) represents potential scattering from the Austin pseudopotential calculated in the first Born approximation.

The remaining term in (20) i.e. $R_\ell(\varepsilon)$ is given by the equation

$$R_\ell(\varepsilon) = \sum^{(R)} \frac{\Gamma_{\ell\alpha'}(\varepsilon)}{E_{\ell\alpha}(\varepsilon) - \varepsilon} \quad (24)$$

where

$$\Gamma_{\ell\alpha}(\varepsilon) = \kappa \left[\int_0^{r_i} r^2 dr \int_0^{r_i} r'^2 dr j_\ell(kr) (r|U_\ell|r') \chi_{\ell\alpha}(r', \varepsilon) \right]^2 \quad (25)$$

Clearly $R_\ell(\varepsilon)$ represents corrections to the use of the first Born approximation for $\tan n'_\ell(\varepsilon)$ and any resonant scattering. The second term on the r.h.s. of equation (18) is therefore a first order approximation to $R_\ell(\varepsilon)$ which can only be valid in a small energy region in the vicinity of ε_0 .

Finally we note that in view of the relationship (c.f. equation (11)) between $\tan n'_\ell(\varepsilon)$ and L_ℓ , any singular behavior in $\tan n'_\ell(\varepsilon)$ must be reflected in the behavior of L_ℓ . In particular a divergence in $\tan n'_\ell(\varepsilon)$ at an energy ε'_0 will be associated with a divergence in L_2 at an energy $\varepsilon'_0 > \varepsilon_0$.

Thus corresponding to equation (18) we may write the equation

$$\gamma_\ell(r_i, \varepsilon) \approx \gamma'_\ell(r_i, \varepsilon) - \frac{\Gamma' g(\varepsilon)}{\varepsilon'_0 - \varepsilon} \delta_{\ell\lambda}, \quad (26)$$

since

$$\gamma_\ell(r_i, \varepsilon) = L_\ell - \frac{j'_\ell(kr_i)}{j_\ell(kr_i)}. \quad (27)$$

The quantities γ'_ℓ , Γ' , ε' , and $g(\varepsilon)$ are clearly the analogues of $\tan \eta'_\ell$, Γ , ε and $f'_0(\varepsilon)$ respectively. By substituting (18) in (11) and using (27) it is possible to obtain more quantitative relationships between the two sets of parameters if required.

Having presented the main results of the KKR theory and discussed at length the phase-shifts which enter the theory we are now in a position to examine the approximate methods which have been used to find the roots of the KKR determinantal equations (12) and (14).

DERIVATION OF APPROXIMATE KKR BAND-STRUCTURE SCHEMES

The central problem to be considered here is the determination of the roots of the KKR determinantal equations i.e., the electronic band-structure. A direct approach to the problem is to evaluate the determinants as a function of energy and use interpolation schemes to locate the zeros. However, in practice this can be a time consuming process even on a fast computer e.g. the IBM 360/65. The principal reasons for this are (i) the time involved in evaluating the complicated elements of the determinant, and (ii) the time involved in interpolating to find the zeros of the determinants (particularly degenerate roots) since the determinants are rapidly varying functions of energy. Here we show how, by using the approximate representations (18) or (26) and the exact KKR equations, it is possible to derive approximate schemes of band-structure which are at least an order of magnitude faster than the exact KKR scheme and which clearly exhibit the physical origin of the band-structure. The following derivations are meant to be indicative of the schemes developed by Hubbard and Dalton (1968), Hum and Wong (1969), Pettifor (1969) and Hubbard (1969). [An alternative scheme proposed by Jacobs (1968) has been found to suffer from convergence problems in practice (Hum and Wong (1969), Pettifor (1969)) and has therefore not been developed to any extent.]

The primary objective of the mathematical manipulations presented below is to transform the exact KKR determinantal equation to an approximate standard secular equation of the form proposed by Hodges e.a. (1966) and Mueller (1967). This may be done either by substituting $\tan n_\ell$ from equation (18) in equation (7), or substituting $\gamma_\ell(\underline{k}, \varepsilon)$ from equation (26) in (13), and rearranging the resulting equations. For ease of comparison and completeness we consider both approaches here starting first with the AM representation and then the RL representation.

AM Representation

From equation (9) we see that $A_{LL'}(\underline{k}, \varepsilon)$ considered as a function of energy ε will diverge for energies $\varepsilon = k_n^2$. To treat this divergence it is convenient to resolve $A_{LL'}(\underline{k}, \varepsilon)$ into two parts i.e.,

$$A_{LL'}(\underline{k}, \varepsilon) = A'_{LL'}(\underline{k}, \varepsilon, \beta) + A''_{LL'}(\underline{k}, \varepsilon, \beta) \quad (28)$$

where

$$A'_{LL'}(\underline{k}, \varepsilon, \beta) = \sum_n \frac{V_L^*(k_n, r_\ell) V_{L'}(k_n, r_\ell)}{k_n^2 - \varepsilon} S_{n\ell}^2(\beta), \quad (29)$$

$$A''_{LL'}(\underline{k}, \varepsilon, \beta) = \sum_n \frac{V_L^*(k_n, r_\ell) V_{L'}(k_n, r_\ell)}{k_n^2 - \varepsilon} [1 - S_{n\ell}^2(\beta)] \\ - \kappa \frac{n_\ell \langle \mathbf{r}_\ell \rangle}{j_\ell \langle \mathbf{r}_\ell \rangle} \delta_{LL'}, \quad (30)$$

and $S_{n\ell}(\beta)$ is an arbitrary 'splitting factor'. Although $A'_{LL'}(\underline{k}, \varepsilon, \beta)$ and $A''_{LL'}(\underline{k}, \varepsilon, \beta)$ both depend upon the parameter β (or any other parameters introduced by the splitting factors) their sum, $A_{LL'}(\underline{k}, \varepsilon)$, is obviously independent of β . If $S_{n\ell}(\beta)$ is chosen so that

$$1 - S_{n\ell}^2(\beta) \sim (\varepsilon - k_n^2) \quad (31)$$

for $\varepsilon - k_n^2 \sim 0$, then $A''_{LL'}(\underline{k}, \varepsilon, \beta)$ will not diverge for any energy ε . On the other hand one would like a choice of $S_{n\ell}(\beta)$ which will make the series for $A'_{LL'}(\underline{k}, \varepsilon, \beta)$ rapidly convergent. In practice four different choices for $S_{n\ell}(\beta)$ have been investigated based upon either a Gaussian (as in the Ewald procedure) or Lorentzian function, or a combination of both. For example Hubbard (1967) and Hubbard and Dalton (1968) adopted the choice

$$S_{n\ell}^2(\beta) = e^{\frac{\varepsilon - k_n^2}{\beta}} \quad (\text{Gaussian splitting}) \quad (32)$$

whereas Jacobs (1968) proposed a splitting factor of the Lorentzian form (i.e.)

$$S_{n\ell}^2(E_0) = \frac{E_0}{E_0 - (\varepsilon - k_n^2)} \quad (\text{Lorentzian splitting}) \quad (33)$$

where E_0 is an adjustable parameter. Hum and Wong (1969) used a combination of the Gaussian and Lorentzian splittings whereas Pettifor (1969) adopted an angular momentum dependent Gaussian splitting (i.e. $S_{n\ell}^2 = 1$, $\ell \neq \lambda$ whereas for $\ell = \lambda$ the form given in (32) was used.) The important point to note here is that each choice of splitting factor introduces a parameter (two parameters in the case of Hum and Wong (1969)) which does not appear in the exact KKR theory. Hence band-structures obtained from schemes which depend upon a specific choice of splitting factor should in principle be independent of the particular numerical values assigned to the parameters. The degree to which the computed levels are insensitive to the actual choice of parameter is therefore a convenient measure of the success of the band-structure scheme. This will be examined in more detail in the following section. Finally, we remark that one of the difficulties associated with the scheme proposed by Jacobs (1968) is due to the lack of convergence arising from the use of the Lorentzian splitting.

Having discussed the choice of splitting factor we now proceed to obtain an approximate form for the KKR determinantal equations. [The following derivation is based upon the papers by Hubbard (1967) and Hubbard and Dalton (1968) and is closely related to the work of Hum and Wong (1969) and Pettifor (1969)].

From equations (7), (8), (18) and (28) we can write

$$\bar{C}_L = - \frac{\tan n_\ell}{\kappa} \sum_{L'} A_{LL'} \bar{C}_{L'} \quad (34)$$

and hence

$$\bar{C}_L = - \frac{\tan \eta'_\ell}{\kappa} \sum_{L'} A'_{LL'} \bar{C}_{L'} - \frac{\tan \eta'_\ell}{\kappa} \sum_{L'} A''_{LL'} \bar{C}_{L'} \\ (\text{plane-waves}) \quad (\text{neglect}) \quad (35)$$

$$- \delta_{\ell\lambda} \frac{\Gamma}{\varepsilon_0 - \varepsilon} \sum_{L'} A_{LL'} \bar{C}_{L'} \\ (\text{resonant function})$$

In writing the above equation we have first separated the terms into those associated with the potential and resonant contributions to the phase-shifts, and secondly, separated the potential terms into those relating to the singular ($A'_{LL'}$) and non-singular ($A''_{LL'}$) parts of $A_{LL'}$. For simplicity we have assumed $f(\varepsilon) \equiv 1$. This particular choice of $f(\varepsilon)$ predicts an incorrect behavior for $\tan \eta_2(\varepsilon)$ at small energies (Hubbard and Dalton (1968) but we shall ignore this for the purposes of the present discussion. The next step is to introduce plane-wave expansion coefficients, b_n , by replacing the summation over L' in the first term on the r.h.s. of (35) using the definition of $A'_{LL'}$ (c.f. (29)) i.e.,

$$\sum_{L'} A'_{LL'} \bar{C}_{L'} = - \sum_{L'n} \sum_{\ell} \frac{V_L^*(k_n, r_\ell) V_{L'}(k_n, r'_\ell)}{k_n^2 - \varepsilon} S_{n\ell}(\beta) \bar{C}_{L'} \\ = - \sum_n t_{n\ell}^*(\beta) b_n \quad (36)$$

where

$$b_n = \frac{1}{(k_n^2 - \varepsilon)} \sum_{L'} t_{n\ell}(\beta) \bar{C}_{L'} \quad (37)$$

and

$$t_{n\ell}(\beta) = V_L(k_n, r_\ell) S_{n\ell}(\beta). \quad (38)$$

In a similar manner we may introduce a set of expansion coefficients a_m ($m = -\lambda, -\lambda+1, \dots, +\lambda$) to describe the resonant part of the wave-function (c.f. (35)) i.e.,

$$\Gamma^{1/2} a_m = \frac{\Gamma}{\varepsilon_0 - \varepsilon} \sum_{L'} (A'_{\lambda m, L'} + A''_{\lambda m, L'}) \bar{C}_{L'}. \quad (39)$$

The factor $\Gamma^{1/2}$ has been introduced for purposes of symmetry only. Finally, we may argue that the remaining term in (35) may be neglected since it includes the product $\tan n'_\lambda A''_{LL}$, which is assumed to be small. [Further discussion on this point can be found in Hubbard and Dalton (1968) and Pettifor (1969).] Combining the above results yields the following expression for the expansion coefficient \bar{C}_L in terms of the new coefficients b_n and a_m i.e.,

$$\bar{C}_L \approx \frac{\tan n'_\lambda}{\kappa} \sum_n t_{n\lambda}^*(\beta) b_n - \Gamma^{1/2} a_m \delta_{\lambda\lambda}. \quad (40)$$

To determine the transformed determinantal equation corresponding to (12) we now seek the condition that the simultaneous equations which b_n and a_m satisfy possess a non-zero solution. These equations are obtained by eliminating the coefficients \bar{C}_L from (37) and (39) using (40) i.e.,

$$\sum_{n'=0}^{\infty} [\kappa_n^2 \delta_{nn'} + \alpha_{nn'}(\varepsilon)] b_{n'} + \sum_{m'=-\lambda}^{+\lambda} h_{m'n}(\varepsilon) a_{m'} = \varepsilon b_n, \quad (41)$$

$$\sum_{n'=0}^{\infty} h_{mn}^*(\varepsilon) b_{n'} + \sum_{m'=-\lambda}^{+\lambda} [\varepsilon_0 \delta_{mm'} + \Gamma A_{\lambda m' m}(\varepsilon)] a_{m'} = \varepsilon a_m + \delta_m, \quad (42)$$

where

$$\alpha_{nn'}(\varepsilon) = \sum_\lambda t_{n\lambda}(\beta) \frac{\tan n'_\lambda(\varepsilon)}{\kappa} t_{n'\lambda}(\beta), \quad (43)$$

$$h_{mn}(\varepsilon) = \Gamma^{1/2} t_{n\lambda}(\beta), \quad (44)$$

and δ_m is a small correction term containing the product $\tan n'_\lambda A_{\lambda m' m}$, which we hereon neglect. Although we have not exhibited the energy dependence of $t_{n\lambda}(\beta)$ explicitly it should be remembered that this quantity is energy dependent through the factors $V_L(k_n, r_\lambda)$ and $S_{n\lambda}(\beta)$ (cf (38), (32) and (10)). The condition that equations (41) and (42) have a solution yields the equation

$$\det | \bar{H}(\varepsilon) - \varepsilon \bar{I} | = 0, \quad (45)$$

where

$$\tilde{H}(\epsilon) = \begin{bmatrix} K_0 & K_1 & K_2 & \dots & K_n & -\lambda, & \dots & \dots & \dots & +\lambda \\ k_n^2 \delta_{nn'} + \alpha_{nn'}(\epsilon) & & & & h_{m'n}(\epsilon) & & & & & \\ \vdots & & & & \vdots & & & & & \\ \vdots & & & & \vdots & & & & & \\ h_{mn}^*(\epsilon) & & & & \varepsilon_0 \delta_{mm'} + \Gamma A_{\lambda m \lambda m'}(\epsilon) & & & & & \end{bmatrix} \quad (46)$$

Equation (45) is the approximate transformation of the exact KKR equation (12) we have been looking for i.e., it has a structure similar to that proposed in the work of Mueller (1967) and Hodges e.a. (1968).

However, $\tilde{H}(\epsilon)$ differs from their model Hamiltonians, H_M , in being energy-dependent. The next step is to use our physical understanding of the nature of the bands in transition metals to eliminate the energy ϵ from $\tilde{H}(\epsilon)$. For example we know that the tight-binding block $\Gamma A_{\lambda m \lambda m'}(\epsilon)$ describes the narrow 3d-band whereas the plane-wave block, $k_n^2 \delta_{nn'} + \alpha_{nn'}(\epsilon)$, describes the s-p conduction bands. Therefore, one way of constructing a model Hamiltonian from $\tilde{H}(\epsilon)$ is to write

$$\tilde{H}_M = \begin{bmatrix} k_n^2 \delta_{nn'} + \alpha_{nn'}(\epsilon_p) & h_{m'n}(\epsilon_d) \\ \vdots & \vdots \\ \vdots & \vdots \\ h_{mn}^*(\epsilon_d) & \varepsilon_0 \delta_{m'm} + \Gamma A_{\lambda m, \lambda m'}(\epsilon_d) \end{bmatrix} \quad (47)$$

where

$$\epsilon_p = k_n^2, \quad \epsilon_d = \epsilon_0. \quad (48)$$

The method of eliminating the energy from $\tilde{H}(\epsilon)$ is to a large extent arbitrary and is one of the features which distinguishes the different approximate KKR schemes. In figure 2 we have attempted to summarise the essential ingredients common to all the approximate KKR schemes in current use. The actual matrix elements obtained for \tilde{H}_M depend upon the particular choice of splitting factor and the method whereby the energy is eliminated from $\tilde{H}(\epsilon)$. The physical origin of the form of the matrix elements is indicated in figure 2 and discussed at length by Heine (Heine (1967), also Heine (1969)).

$$\det / H_M - \epsilon I / = 0$$

[PSEUDOPOTENTIAL THEORY]

CONDUCTION BANDS
ARISING FROM POTENTIAL
SCATTERING

$$H_M = \begin{bmatrix} \psi_{\mathbf{k}\mathbf{k}'} & (\mathbf{k}+\mathbf{k}')^2 \theta_{\mathbf{k}\mathbf{k}} - \frac{4\pi}{r} \sum_l (2l+1) \tan \eta_l^1(\mathbf{k}^2) A_l(\mathbf{k}) P_l(\cos \theta_{\mathbf{k}\mathbf{k}'}) \\ h_{m\mathbf{k}} & h_{m\mathbf{k}} \end{bmatrix}$$

$$h_{m\mathbf{k}} = \frac{4\pi}{r^2} \Gamma^{\frac{1}{2}} Y_L(\mathbf{k}+\mathbf{k}') \Gamma(\epsilon_0, \mathbf{k})$$

$$\epsilon_0 \delta_{mm} + \Gamma_{mm}(\epsilon_0)$$

- ① HYBRIDISATION INTERACTION RESULTS FROM DECAY
OF ATOMIC D-STATE INTO THE $l=2$ COMPONENT
OF THE PLANE WAVE $|\mathbf{k}+\mathbf{k}'\rangle$

$$h_{m\mathbf{k}} \sim \langle \Phi_m | V | \mathbf{k}+\mathbf{k}' \rangle \sim \int \Phi(\epsilon_0) V Y_L(\mathbf{k}+\mathbf{k}') j_l(\mathbf{k}+\mathbf{k}') r^2 dr$$

$$\sim r^{\frac{1}{2}} Y_L(\mathbf{k}+\mathbf{k}')$$

- ① D-BANDS ARISE FROM RESONANT SCATTERING
② STRUCTURE OF BANDS DEPENDS UPON
LATTICE STRUCTURE AND ϵ_0 .
③ WIDTH OF BANDS IS PROPORTIONAL TO
 $\Gamma \sim [\int_0^{r_i} j_l(\mathbf{k}_0') V \Psi_l(\epsilon_0)]^{\frac{1}{2}}$
④ $\Gamma_{mm}(\epsilon_0)$ CAN BE EXPRESSED IN TERMS OF
TWO-CENTRE OVERLAP INTEGRALS

MODEL HAMILTONIAN FOR TRANSITION METALS

From (45) and (47) we may regard the model Hamiltonian as a first step in the iterative solution of the approximate KKR equation. Thus one can begin by replacing $\tilde{H}(\epsilon)$ by $\tilde{H}(\tilde{\epsilon}_1)$ where $\tilde{\epsilon}_1$ denotes an approximate set of roots for equation (15). One then solves the secular equation with $\tilde{H}(\epsilon)$ replaced by $\tilde{H}(\tilde{\epsilon}_1)$ to determine a more accurate set of roots $\tilde{\epsilon}_2$, and so on. This procedure has not been investigated in detail to date, but would seem to be a plausible way for finding accurate roots of equation (45). There would obviously be some arbitrariness in how one should replace the energy in the off-diagonal matrix elements but this should not be serious.

In the following section we discuss the numerical results obtained from approximate schemes such as those described above. However, for comparison, we now examine an alternative approach developed by Hubbard (1969) which avoids the introduction of a splitting factor, $S_{n\ell}(\beta)$, by using the RL representation of the KKR theory instead of the AM representation.

RL Representation

To derive an approximate KKR scheme in the RL representation we begin with the approximate representation (26) in conjunction with equations (13) to (16) and proceed as above. Thus, following Hubbard (1969), we write

$$(k_n^2 - \epsilon) D_n = - \sum_n' \Gamma_{nn'} D_{n'},$$

$$\text{i.e., } (k_n^2 - \epsilon) D_n = \frac{4\pi}{\tau} \sum_{n'} \sum_{\ell} (2\ell+1) r_{\ell}^2 \gamma_{\ell}' j_{\ell} (k_n r_{\ell}) j_{\ell} (k_n, r_{\ell}) P_{\ell} (\cos \theta_{nn'}) D_r \\ \text{(plane-waves)} \\ + \frac{4\pi r_{\lambda}^2}{\epsilon' - \epsilon} \sum_{n'} \sum_{\lambda} (2\lambda+1) j_{\lambda} (k_n r_{\lambda}) j_{\lambda} (k_n, r_{\lambda}) P_{\lambda} (\cos \theta_{nn'}) D_r \\ \text{(resonant function)} \quad (49)$$

As before (cf (35)) the terms associated with the 'potential' part of $\gamma_{\ell}(r_{\ell}, \epsilon)$ have been separated from those associated with the singular part of $\gamma_{\ell}(r_{\ell}, \epsilon)$. Here again, for simplicity, we have set $g(\epsilon)$ in equation (26) equal to unity. By comparing these equations with their analogues in the AM representation we note an immediate advantage of starting from the RL representation i.e., there is no need to introduce a splitting factor, $S_{n\ell}(\beta)$, and hence an adjustable parameter, owing to the non-singular behavior of $\Gamma_{nn'}(k, \epsilon)$. This is because the singular behavior

associated with $A_{LL}'(k, \varepsilon)$ at $\varepsilon=k^2$ has already been transformed away in passing from the AM to the RL representation. Furthermore whereas in the previous case we defined two sets of coefficients, i.e., the b 's and a 's, here it is necessary to introduce only one set, i.e., the a 's, since the coefficients D are already defined. Thus the resonant function in equation (49) is rewritten as follows,

$$\begin{aligned} & \frac{4\pi r_\lambda^2}{\tau} \frac{\Gamma'}{\varepsilon_0' - \varepsilon} \sum_n' (2\lambda+1) j_\lambda(k_n r_\lambda) j_\lambda(k_n, r_\lambda) P_\lambda(\cos \theta_{nn'}) D_n' = \\ &= \frac{(4\pi)^2 r_\lambda^2}{\tau} \frac{\Gamma'}{\varepsilon_0' - \varepsilon} \sum_n' \sum_m j_\lambda(k_n r_\lambda) j_\lambda(k_n, r_\lambda) Y_{\lambda m}(k_n) Y_{\lambda m}^*(k_n) D_n' \\ &= - \sum_m h_{nm} a_m \end{aligned} \quad (50)$$

$$\text{where } h_{nm} = -4\pi r_\lambda \left(\frac{\Gamma'}{\tau} \right)^{1/2} j_\lambda'(k_n r_\lambda) Y_{\lambda m}(k_n), \quad (51)$$

$$a_m = \frac{4\pi r_\lambda}{\varepsilon_0' - \varepsilon} \left(\frac{\Gamma'}{\tau} \right)^{1/2} \sum_n j_\lambda(k_n r_\lambda) Y_{\lambda m}^*(k_n) D_n. \quad (52)$$

In deriving equation (52) we have used the standard formula relating the degendre polynminal $P_\lambda(\cos \theta_{nn'})$ to the spherical harmonic $Y_{\lambda m}(k_n), Y_{\lambda m}^*(k_n)$ in order to group together all terms depending on the wave-vector k_n . Defining $V_{nn'}(\varepsilon)$ by the equation

$$V_{nn'}(\varepsilon) = \frac{4\pi}{\tau} \sum_\lambda (2\lambda+1) r_\lambda^2 \gamma_\lambda'(\varepsilon) j_\lambda(k_n r_\lambda) j_\lambda(k_n, r_\lambda) P_\lambda(\cos \theta_{nn'}) \quad (53)$$

results in the following alternative form for equations (49) and (52),

$$\sum_n [(k_n^2 - \varepsilon) \delta_{nn'} + V_{nn'}(\varepsilon)] D_n' + \sum_{m=-\lambda}^{+\lambda} h_{nm} a_m = 0 \quad (54)$$

$$\sum_n h_{n'm}^* D_n' + \sum_{m=-\lambda}^{+\lambda} (\varepsilon_0' - \varepsilon) \delta_{mm'} a_m = 0 \quad (55)$$

The condition that these equations have a solution gives rise to the determinantal equation

$$\det |\bar{H}'(\varepsilon) - \varepsilon I| = 0 \quad (56)$$

where

$$\bar{H}'(\varepsilon) = \begin{bmatrix} (k_n^2) \delta_{nn'} + V_{nn'}(\varepsilon) & h_{nm} \\ h_{n'm}^* & \varepsilon_0 \delta_{mm'} \end{bmatrix}. \quad (57)$$

It is worth stressing, that, apart from the particular representation adopted for $\gamma_\ell(\varepsilon)$, the above equation is exact. At this stage $\bar{H}'(\varepsilon)$ is not quite in the form of \bar{H}_M , i.e., the model Hamiltonian used by Mueller (1967) and Hodges et al. (1966), owing to the diagonal nature of the tight-binding block i.e., $\varepsilon_0 \delta_{mm'}$. From a purely computational point of view, however, it might be worthwhile attempting to find the roots of equation (56) using the iterative scheme mentioned earlier. This has not yet been done. An alternative procedure, and the one adopted by Hubbard (1969), is to reduce the infinite order equation (56) to a low order one using the 'folding' technique of Heine (1967). In the application of this technique one divides the vectors k_n into two sets: a 'preferred set' P and the 'remainder' set R . The set P is chosen so that the energy of a vector in the remainder set R , is far removed from the energy range of interest e.g. 0-1 rydberg. This ensures that $\varepsilon - k_n^2$ never becomes small for vectors k_n in the set R . By proceeding in this way Hubbard (1969) was able to reduce $\bar{H}'(\varepsilon)$ (approximately) to the form $\bar{H}_1'(\varepsilon)$ where

$$\bar{H}_1'(\varepsilon) = \begin{bmatrix} K_1 K_2 \dots \dots \dots \dots \dots \dots K_{P_i} & \\ & | \\ (k_n^2) \delta_{nn'} + V_{nn'}(\varepsilon) & h_{nm} \\ h_{nm}^* & \varepsilon_0 \delta_{mm'} + \sum_{mm'}^{(R)} \end{bmatrix}, \quad (58)$$

and

$$\sum_{mm'}^{(R)} = -\Gamma \frac{(4\pi r_\lambda)^2}{\tau} \sum_n^{(R)} \frac{j_\lambda^2(k_n r_\lambda)}{k_n^2 - \varepsilon} Y_{\lambda m}^*(k_n) Y_{\lambda m'}(k_n). \quad (59)$$

The main advantage of $\bar{H}_1'(\varepsilon)$ compared with $\bar{H}_1(\varepsilon)$ is its smaller size. The size of $\bar{H}_1'(\varepsilon)$ depends upon the number of vectors in the set P which in practice is 4 for a face-centered cubic (f.c.c.) lattice and 7 for a body-centered cubic (b.c.c.) lattice. The choice of the set P, as discussed above, ensures that the matrix elements $\sum_{mm'}(R)$ do not diverge for energies ε in the energy range of interest. It is worth noting here that the matrix elements h_{nm} in the present scheme are automatically independent of the energy ε . This is also true of the scheme developed by Hubbard and Dalton (1968) (but not in the schemes of Hum and Wong (1969) or Pettifor (1969)) but only after making a particular choice for the function $f(\varepsilon) \neq 1$ (c.f. (18)). Starting from $\bar{H}_1'(\varepsilon)$ it is now possible to derive a variety of model Hamiltonians, as outlined earlier, by suitably eliminating the energy from the matrix elements. For details of the specific scheme proposed by Hubbard we refer to his original paper (Hubbard (1969)).

We now examine some of the numerical results obtained using the approximate KKR schemes to assess their suitability (in terms of speed and accuracy) for calculating electronic band-structures.

NUMERICAL RESULTS FROM APPROXIMATE KKR SCHEMES

The success of the approximate KKR schemes is clearly going to depend markedly on the accuracy with which the energy levels can be calculated. As a guideline we recall that published results obtained using the exact KKR method have uncertainties of the order 0.001-0.005 ryds. The uncertainty in a computed energy level of course depends upon the degree to which the calculation has converged i.e. upon the amount of computer time allowed for the calculation. However the figures quoted for the uncertainty are typical of the accuracy arrived at in a first-principle calculation. It is unlikely that an approximate KKR scheme will yield energy levels to this accuracy owing to the number of approximations involved in deriving the final energy-independent model Hamiltonian. (c.f. fig. 2) However one of the arguments for deriving such schemes in the first place was that such high accuracy (i.e. ~ 0.001 ryds) was not really required in many cases in view of the uncertainties in the one-electron potential used in the calculation. Replacing the approximate potential by the 'true' potential would probably result in changes in the energy levels ~ 0.01 ryds (or more) so that it is not necessary to obtain energy levels to better than 0.01 ryds. Therefore the accuracy to be aimed at in these schemes is 0.01-0.02 ryds. On the other hand to compensate for the loss of accuracy one would hope to gain on speed and flexibility. Since the time to calculate the energy levels at a point k depends markedly on the particular computer used it is important to specify the latter

when quoting times. In the present paper estimates for times are understood to refer to the IBM 360/65 machine. For purposes of comparison, we note that the time to calculate the lowest six energy levels at a general point in k -space using the exact KKR scheme is of the order 10 secs on this machine. Here again this figure is intended only as a guideline since the actual times will depend upon the specific k -point considered, on whether degenerate roots are present, and on the specific programming techniques used. [NB. Using an IBM 360/91 machine and streamlined programming methods it is currently possible to reduce the figure of 10 secs to ~1 sec (H. L. Davis-private communication).] A realistic aim for the approximate KKR schemes is therefore a time of 0.1 secs (IBM 360/65 machine) to calculate the lowest six or seven levels at an arbitrary k -point.

Below we first outline the steps required to obtain the energy levels at a given k -point using the approximate KKR scheme and then discuss each step in turn.

Computational Procedure

A. Preliminary

1. Tabulate the phase-shifts as a function of energy for each angular momentum ℓ , i.e.,

$$\frac{\tan \eta_\ell(\varepsilon)}{k^{2\ell+1}} \quad \text{vs. } \varepsilon, \ell = 0, 1, 2$$

2. Determine the resonance parameters Γ and ε_0 associated with $\tan \eta_2$ and tabulate the residual phase-shifts (c.f. (18)) i.e.,

$$\frac{\tan \eta'_2(\varepsilon)}{k^5} \quad \text{vs. } \varepsilon$$

3. Determine the size of the conduction block i.e. the number of reciprocal lattice vectors to be included.
4. Determine the adjustable parameters (e.g. β).

B. Iterative

5. Calculate the conduction block matrix elements.
6. Calculate the hybridisation matrix elements.
7. Calculate the tight-binding matrix elements.
8. Determine the eigenvalues of the standard secular equation.

As can be seen above we have divided the computational procedure into two parts, A and B. The first part contains those steps in the calculation which need only be executed once whereas the remaining steps (i.e. 5-8) must be repeated at every k-point. The only change in steps 1-8 above in going from schemes based upon the AM representation to those based on the RI representation is the replacement of $\tan \eta_\ell(\epsilon) / 2\ell+1$ by $\gamma_\ell(\epsilon, r_\ell)$. In section 2 we have already discussed the calculation of the phase-shifts from a given one-electron potential and the determination of the resonance parameters Γ and ϵ (steps 1 and 2). Here we need only add that an energy interval of 0.01 ryds i.e. = $0.0(0.01)2.0$ ryds is usually adequate for the tabulation of the phase-shifts although in order to obtain accurate estimate for Γ and ϵ an energy interval of 0.001 ryds is preferable. Furthermore we note that in calculating the potential part of $\tan \eta_2$, i.e. $\tan \eta_2'$, by subtracting off the resonant term, small oscillations are introduced into $\tan \eta_2'$ in the vicinity of the resonant energy ϵ . It is convenient to eliminate these by some simple interpolation procedure before calculating the conduction block matrix elements (step 5 above).

The determination of the size of the conduction block and the adjustable parameters (steps 3 and 4) can only be done by intelligent trial and error i.e., choosing particular sets of reciprocal lattice vectors and specific values for the parameters and calculating the corresponding eigenvalues. By repeating this procedure several times it is then possible to decide upon the most suitable choice for the vectors and the parameters. This is not such a formidable program as it may sound for we know from experience that the only reciprocal lattice vectors which we need consider are those belonging to the nearest-neighbor or next-nearest neighbor sets (including the vector at the origin). In some isolated cases it may be necessary to include one or two representative lattice vectors from the third-neighbor set to achieve high accuracy at particular points in the zone, but in general these need not be considered. The choice of the value of the adjustable parameter is linked to some extent with the choice of reciprocal lattice vectors since it plays the role of a convergence parameter. However from our previous discussion we would hope that for a given choice of lattice vectors we would be able to find a range of values for the parameter in which the computed energy levels were not sensitive to the particular choice of the parameter. For definiteness we record here the minimal sets of lattice vectors that have been used in the model Hamiltonian schemes:

$$\text{f.c.c. } 4\text{K-vectors } \underline{K}_1 = \frac{2\pi}{a} (000), \underline{K}_2 = \frac{2\pi}{a} (\bar{1}\bar{1}\bar{1}), \underline{K}_3 = \frac{2\pi}{a} (\bar{1}\bar{1}0), \\ \underline{K}_4 = \frac{2\pi}{a} (0\bar{2}0).$$

$$\text{b.c.c. } 7\text{K-vectors } \underline{K}_1 = \frac{2\pi}{a}(000), \underline{K}_2 = \frac{2\pi}{a}(0\bar{1}1), \underline{K}_3 = \frac{2\pi}{a}(0\bar{1}\bar{1}), \\ \underline{K}_4 = \frac{2\pi}{a}(1\bar{1}0), \underline{K}_5 = \frac{2\pi}{a}(\bar{1}10), \underline{K}_6 = \frac{2\pi}{a}(\bar{1}0\bar{1}), \underline{K}_7 = \frac{2\pi}{a}(0\bar{2}0).$$

Thus the minimal set in the f.c.c. case contains two nearest and one next-nearest neighbor reciprocal lattice vectors whereas the b.c.c. case requires five nearest and one next-nearest neighbor reciprocal lattice vectors.

The procedure adopted to find the best choice of the adjustable parameter was the following:

- (i) Choose a set of reciprocal lattice vectors (e.g. the minimal set) for the conduction block.
- (ii) For several values of the parameter evaluate the energy levels at a few symmetry points (e.g. $\Gamma(000)$, $X(010)$, $L(\frac{1}{2}\frac{1}{2}\frac{1}{2})$, $K(3/2, 3/2, 0)$, $W(\frac{1}{2}10)$, $\Delta(0\frac{1}{2}0)$, $\Lambda(\frac{1}{4}\frac{1}{4}\frac{1}{4})$ and $\Sigma(3/8 3/8, 0)$ for the f.c.c. lattice, and $\Gamma(000)$, $H(010)$, $\Sigma(\frac{1}{4}\frac{1}{4}0)$, $N(\frac{1}{2}\frac{1}{2}0)$, $\Delta(0\frac{1}{2}, 0)$, $\Lambda(\frac{1}{4}\frac{1}{4}\frac{1}{4})$, $P(\frac{1}{2}\frac{1}{2}\frac{1}{2})$, $D(\frac{1}{2}\frac{1}{2}\frac{1}{4})$ and $G(\frac{1}{4}, 3/4, 0)$ for the b.c.c. lattice)
- (iii) By examining the sensitivity of several levels to the choice of the parameter determine those values of the parameter for which the levels are least sensitive.
- (iv) Choose that value of parameter which lies in the middle of the range of least sensitivity for the largest number of levels.

To illustrate the above procedure (i.e. (i)-(iv)) we have recorded the levels at certain symmetry points, for a given size of conduction block, as a function of the adjustable parameter in Tables II, III, IV. Table II is taken from Hubbard and Dalton (1968), Table III from Pettifor (1969) and Table IV from Wong (1970). An examination of these tables shows that whereas the d-like levels (e.g. X_5) are almost independent of the adjustable parameter (c.f. column 5 of Table II and column 4 of Table III) the low-lying levels at L (c.f. L_1 in Tables II and III) vary markedly with β . For example in Pettifor's scheme (Table III) L_1 varies from 0.227 to 0.386 ryds (exact value = 0.371 ryds) as β varies from 0.3 to 0.9 ryds. However it should be noted that at the expense of introducing two parameters (instead of one) and using a fairly large conduction block it is possible to practically eliminate any variation in the energy levels over quite a wide choice of parameters (cf Table IV). An important additional consideration in deciding upon the optimum choice of the parameter is the numerical accuracy of the levels. Therefore at the bottom of each column in Tables II and III we have added the results of Wood (1962) using the exact APW scheme. On the basis of this discussion and the results of Tables II to IV (which are merely representative of a

TABLE II. Dependence of energy levels on the parameter β

[(HD scheme - 15K-vectors), f.c.c. iron Wood (1962)]

β	Γ_{25}'	Γ_{12}	x_1	x_5	L_1	L_3	L_3'	K_1	K_1'
0.6	0.604	0.711	0.363	0.821	0.376	0.571	0.849	0.438	1.164
0.7	0.604	0.708	0.373	0.817	0.388	0.581	0.826	0.442	1.163
0.8	0.605	0.706	0.381	0.818	0.396	0.589	0.817	0.446	1.162
0.9	0.607	0.705	0.388	0.824	0.403	0.596	0.817	0.450	1.163

Wood
(1962) 0.606 0.712 0.355 0.807 0.371 0.595 0.787 0.421 1.168

TABLE III. Dependence of energy levels on the parameter β

[(P scheme - 4K-vectors), f.c.c. iron Wood (1962)]

β	x_1	x_3	x_5	L_1	L_3	L_3'	Δ_1	Δ_2'	Δ_5
0.3	0.212	0.398	0.793	0.227	0.596	0.780	0.245	0.510	0.687
0.5	0.303	0.416	0.790	0.327	0.599	0.777	0.257	0.513	0.695
0.7	0.346	0.441	0.790	0.366	0.605	0.777	0.260	0.525	0.699
0.9	0.372	0.462	0.790	0.386	0.611	0.777	0.261	0.537	0.704

Wood
(1962) 0.355 0.430 0.807 0.371 0.595 0.787 0.267 0.516 0.692

TABLE IV. Dependence of energy levels on the parameter
 [(HW scheme - 23K-vectors, $n = 1.0$ ryds)[†]]
 [h.c.p. Cobalt, Hodges & Ehrenreich (1968)]

β	Γ_{1+}	Γ_{4-}	Γ_{5+}	Δ_1	Δ_2	Δ_5	H_3	H_2	H_1
2.0	0.062	0.278	0.466	0.104	0.240	0.463	0.372	0.438	0.482
4.0	0.062	0.292	0.467	0.104	0.246	0.462	0.374	0.439	0.489
6.0	0.062	0.297	0.467	0.104	0.249	0.462	0.376	0.440	0.491
8.0	0.062	0.300	0.468	0.104	0.250	0.462	0.377	0.440	0.492

[†]This table is taken from Wong (1970)

more complete investigation) it is reasonable to choose the values 0.8, 0.6 and 5.0 for the HD, P and HW schemes respectively. Different sets of K-vectors will lead to slightly different choices for the optimum parameters but the values quoted are in current use. We stress here that it is assumed that the same value of the parameter is applicable to both the f.c.c. and b.c.c. lattices (and for the h.c.p. lattice, Wong (1970)) so that there is no question of having to determine new values for the parameter every time a new potential or lattice is used. So far no mention has been made of the Hubbard (1969) H-scheme, for obvious reasons. As explained in the previous section this scheme did not require the introduction of a splitting factor, and hence adjustable parameters, and is one of its main advantages over the other schemes.

Before discussing the accuracy of the schemes we briefly consider here the calculation of the matrix elements (i.e. steps 5, 6, and 7 of the computational procedure). The form of the matrix elements is given in Figure 2. The conduction block and hybridisation matrix elements present no problem since the former are readily evaluated using the tables of phase-shifts (c.f. step 1) (only three terms are included in the sum $\ell = 0, 1, 2$), and the latter consists of only a single term. Although we have not given explicit formulae for the tight-binding matrix elements (c.f. T_{mm} , in figure 2) they are easily evaluated except possibly for the 'overlap' integrals which contain integrals of the form

$$I(\alpha, \gamma) = \int_{\alpha}^{\infty} x^L e^{-x^2 + \frac{\gamma}{x^2}} dx \quad (60)$$

However these integrals are familiar from the theory of heat-diffusion (Carslaw and Jaeger (1959) and can be expressed in terms of the error function with imaginary argument. The latter can be evaluated using standard Library sub-routines. Only one or two of these integrals (corresponding to a nearest or next-nearest neighbor overlap integral respectively) need be calculated in practice. In Hubbard's scheme the matrix elements $T_{mm'}$ are expressed in terms of a sum over reciprocal lattice vectors (c.f. (59)) instead of a direct lattice sum involving overlap integrals. However the reciprocal lattice sum can be readily evaluated by approximating part of the sum by an integral (i.e. the sine integral). Thus in practice the time to evaluate the matrix elements is approximately independent of the particular scheme used. The main factor determining the time to evaluate the energy levels at a general \mathbf{k} -point is therefore the size of the model Hamiltonian. [NB This statement is true only if we are considering the procedure as outlined above. To improve the accuracy of the energy levels Hubbard used in addition, a variational principle, thus requiring further computer time. This explains the discrepancy between the times quoted for Pettifor (1969) and Hubbard (1969) in Table VII.]

Having decided upon the choice of parameters and the size of the conduction block we now examine the accuracy of the schemes. In Table V we have recorded a selection of levels for copper and b.c.c. iron for comparison with the first-principles results of Burdick (1963) and Wood (1962) respectively. It will be noticed that most of the levels given by the approximate schemes differ by less than 0.02 ryds from the exact results and many of them differ by no more than 0.01 ryds. This information is shown in a more transparent form in Table VI and Figure 3.

[NB Table VI is taken from Hubbard and Dalton (1968) and is representative of all schemes considered. Besides showing the distribution of errors for $\beta=0.8$ it includes corresponding results for different sets of \mathbf{k} -vectors.]

To conclude this section we have summarized various aspects of the four different computational schemes in Table VII. In the second row we have recorded the size of the conduction blocks required to obtain energy levels to an accuracy shown in the fourth and fifth rows. The corresponding values of the adjustable parameters are shown in the third row. It is clear from this table that Hubbard's scheme is superior to the other schemes in that it does not involve any adjustable parameters and is more accurate than the other schemes. However it should be pointed out that to attain the accuracy quoted Hubbard (1969) included various refinements to his scheme which have not been discussed in this paper. Pettifor's scheme, although giving accurate results for the f.c.c. lattice (the application of his scheme to other lattices has not yet been

TABLE V. A comparison of the energy levels for copper and
(b.c.c.) iron.

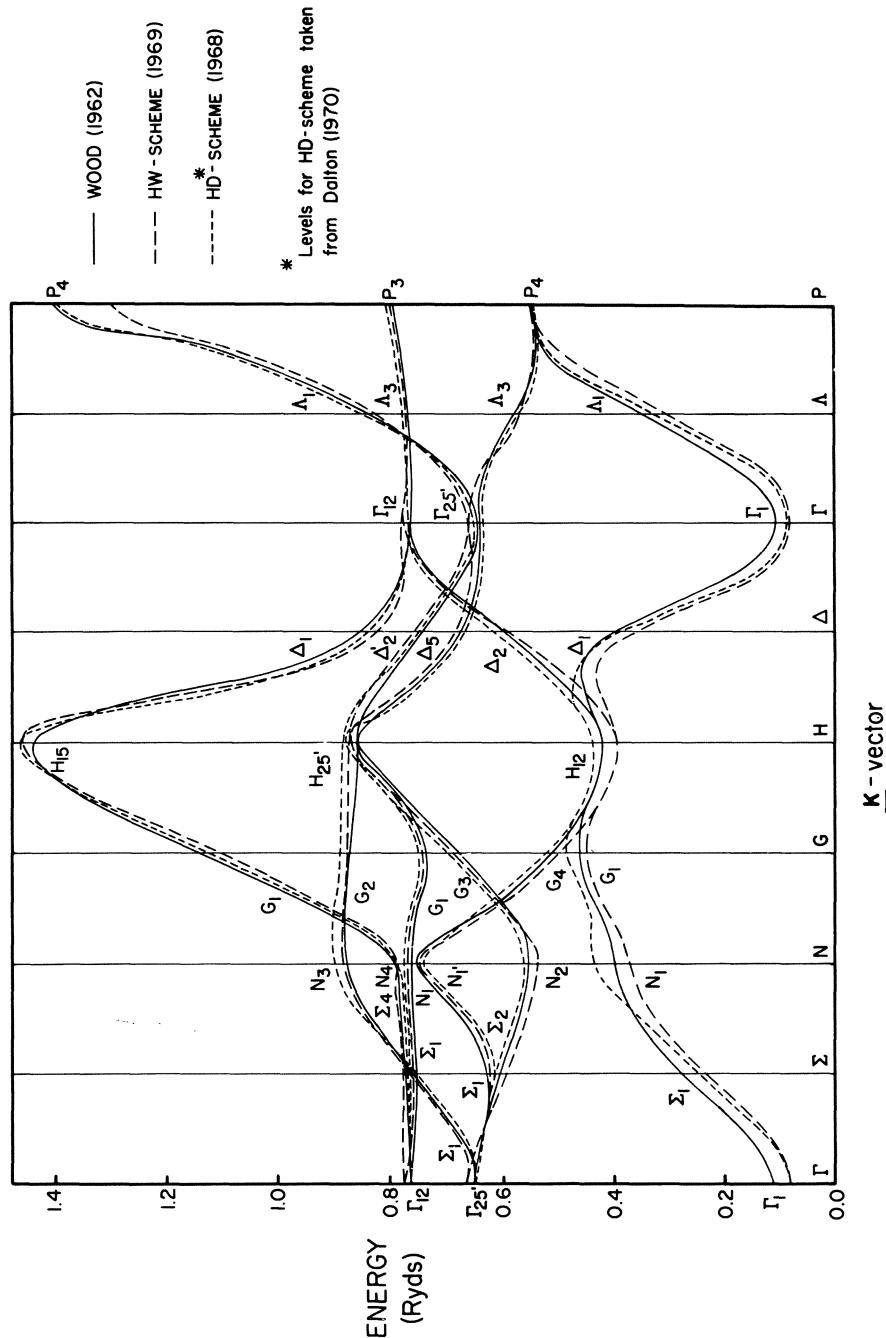
ENERGY LEVELS	BURDICK (1963) COPPER	HD [†] $\beta=0.8$ <u>15K</u>	P $\beta=0.7$ <u>4K</u>	ENERGY LEVELS	IRON WOOD (1962)	HD [†] $\beta=0.8$ <u>19K</u>	HW $\beta=5.0$ $\eta=1.0$ <u>19K</u>
X(108) 1	0.163	0.185	0.170	$\Gamma(000)$ 1	0.108	0.078	0.076
3	0.200	0.211	0.214	25'	0.640	0.639	0.652
2	0.399	0.401	0.395	12	0.762	0.761	0.773
5	0.412	0.413	0.406				
5	0.412	0.413	0.406	H(080) 12	0.413	0.435	0.386
4'	0.704	0.701	0.716	25'	0.850	0.869	0.860
				15	1.435	1.458	1.455
L(444) 1	0.164	0.192	0.178				
3	0.297	0.298	0.299	$\Sigma(220)$ 1	0.275	0.254	0.236
3	0.297	0.298	0.301	2	0.594	0.589	0.591
3	0.401	0.404	0.398	1	0.618	0.612	0.616
3	0.401	0.404	0.398	3	0.751	0.743	0.754
2'	0.510	0.501	0.517	1	0.753	0.746	0.755
				4	0.773	0.769	0.778
$\Delta(040)$ 1	0.080	0.105	0.098				
2'	0.243	0.254	0.258	N(440) 1	0.399	0.428	0.369
1	0.328	0.330	0.332	2	0.545	0.548	0.536
5	0.342	0.349	0.354	1'	0.748	0.744	0.748
5	0.342	0.349	0.354	1	0.765	0.763	0.768
2	0.371	0.374	0.375	4	0.787	0.787	0.789
				3	0.877	0.889	0.875

[†]These levels are taken from Dalton (1970) and differ slightly from those given in Hubbard and Dalton (1968). (cf. footnote in Dalton (1970)).

	COPPER			IRON (f.c.c.)			IRON (b.c.c.)		
Number of K-vectors	4	15	27	4	15	27	5	14	19
Number of R-vectors	12	18	42	12	18	42	14	14	26
β	0.5	0.8	0.8	0.5	0.8	0.8	0.5	0.6	0.8
$0.00 \leq \Delta\varepsilon \leq 0.01$	26	34	40	9	26	32	10	21	30
$0.01 < \Delta\varepsilon \leq 0.02$	1.3	11	6	15	12	12	9	8	16
$0.02 < \Delta\varepsilon \leq 0.03$	9	3	2	12	8	4	9	14	7
$0.03 < \Delta\varepsilon$	0	0	0	12	2	0	26	11	1
r.m.s. error	0.014	0.010	0.009	0.028	0.014	0.011	0.048	0.027	0.014

[†] Taken from Hubbard and Dalton (1968) (c.f. Dalton (1970))

Table VI. Distribution of Errors for Approximate KKR Scheme [†]



A COMPARISON OF THE BAND-STRUCTURES FOR (b.c.c.) IRON

Band-Structure Scheme	Hubbard and Dalton (1968)	Hum and Wong (1969)	Pettifor (1969)	Hubbard (1969)
Size of Conduction Block K - vectors	15 (f.c.c.) 19 (b.c.c.)	15 (f.c.c.) 19 (b.c.c.)	4 (f.c.c.) *	4 (f.c.c.) 7 (b.c.c.)
Convergence Parameters	β (≈ 0.8)	β , η (≈ 5.0) (1.0)	β (≈ 0.6)	NONE
Fraction of Levels with errors $\Delta\epsilon \leq 0.02$ ryds.	94% (f.c.c.) 85% (b.c.c.)	- 76% (b.c.c.)	97% (f.c.c.) -	100% (f.c.c.) 100% (b.c.c.)
Root Mean Square Error (50-70 Levels)	0.010 (f.c.c.) 0.014 (b.c.c.)	- 0.015 (ryds) (b.c.c.)	0.0009 (ryds) (f.c.c.) -	0.007 (f.c.c.) 0.006 (b.c.c.)
Average Time Per k - point (IBM 360/65)	~ 1.0 secs.	-	~ 0.2 secs.	~ 1.0 secs.
Applications	Cu,Ni, (1) Fe,Cr (1)	Fe,Cu (2) (h.c.p.)	Cu,Fe(f.c.c.)	Cu,Fe,Cr,Co (3)

- 1) Dalton (1970);
 2) Wong (1970);
 3) Wakoh (1970).

* A dashed entry indicates that the omitted data was not available at the time of writing.

Table VII. APPROXIMATE KKR SCHEMES FOR TRANSITION METALS
 (f.c.c. refers to copper; b.c.c. refers to iron)

made), still has levels which are strongly parameter dependent (c.f. L₁ in Table III). The remaining two schemes unfortunately require larger size matrices to obtain an accuracy of 0.01-0.02 ryds. Hum and Wong's (1969) scheme suffers also from the disadvantage of having two adjustable parameters instead of one

The main conclusion to be drawn from this analysis is that although there are several ways of deriving approximate KKR schemes, all of which are capable of producing band-structures accurate to 0.01-0.02 rydbergs, the scheme developed by Hubbard (1969) appears to be the most successful currently in use.

SUMMARY AND CONCLUSIONS

In this paper we have reviewed and critically compared recent approximate methods for computing transition metal band-structures. The factor common to these methods is that they are based upon a particular representation for the scattering phase-shifts and the exact KKR theory. Besides outlining the essential steps in the derivation of approximate KKR band-structure schemes (section 3) we have considered in detail the schemes developed by Hubbard and Dalton (1968), Pettifor (1969), Hum and Wong (1969) (section 4). If formal simplicity, speed of computation, and accuracy are the criteria whereby we judge the success of a band-structure scheme, then the scheme proposed by Hubbard (1969) appears to be the most successful developed to date.

The attempts at simplifying the exact KKR theory discussed in this paper have proved fruitful in two ways. First they have lead to a variety of computational schemes which, although inherently less accurate than the exact KKR theory, are faster and more flexible. However it should be emphasized that the four schemes presented here are meant to be indicative of the practicability of such schemes, and that many alternative schemes can be derived by developing the theory of section 3 of this paper. Second, they have provided a formal justification for the successful interpolation schemes ('model Hamiltonians') proposed by Hodges e.a. (1966) and Mueller (1967). In particular they have shown that the ten to fifteen adjustable parameters needed to define their model Hamiltonians can be expressed in terms of a much smaller number, and furthermore, have shown how these parameters are related to the scattering properties (i.e. phase-shifts) of the electrons.

It may be inquired as to whether equivalent schemes may be derived using the exact APW or OPW schemes. In view of the equiva-

lence of the KKR and APW schemes (Johnson(1966)) one would expect it to be relatively straight forward to derive approximate APW schemes along the lines of section 3. In practice the surface Bessel function term $j_1(|k - k'| |r_i|)$ may cause some difficulty. On the other hand it is not clear how to modify the OPW scheme since the matrix elements are already independent of energy. A suitable form for the resonant wave-function in conjunction with the Löwdin (1951) technique might prove a useful starting point.

Finally we remark that future work on these schemes should be directed towards the eigenvectors (rather than the eigenvalues) since these are essential for calculating matrix elements associated with dielectric constants, susceptibilities and other physical properties. The extension of these schemes to include relativistic effects (e.g. Dalton (1968)) is also worthwhile in order to study metals in the second and third transition series. To take advantage of the flexibility of the model Hamiltonian approximation, effort should also be directed towards calculating the physical properties of transition metals by analogy with the use of pseudo potentials for calculating the properties of non-transition metals (Harrison (1966)).

Acknowledgements

It is a pleasure to acknowledge many valuable conversations with Dr. J. Hubbard relating to the material presented in this paper.

REFERENCES

- Asano S. and Yamashita J. (1967), "Band Theory of Antiferromagnetic Chromium," *J. Phys. Soc. Japan*, 23, pp. 714-736.
- Austin B. J., Heine V., and Sham L. J. (1962), "General Theory of Pseudopotentials," *Phys. Rev.*, 127, pp. 276-282.
- Burdick G. A. (1963), "Energy Band-structure of Copper," *Phys. Rev.*, 129, pp. 138-149.
- Carslaw H. S. and Jaeger J. C. (1959), "The Conduction of Heat in Solids," (Clarendon Press).
- Cooper B. R., Kreiger E. L., and Segall B. (1970), "Phase-shift Parameterisation: Band Structure for Ag.," (published in this volume).
- Dalton N. W. (1968), "Adaptation of the Kapur-Peierls Theory of Resonances to Relativistic Band-structures," *Physics Letters*, 28A, pp. 280-281.

Dalton N. W. (1970), "Approximate Calculation of Electronic Band-structures. IV Density-of-states for Transition Metals," J. Phys.C (Solid State Phys.) (to be published).

Dalton N. W. and Hubbard J. (1968) "Remarks on the Approximate Calculation of Electronic Band-structures" A.E.R.E Rep. No. TP/237.

Harrison W. A. (1966) "Pseudopotentials in the Theory of Metals" (Benjamin, New York).

Heine V. (1967) "s-d Interaction in Transition Metals" Phys. Rev. 153, pp. 673-682.

Heine V. (1969) "Electronic Structure of Metals" (Ch. 1 "The Physics of Metals" Ed. Ziman - Camb. Univ. Press).

Herman F. and Skillman S. (1963) "Atomic Structure Calculations" (Prentice-Hall Inc.).

Hodges L., Ehrenreich H. and Lang N. D. (1966) "Interpolation Scheme for Band-structure of Noble and Transition Metals: Ferromagnetism and Neutron Diffraction in Ni" Phys. Rev. 152, pp. 505-526.

Hubbard J. (1967) "The Approximate Calculation of Electronic Band-structures" Proc. Phys. Soc. 92, pp. 921-937.

Hubbard J. and Dalton N. W. (1969) "The Approximate Calculation of Electronic Band-structures II. Application to Copper and Iron" J. Phys. C. (Solid State Phys.) [2], pp. 1637-1649.

Hubbard J. (1969) "The Approximate Calculation of Electronic Band-Structures III" J. Phys. C. (Solid State Phys.) [2], pp. 1222-1229.

Hum D. and Wong K. C. (1969) "Calculation of Transition Metal Band-structures" J. Phys. C. (Solid State Phys.) [2], 2, pp. 833-840.

Jacobs R. L. (1968) "The Theory of Transition Metal Band-structures" J. Phys. C. (Proc. Phys. Soc.) [2], 1, pp. 492-506.

Johnson K. H. "Relationship Between the Augmented-Plane-Wave and Korringa-Kohn-Rostoker Methods in Band Theory" Phys. Rev. 150, pp. 429-440.

Kapur P. L. and Peierls R. E. (1938) "The Dispersion Formula for Nuclear Reactions" Proc. Roy. Soc. A, 166, pp. 277-295.

- Kohn W. and Rostoker N. (1954) "Solution of the Schrodinger Equation in Periodic Lattices with an Application to Metallic Lithium" Phys. Rev. 94, pp. 1111-1120.
- Korringa J. (1947) "On the Calculation of a Bloch Wave in a Metal" Physica 13, pp. 392-404.
- Löwdin P. (1951) "A note on the Quantum-Mechanical Perturbation Theory" Journal Chem. Phys. 19, pp. 1396-1401.
- Mueller F. M. (1967) "Combined Interpolation Scheme for Transition and Noble Metals" Phys. Rev. 153, pp. 659-669.
- Newton R. G. (1966) "Scattering Theory of Waves and Particles" (McGraw-Hill).
- Pettifor D. G. (1969) "An Energy-independent Method of Band-Structure Calculation for Transition Metals" J. Phys. C. (Solid State Phys.) [2], 2, pp. 1051-1069.
- Schiff L. I. (1955) "Quantum Mechanics" (McGraw-Hill).
- Slater J. C. (1937) "Wave Functions in a Periodic Potential" Phys. Rev. 51, pp. 846-851.
- Slater J. C. (1966) "Green's-Function Method in the Energy-Band Problem" Phys. Review 145, pp. 599-602.
- Slater J. C. and Koster G. F. (1954) "Simplified LCAO Method for the Periodic Potential Problem" Phys. Rev. 94, pp. 1498-1524.
- Wakoh S. (1970) "Generalization of Hubbard's Model Hamiltonian Method: Application to h.c.p. cobalt" J. Phys. C. (Solid State Phys.) (to be published).
- Wong K. C. (1970) "Calculation of Transition Metal Band-structures. II" J. Phys. C. (Solid State Phys.) [3], 2, pp. 378-386.
- Wood J. H. (1962) "Energy Bands in Iron via the Augmented-Plane-Wave Method" Phys. Rev. 126, pp. 517-527.
- Ziman J. M. (1965) "The T-matrix, the K-matrix, d-bands and k -dependent Pseudo Potentials in the Theory of Metals" Proc. Phys. Soc. 86, pp. 337-353.

OPTICAL PROPERTIES OF THE ALKALIS USING THE KKR-Z METHOD*

A. R. Wilson, G. Dresselhaus and C-Y. Young

Lincoln Laboratory, Massachusetts Institute of Technology

Lexington, Massachusetts 02173

We have calculated the interband absorption of potassium and sodium by means of a KKR-Z pseudopotential^{1,2} obtained from a fit to the experimental Fermi surface data. Previous calculations of the optical properties of the alkali metals have relied either on a local pseudopotential³ or else introduced non-locality as arising solely from the corrections due to the atomic core.⁴ However, it has been demonstrated that the local pseudopotential formulation⁵ is inadequate to account for the measured Fermi surface of K. It has also been shown⁶ that the pseudopotential that appears in the optical matrix element will incorporate the effects of many-body correlations only if it is consistent with the experimental data of Fermi surface measurements (such as dHvA results). The method used here was applied with this requirement being satisfied. Recent theoretical results⁶ and the successful application of a phenomenological model for the optical properties of Al⁷ suggest that a one-electron potential which has been fitted to the experimental Fermi surface should implicitly contain electron-electron effects in the optical absorption. The measurement of the interband absorption of the alkalis, although difficult to carry out, is nevertheless of some importance as the absorption depends critically on the pseudopotential. The shapes of the sodium and potassium Fermi surfaces have been measured with great precision by de Haas-van Alphen experiments.^{8,9,5} These experiments are sufficiently accurate to show that the local pseudopotential of Ashcroft, while satisfactory in the case of Na, is unable in K to account for the electronic energy levels in the neighborhood of the Fermi energy.^{5,10,11} We have investigated the optical properties of these metals using a more general pseudopotential to be described below.

*This work was sponsored by the Department of the Air Force.

In this work the Ziman form of the KKR-method non-local pseudo-potential is used. It was selected instead of other methods for several reasons. Unlike the ordinary KKR method the wave-function external to the muffin-tin radius R_S is expressed as a sum of plane waves - a feature of some value in the calculation of the optical properties. This property is also true with the augmented-plane-wave (APW) method. However, the plane wave matrix elements in the APW procedure are both more lengthy to compute and, more important, are quite sensitive to energy dependence. To see the reason for this preference more clearly, let us consider briefly both methods. At points outside the muffin-tin radius $r > R_S$ the wavefunction of state \vec{k} in either scheme is expressed in terms of plane waves, given by

$$\psi_{\vec{k}}(\vec{r}) = \sum_{\vec{G}} \phi_{\vec{k} + \vec{G}} e^{i(\vec{k} + \vec{G}) \cdot \vec{r}} \quad (1)$$

where \vec{G} are the (bcc) reciprocal lattice vectors. For a spherically symmetric potential of range R_S the components $\phi_{\vec{k} + \vec{G}}$ of the external wavefunction must satisfy

$$\sum_{\vec{G}'} \phi_{\vec{k} + \vec{G}'} \left\{ \left[(\vec{k} + \vec{G})^2 - \kappa^2 \right] \delta_{\vec{G}, \vec{G}'} + \langle \vec{k} + \vec{G} | \Gamma | \vec{k} + \vec{G}' \rangle \right\} = 0 \quad (2)$$

where we have used atomic units (κ^2 is the energy). The matrix elements in the KKR-Z method are

$$\begin{aligned} \langle \vec{k} + \vec{G} | \Gamma_{KKR-Z} | \vec{k} + \vec{G}' \rangle &= \frac{-3}{\kappa R_S^3} \sum_{\ell} (2\ell+1) \tan \gamma_{\ell} P_{\ell}(\cos \theta) \\ &\times \frac{j_{\ell}(|\vec{k} + \vec{G}| R_S) j_{\ell}(|\vec{k} + \vec{G}'| R_S)}{j_{\ell}^2(\kappa R_S)} \end{aligned} \quad (3)$$

where the j_{ℓ} are spherical Bessel functions, the P_{ℓ} are Legendre polynomials and θ is the angle between $\vec{k} + \vec{G}$ and $\vec{k} + \vec{G}'$. The γ_{ℓ} are phase shifts which characterize the muffin-tin potential. The corresponding APW matrix elements are

$$\begin{aligned} \langle \vec{k} + \vec{G} | \Gamma_{APW} | \vec{k} + \vec{G}' \rangle &= -\frac{3}{R_S} \left\{ \sum_{\ell=0}^{\infty} (2\ell+1) \frac{R_{\ell}(R_S)}{R_{\ell}(R_S)} j_{\ell}(|\vec{k} + \vec{G}| R_S) j_{\ell}(|\vec{k} + \vec{G}'| R_S) \right. \\ &\times P_{\ell}(\cos \theta) - \left[(\vec{k} + \vec{G}) \cdot (\vec{k} + \vec{G}') - \kappa^2 \right] \frac{j_{\ell}(|\vec{G} - \vec{G}'| R_S)}{|\vec{G} - \vec{G}'| \} \right\} \end{aligned} \quad (4)$$

where the R_ℓ are radial wavefunctions which, for a spherically symmetric muffin-tin potential, can be expressed in terms of phase shifts η_ℓ^{APW} analogous to the η_ℓ .

Unlike the KKR-Z scheme the matrix elements Γ_{APW} are separated into two sets of terms which partially cancel each other. This accounts for the greater sensitivity, at fixed η_ℓ , of the Γ_{APW} on small changes of the energy (κ^2). This partial cancellation also means that the number of terms in the sum over ℓ is not determined by the number of phase shifts used (3 in our calculations) but rather by the convergence of the sum itself - typically 10 or 12 terms are needed for 1% accuracy. The generation of the j_ℓ becomes increasingly slower at large ℓ . In addition the Neumann functions n_ℓ which are included in the definition (cf¹) of the radial functions R_ℓ need to be calculated in the APW method.

The determinant resulting from the conditions (2) was solved for $\kappa^2(\vec{k})$ using 19 plane waves $\vec{G} = (000)$, (110) , (200) , and their equivalents.¹⁰ The solution at the Fermi surface $\kappa^2(\vec{k}) = \kappa^2_F$ was obtained by successive reductions $N^2 \rightarrow (N-1)^2$ in the size of the determinant such that the $N = 1$ resultant was the eigenvalue desired. This is always possible because the off-diagonal Γ 's are much smaller than the diagonal entries $\Gamma/\kappa^2_F \sim 0.1$. It is important in setting up the determinant that the $\vec{G} = \vec{G}' = \vec{0}$ entry be situated so that at no stage in the reduction process division by this term can take place. Numerical error may otherwise be large.

Using as phenomenological parameters the phase shifts η_0 , η_1 , η_2 and the Fermi energy - $\langle \vec{k}+\vec{0} | \Gamma | \vec{k}+\vec{0} \rangle$ the fractional change in the Fermi radius from the free-electron value was calculated and mean-square-fitted to the experimental results of Lee⁸ (in Na) and Lee and Falicov⁵ and Shoenberg and Stiles⁹ (in K). The fit shown in Fig. 1 is the fractional area change $\Delta A(\vec{k})/A_0$ for potassium. The model fits the experiments to within the experimental error. The phase shifts obtained appear in Table I. The d-phase shift η_2 is of particular importance for K¹⁰ but

TABLE I

Parameters obtained from fitting the Fermi surface to experimental data. There is no unique choice of parameters and a second set was obtained for K. Set (a) was used in the calculation of the absorption.

	η_0	η_1	η_2	Fermi Energy (ryd)	R_s (in a. u.)
Na	0.03542	-0.00417	-0.00120	0.23607	2.40
K(a)	0.00216	0.00244	0.00203	0.15389	2.53
K(b)	0.00282	-0.00034	-0.00326	0.15315	2.53

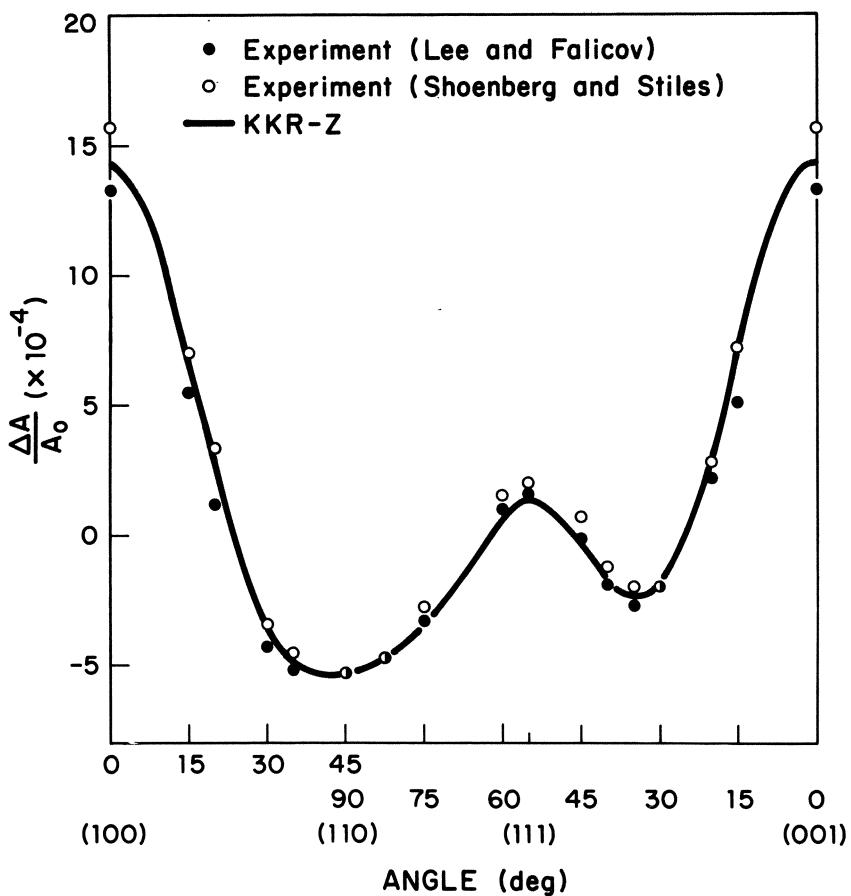


Fig. 1. Fractional area change $\Delta A / A_0$ from the free electron extremal cross-sectional area of potassium. The parameters used to obtain this fit are set (a) in Table I.

somewhat smaller (compared with the s-phase shift) for Na. In either case the Fermi energy turned out to be very close to the free-electron value - a result which was not true for local pseudopotential calculations.¹¹ Also the effective mass calculated, neglecting energy dependence of the pseudopotential, was found to be isotropic over the Fermi surface to within 2%.

One additional parameter in the fitting which was held fixed throughout is the radius R_s . For both Na and K it was taken to be the respective atomic core radius. By this choice the planewave representation of the wavefunction described earlier deviates from the true wavefunction over a fraction $(R_s/R_{\text{Wigner-Seitz}})^3 \sim 0.12$ of the entire atomic volume in K, for instance. We have not attempted to include the core corrections on this wavefunction in this small volume.⁴

The optical interband conductivity was calculated using Butcher's result³ generalized to include lifetime effects

$$\sigma_{ii} = \frac{2\pi e^2 n\omega}{i3m^2} \sum_{\vec{G}} \int_{|\vec{k}| \leq k_F} d\vec{k} \frac{|\langle \phi_{\vec{k}} | P_i | \phi_{\vec{k}+\vec{G}} \rangle|^2}{E(\vec{k}+\vec{G}) - E(\vec{k})} \\ \times \frac{1}{[E(\vec{k}+\vec{G}) - E(\vec{k})]^2 - \hbar^2 (\omega + i/\tau)^2} \quad (5)$$

Here P_i is the momentum operator which, with plane waves, has eigenvalues $\vec{k}_i + \vec{G}_i$. $E(\vec{k}+\vec{G})$ is the energy of band G at momentum \vec{k} and ω is the light frequency. In this situation we must carry out a volume integral over that region of k -space enclosed by the Fermi surface. A Monte Carlo integration scheme was used, the coefficients $\phi_{\vec{k}+\vec{G}}$ being evaluated from

the diagonalization of the 19×19 matrix in (2). The Brillouin zone was divided into 8 symmetrically inequivalent cubes. The number of points randomly taken in each cube was weighted by its relative contribution to $\sigma(\omega)$. Thus, only a small volume of the zone need be sampled extensively, since the interband contribution in the alkali metals is mainly from the region about the Σ -axis. The inclusion of the relaxation time τ render our results more suitable for comparison with experimental data. Also we can compute without additional effort the conductivity at an arbitrary number of values for ω . If we replace the resonant denominator by an energy delta-function and reduce the dimensions of the integral by one, some of that advantage is lost in calculating the spectrum $\sigma(\omega)$ as a different surface of constant energy difference is required each time. With $\tau \sim 2 \times 10^{-14}$ sec. the Monte Carlo calculations (Figs. 2, 3) converged with the error bars shown after 1200 points had been selected - the entire process requiring about one hour of IBM 360-67 time.

As indicated in Table I, a second set of parameters was found which provided a fit to the K-Fermi surface. That there is no unique set is hardly surprising when we consider the complexity of the dependence on the phase-shifts in the Γ -matrix elements(3). This second set of parameters however provided an optical absorption spectrum rather similar to the one shown here.

The absorption maximum in sodium is about 3 eV and with magnitude $\text{Re } \sigma \sim 3 \times 10^{13} \text{ sec}^{-1}$ for $1/\tau = 1.5 \times 10^{14} \text{ sec}^{-1}$ which is in agreement with the results of Smith¹² as analyzed by Powell.¹³ The absorption in potassium agrees with Smith's results in the position of the maximum but not in magnitude - our results are about four times smaller.¹⁴ A possible source of the discrepancy lies in our neglect of the energy dependence of the phenomenological phase-shifts. These were obtained at the Fermi energy and to complete the phenomenology their energy-dependence must be sought from further experimental results - possibly the optical

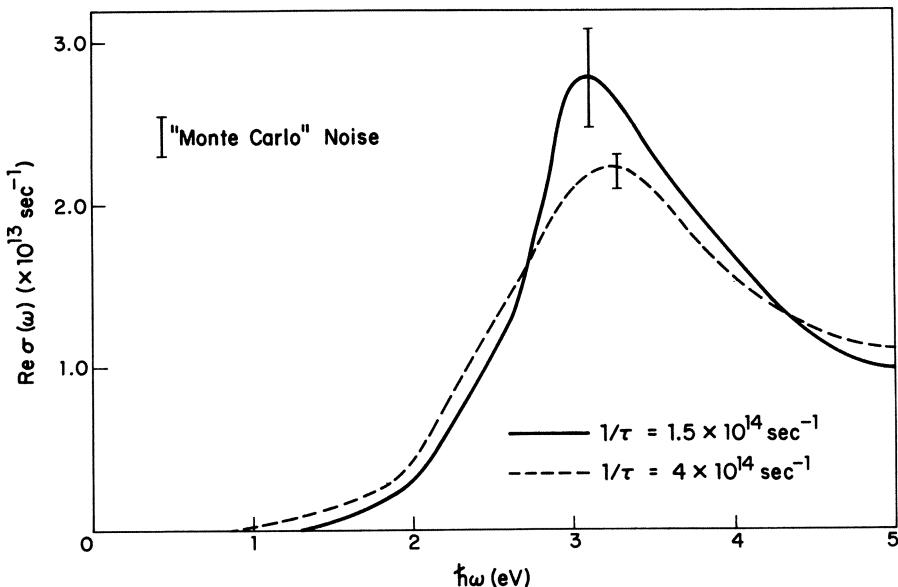


Fig. 2. The optical absorption $\text{Re } \sigma(\omega)$ in sodium. Calculated with the two different collision rates indicated.

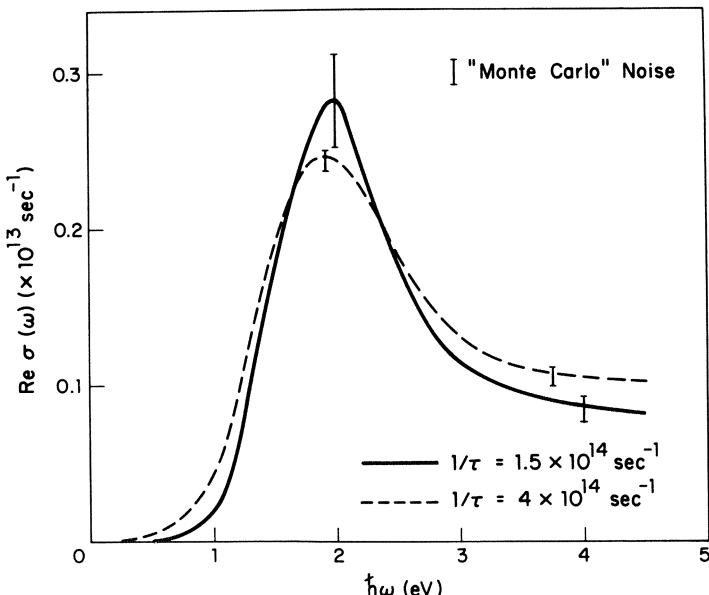


Fig. 3. The optical absorption $\text{Re } \sigma(\omega)$ in potassium calculated with the two different collision rates indicated.

absorption data itself. Therefore the optical data may provide useful information on the energy-dependence of the phase shifts. In contrast sodium presents no such difficulties. Even in fitting the Fermi surface, a local pseudopotential is satisfactory.^{5,8}

References

1. G. J. Morgan, Proc. Phys. Soc. 89, 365 (1966).
2. J. M. Ziman, Proc. Phys. Soc. 86, 337 (1965).
3. P. N. Butcher, Proc. Phys. Soc. 64, 765 (1951).
4. J. A. Appelbaum, Phys. Rev. 144, 435 (1966); A. O. E. Animalu, Phys. Rev. 163, 557 (1967).
5. M. J. G. Lee and L. M. Falicov, Proc. Roy. Soc. A304, 319 (1968).
6. C-Y. Young, Bull. Am. Phys. Soc. Series II, 15, 368 (1970) and also to be published.
7. G. Dresselhaus, M. S. Dresselhaus and D. Beaglehole, Proc. Density of States Conf. (1969).
8. M. J. G. Lee, Proc. Roy. Soc. A295, 440 (1966).
9. D. Shoenberg and P. J. Stiles, Proc. Roy. Soc. A281, 62 (1964).
10. M. J. G. Lee, Phys. Rev. 178, 953 (1969). In his fit of the Fermi surface, Lee has used the APW method with 30 plane waves. We have used as few as possible (19) to allow us to obtain a reasonably accurate fit, since the computation time involved in the calculation of the optical absorption would be prohibitive with a larger matrix. Because a different number of plane waves were used, the phase shifts obtained by us cannot be directly compared with those of Lee.
11. N. W. Ashcroft, Phys. Rev. 140, A935 (1965).
12. N. V. Smith, Phys. Rev. 183, 634 (1969).
13. C. J. Powell, Bull. Am. Phys. Soc. Series II, 15, 576 (1970) and also to be published.
14. A major contribution to the volume integral comes from the matrix element $\langle \vec{k} + \vec{G}_{110} | \Gamma | \vec{k} \rangle$ in the direction of backscattering
 $\cos \theta_{\vec{k}, \vec{k} + \vec{G}_{110}} = \pi$. The value of the corresponding matrix element in Lee's calculation¹⁰ is about the same as ours. We would consequently expect that the optical absorption in K calculated by the two methods be similar.

Discussion

Saffren: Have you, in fact, neglected the core contributions to the conductivity?

Wilson: We have neglected the core corrections largely on the evidence of previous work on this question. Appelbaum⁴ calculated for Na the optical absorption by solving numerically a determinant and using wavefunctions obtained by orthogonalizing plane waves (corresponding to the local pseudopotentials V_{110}, V_{200}) to core states. He found a reduction by 80% from the original result of Butcher³ where no core corrections had been made.

However, the discrepancy is not due simply to the core - it is mainly due to the choice of relative sign of the two pseudopotentials V_{110} and V_{200} (only V_{110} was used by Butcher). Animalu⁴ repeated the calculation using V_{110} alone and found an enhancement over the Butcher result by 40%. However, Animalu did not solve a determinant, but used perturbation theory instead. His result may be inaccurate on this account since Appelbaum found that Butcher's result for the absorption, also derived from perturbation theory, was decreased by an order of magnitude if a determinant were solved instead. We concluded from this that if the full determinant had been used and in addition an appropriate choice of V_{110} , V_{200} made (namely, with opposite signs) then the core corrections should be considerably smaller than claimed by these authors.

Callaway: I should have thought that in the core region the electrons, having relatively large momenta, would contribute substantially to the optical matrix elements.

Wilson: That may be so. However, the core occupies only an eighth of the atomic volume.

Dresselhaus: I would like to point out that optical matrix elements used here appear to be consistent with the effective mass sum rule as applied to the nearly free-electron band structure.

Williams: I agree with the speaker's remarks. It appears from calculations by various authors (e.g. Ehrenreich) that the core corrections are on the order of 10%.

3. LCAO METHODS IN THE BAND PROBLEM

DISCRETE VARIATIONAL METHOD FOR THE ENERGY BAND PROBLEM*

D. E. Ellis[†]

Physics Department, Northwestern University
Evanston, Illinois 60201

and

G. S. Painter**

Quantum Theory Project, University of Florida
Gainesville, Florida 32601

One may well ask, why set forth yet another scheme to solve the periodic one-electron Schrodinger equation for a crystal? Our chief purpose is to develop a flexible and efficient approach which can be applied to systems treated only with difficulty by other methods; in particular the crystal potential should be unrestricted in form. Since solution of the band problem is only a starting point for evaluating crystalline properties, we require a method in which the wave functions are accessible and easily manipulated. The numerical-analytic scheme presented here is intended as an example of a large class of methods which can meet these needs.

Discrete variational methods (DVM) have a long history; among the applications to wave mechanics we may mention the work of Frost et al. on molecular structure¹, Carlson on the hydrogen molecule ion², and Bell et al. on the nuclear three body problem.³ Common to these methods is the selection of a discrete set of sample points in co-ordinate space, definition of an error functional connected with the Schrodinger equation, and minimization of the error functional over the sample points by a variational procedure. While de-

* Research supported by the National Science Foundation, the Air Force Office of Scientific Research and by the Advanced Research Projects Agency through the Northwestern University Materials Research Center.

† Alfred P. Sloan Research Fellow

** Now at Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

tails of application vary from one problem to another, the DVM is clearly a general and powerful method. A detailed presentation of our method has been submitted for publication elsewhere; here we outline the most important results. In the following paper we discuss our practical experience with applications of the method.

I. The One-Electron Model and Variational Method

We start with the familiar one-electron Hamiltonian

$$H(\mathbf{r}_\sim) = -\frac{1}{2}\nabla^2 + V(\mathbf{r}_\sim) \quad (1)$$

(in Hartree atomic units)

with periodic potential $V(\mathbf{r})$. It is important to note at the outset that our method depends in no way upon the specific form of the potential; however, for the sake of definiteness we adopt the usual superposition model as a starting point,

$$V(\mathbf{r}) = \sum_{v,i} v_i (\mathbf{r} - \mathbf{R}_v - \mathbf{u}_i) + v_x(\mathbf{r}) \quad (2)$$

This potential consists of the sum of atomic Coulomb potentials v_i and a local exchange approximation, v_x , for which we take the scaled "1/3" form⁴

$$v_x^{\uparrow,\downarrow}(\mathbf{r}) = -3\alpha \left(\frac{3}{4\pi} \rho^{\uparrow,\downarrow} \right)^{1/3} \quad (3)$$

Here the arrows denote either spin, $\rho^{\uparrow,\downarrow}$ is the crystal spin density, and α is a scaling parameter.⁵ The assumed initial crystal spin density is found by summing atomic densities, ρ_i ,

$$\rho^{\uparrow,\downarrow} = \sum_{v,i} \rho_i^{\uparrow,\downarrow} (\mathbf{r} - \mathbf{R}_v - \mathbf{u}_i) \quad (4)$$

This model can, of course, be made self-consistent, and is the basis of many present day band calculations.

We may now seek approximate solutions of the Schrodinger equation by expanding the eigenfunctions as

$$\psi_i(\mathbf{k},\mathbf{r}) = \sum_j \chi_j(\mathbf{k},\mathbf{r}) c_{ji}(\mathbf{k}) \quad (5)$$

where the basis functions χ_j are a given set of Bloch functions, and the coefficients C_{ji} are parameters to be determined variationally. We may define the error at any point in space, for a given state, as

$$d_i(r) = (H - E_i) \psi_i(r) \quad (6)$$

and devise any number of criteria for the accuracy of the expansion (5) by defining and minimizing appropriate error functionals. We shall choose to minimize the weighted mean error, in a manner entirely analogous to the Rayleigh-Ritz variational method of continuum theory. Thus we define matrix elements by writing down mean error functionals as

$$\Delta_{ij} = \langle \psi_i | (H - E) | \psi_j \rangle = \sum_m w(r_m) \psi_i^*(r_m) (H - E) \psi_j(r_m) \quad (7)$$

where $w(r)$ is a (positive) weight function, and the summation is over a discrete set of sample points r_m . Then finding the extremum (hopefully a minimum) by requiring

$$\frac{\partial \Delta_{ij}}{\partial C_{pq}} = 0 \quad (8)$$

we are led to the familiar secular matrix equation

$$H(k)C(k) = E(k)S(k)C(k) \quad (9)$$

determining the band structure $E_i(k)$ and the eigenvector coefficients. However, the matrix elements of (9) are defined by (7) as

$$H_{ij} = \langle \chi_i | H | \chi_j \rangle \quad S_{ij} = \langle \chi_i | \chi_j \rangle \quad (10)$$

and are not integrals, as required in the Rayleigh-Ritz procedure. Thus any operators or functions may be included in the Hamiltonian, so long as one can evaluate $H\chi_j(r)$ at specific points in space. This freedom is obtained at a price; the eigenvalues E_i of (9) are not in general upper bounds to the true eigenvalues of the Schrödinger equation, which is a great merit of the integral Rayleigh-Ritz method. Only experience and careful construction of weight functions $w(r)$ and sets of sample points r_m can show whether the

method is competitive in accuracy for a given level of computational effort. Certainly, one would also wish to investigate other error criteria such as least square, min-max, etc.

II. Selection of Weight Function and Sample Points

The matrix elements, as defined by (7), depend both upon the weight function $w(\underline{r})$ and the number and distribution of sample points \underline{r}_m . One simple choice is to view $w(\underline{r})$ as the local volume per point, and to view the point summation as a numerical integration procedure; in this way we may require that the matrix elements (10) converge toward the integral Rayleigh-Ritz matrix elements as the number of sample points is increased. This procedure has considerable merit, since the solutions of (9) should then converge toward the corresponding integral Rayleigh-Ritz solutions which are an upper bound to the exact eigenvalues. This expectation is borne out in practice, and our remaining discussion is presented from this point of view. However, we should note that freedom in the choice of weights and points can also be used to good advantage to minimize the error functionals in specific regions of space. This can be of importance if one has limited variational freedom, for example, and wishes to obtain accurate valence levels at the expense of less accurate core states.

As we have mentioned, the weight function $w(\underline{r})$ can be interpreted as the local volume per point, the inverse of the sampling point density function, and this defines an integration rule. The point density function which is used to generate the sample points \underline{r}_m represents a compromise between computational convenience and the desire to produce a stable rapidly convergent integration rule for the class of functions under consideration. Since we have to treat operators H with Coulomb singularities at each nucleus, and basis functions χ_i , with some nodal structure around each atomic site, it is apparent that a uniform point density will be far from optimum. We have found two procedures which work rather well:

1. Divide the unit cell into "muffin tin" regions of atomic spherical volumes, and the interatomic region. Treat each region separately, by three dimensional numerical quadrature, taking care to treat the nuclear Coulomb singularity properly. Uniform (Simpson rule) sampling in the interatomic region, and random sampling (with uniform radial density) inside the atom spheres has proved adequate; more elaborate schemes may prove more efficient.
2. Define the point density function $\mu(\underline{r})$ as a sum of simple density functions based on each atom in the crystal

$$\mu(\underline{r}) = w^{-1}(\underline{r}) = \sum_{v,i} \mu_i (\underline{r} - \underline{R}_v - \underline{u}_i) \quad (11)$$

Note that μ should diverge more rapidly than $|r - R|^{-1}$ as one approaches the nucleus at R . The results are relatively insensitive to the particular form of μ_j ; we tend to use a Fermi radial distribution with characteristic radius near the "muffin-tin" radius in the crystal. The chief advantage of this atomic density sampling model is that one builds up the sample density along the atomic bonds, as well as about the individual atoms.

Finally, we mention two useful symmetrization procedures:

1. As defined in (10) the matrices H and S will not be Hermitean in general; however, one can conveniently use the symmetrized form $H'_{ij} = (H_{ij} + H_{ji})/2$ (similarly for S) to ensure that the eigenvalues of (9) are real.
2. It is often inconvenient to construct basis functions x_j which form a basis for the irreducible representations of the crystal space group (and hence a partial factorization of (9)) by the usual projection operator techniques. This can be accomplished alternatively by choosing a symmetrized point set r_m in the integration scheme; i.e., with each point r , include all points Rr with R symbolizing operations of the point group. Transformation to an orthogonal basis by diagonalizing the overlap matrix S will produce the desired result.

REFERENCES

1. A. A. Frost, J. Chem. Phys. 10, 240 (1942);
A. A. Frost, R. E. Kellogg and E. C. Curtis, Rev. Mod. Phys. 32, 313 (1960).
2. C. M. Carlson, J. Chem. Phys. 47, 862 (1967);
D. K. Harriss and C. M. Carlson, J. Chem. Phys. 51, 5458 (1969).
3. D. H. Bell and L. M. Delves, J. Computational Physics 3, 453 (1969).
4. J. C. Slater, Phys. Rev. 81, 385 (1951); 165, 658 (1968).
5. J. C. Slater, T. M. Wilson and J. H. Wood, Phys. Rev. 179, 28 (1969).

DISCRETE VARIATIONAL METHOD FOR THE ENERGY BAND PROBLEM
WITH LCAO BASIS AND NON-SPHERICAL LOCAL POTENTIAL*

G. S. Painter

Metals and Ceramics Division, Oak Ridge National
Laboratory, Oak Ridge, Tennessee 37830
and

D. E. Ellis

Physics Department, Northwestern University
Evanston, Illinois 60201

While the discrete variational method (DVM)^{1,2} is not restricted with regard to the form of the basis set, and this is one advantage of the technique, the method has so far been applied to the one electron model Hamiltonian using a linear combination of atomic orbitals (LCAO) basis. The DVM seems most promising for treating compounds, where non-spherical potential terms are significant. The LCAO basis is most appropriate for forming crystal wavefunctions for those systems in which the atomic character of the constituent atoms is maintained to a large degree. The results that have been obtained thus far for bcc Li, diamond, single and multi-layer graphite, LiF, MgO, SiC and TiC show that the DVM is in fact a feasible scheme and works quite well with surprisingly few integration points. Moreover, the LCAO basis has been found to converge the occupied and lower conduction band structures with only slight extensions beyond minimal basis sets.

I. BLOCH BASIS SET

Experience has shown that composite basis sets such as APW, OPW, and combined plane-wave tight binding³ sets can describe both atomic core and interatomic regions of the crystal adequately. Furthermore, recent work^{4,5} demonstrates that even the simple LCAO scheme is capable of high accuracy in the eigenvalues with limited basis size, if restrictive approximations are not made. The usual first principles techniques derive their efficiency from the use of muffin tin model potentials, which are suspect for covalent and ionic crystals. If we retain non-spherical terms in the crystal potential, the LCAO approach becomes advantageous.

* Research sponsored by U. S. Atomic Energy Commission under contract with the Union Carbide Corporation.

In the LCAO representation, the basis of the variational problem is a Bloch sum of atomic-like functions

$$\chi_i(\mathbf{k}, \mathbf{r}) = M^{-1/2} \sum_{\mathbf{v}} e^{i\mathbf{k} \cdot \mathbf{R}_v} a_i(\mathbf{r} - \mathbf{R}_v - \mathbf{u}_i) \quad (1)$$

where the \mathbf{u}_i are non-primitive vectors locating atoms in the cell at \mathbf{R}_v . In our calculations, we have chosen exponential Slater-type orbital (STO) functions a_i of the form $x^m y^m z^n r^p e^{-\alpha r}$ located at each atom site. From existing atomic calculations with STO bases^{6,7}, we have readily available optimized approximate Hartree Fock basis sets for use as starting functions for the crystal calculation. Such a Bloch basis is very convenient for eigenstate analysis. The eigenfunction solutions can be studied by performing a population analysis of the atomic functions, to examine ionicity and hybridization properties of the crystal.

Let us examine the difficulties associated with the LCAO method and see what advantages the DVM offers. For a matrix element of the crystal potential for a compound, we obtain

$$V_{ij}(\mathbf{k}) = \langle \chi_i | V | \chi_j \rangle \propto \sum_{\mathbf{v}} e^{i\mathbf{k} \cdot \mathbf{R}_v} \langle a_i(\mathbf{r} - \mathbf{u}_i) | V(\mathbf{r}) | a_j(\mathbf{r} - \mathbf{R}_v - \mathbf{u}_j) \rangle \quad (2)$$

if one takes the standard LCAO approach of expanding X_i in terms of the atomic constituents according to (1). This approach yields a sum of multicenter integrals with the compounded difficulty of finding an accurate local representation of the non-spherical terms in $V(\mathbf{r})$. These integrals must be evaluated by some analytic approach, which can restrict the form of the basis and the representation of $V(\mathbf{r})$, or by some numerical technique. Furthermore, inaccuracies can arise from a loss of significant figures in performing the lattice sum over integrals.

Since one is effectively led to a numerical evaluation of the constituent integrals, an efficient and straightforward approach to the problem is the direct evaluation of the Bloch functions χ_i and formation of matrix elements such as $\langle \chi_i | V | \chi_j \rangle$ by a numerical quadrature in a single unit cell of the crystal, constructing the potential function values $V(\mathbf{r})$ directly over the integration grid. The advantages of such a numerical approach are clear - the important question that must be answered however is:

Is the integration scheme sufficiently convergent for reasonably small numbers of integration points to allow the treatment of large systems in reasonable computation times.

The results which we have obtained demonstrate that this is indeed so. Before discussing some applications however, let us comment briefly on basis set convergence, since most work has been done with rather small basis size. As a rule of thumb, we take available atomic calculations^{6,7} and select the dominant exponential screening constant for each core state. For each of the valence orbitals we pick a pair of screening constants to bracket the dominant atomic orbital parameter, thus allowing for orbital expansion or contraction upon crystal formation. We then proceed to test basis convergence by adding more Bloch functions as judged necessary to stabilize the band structure in the energy region of interest (typically the valence and lower conduction bands). Just as we are free to augment the basis to improve the band structure in a given energy range, it is a simple matter to choose the integration point distribution so as to weight the accuracy of the results in the various bands.

II. RESULTS

A. Lithium

As an initial test for both the method and the programs developed we took the case of bcc lithium, using the muffin potential expansion given by Lafon and Lin⁴. The energy bands were found to converge rapidly with the number of integration points - the uncertainty of less than 0.002 Ry inferred for the 300 point sample is comparable to the magnitude of errors from the truncation of the Bloch lattice summation (1) at a radius of 20 a.u. The splittings of degenerate levels, which occur due to accumulated error and the lack of symmetry in the integration grid gives a comparable error magnitude. Such splittings can be eliminated by symmetrizing the integration grid; however, our experience indicates that the splittings serve as a reliable upper bound to the numerical error of the method.

Convergence of the energy bands with respect to the basis was studied by considering several basis sets, rather than carrying through a systematic variation of exponential parameters. A basis of six Bloch functions gives a rather accurate description of the core level, conduction and lower excited bands over the Brillouin zone with 300 integration points. Extending the basis to ten Bloch functions yielded changes of less than 0.005 Ry for the occupied states, but serve primarily to fill in the excited "sp" bands.

In Table 1 we compare the results of our calculation for lithium with those of Ham⁸ (Green's function) and Lafon and Lin⁴. With ten basis functions and 300 integration points, we have ade-

quately converged the occupied and lower excited band structures (the higher excited state Γ_{15} lies at 0.532 Ry differing from the Green's function result by about 0.026 Ry).

Table 1. Comparison of the conduction band energies of lithium calculated by different methods (lattice constant $a=6.65$ a.u.). Energies in Rydbergs.

$ak/2\pi$	DVM (10 functions, 300 points)	Green's function ^a	LCAO Gaussian transform ^b
$[100] \Delta_1$			
0.0000	-.682	-.681	-.674
0.2500	-.640	-.640	-.629
0.5000	-.512	-.512	-.500
0.6250	-.414	-.414	-.407
0.7500	-.290	-.294	-.290
1.0000	-.062	-.061	-.065
$[110] \Sigma_1$			
0.2500	-.598	-.598	-.587
0.3750	-.498	-.497	-.485
0.5000	-.414	-.411	-.399
$[111] \Lambda_1$			
0.1250	-.650	-.651	-.641
0.2500	-.556	-.556	-.545
0.3125	-.486	-.486	-.478
0.3750	-.400	-.400	-.395
0.5000	-.192	-.191	-.190

^aReference 8

^bReference 4

B. Diamond

One main advantage of the DVM is the facility for using crystal potentials of unrestricted form, e.g. non-spherical potentials. Diamond is a classic example of the covalent crystal in which non-spherical potential terms are large. We have performed a number of calculations on diamond, details of which will be reported elsewhere. As an example of the gross oversimplification of the muffin tin model for such potentials, we constructed a crystal potential from a direct superposition of free atom Coulomb potentials and charge densities, applying the Slater " $\rho^{1/3}$ " exchange approximation, and compared this with the results of a muffin tin average. We formed the muffin tin model by a direct spherical average inside touching carbon atom spheres and fit the potential over the remainder of the cell to a constant. The variation in the non-spherical potential is greater than 1.0 Ry from the bond center in the unit cell to a point 180° away about the atom center. This amounts to a 35% variation about the muffin tin value, and the deviation of the non-spherical potential at the bond center from the muffin tin value exceeds 0.6 Ry. The band structure is sufficiently sensitive to this averaging to cause a 16% reduction in the direct band gap at Γ compared to the gap found with the complete potential. The LCAO basis appears to converge rapidly; we find that supplementing an eighteen function basis set with 3p and 3d orbitals has only a slight effect on occupied and low lying excited levels.

C. Graphite

There have been several calculations for graphite using the muffin tin model, although the non-spherical potential terms are as significant as in diamond. There have been numerous LCAO treatments, generally making the typical tight-binding approximations or some adjustment of integrals to fit the observed data. Unfortunately, properties such as optical structure are not satisfactorily explained. The results of our calculations along "ab initio" lines with the DVM show good agreement with the observed reflectivity structure for graphite⁹ and are reported elsewhere¹⁰. We present here results of an integration point convergence study carried out for single-layer graphite and comment on the question of electron-hole occupancy at the point K (or P) in the Brillouin zone for multilayer graphite.

In Fig. 1 we plot the energies of the valence and lower conduction bands at Γ versus the total number of integration points in the unit cell. The number of atoms per unit cell is two so that the number of points per atom is half the value appearing on the abscissa. It is striking first of all that the gross band

structure is obtained with only 150 points (75 per atom). The valence bands are converged to a few thousandths of a Hartree (~ 0.005 Hartree spread in doubly degenerate state Γ_{3g}^+) and the lower conduction bands (e.g. doubly degenerate level Γ_{3u}^+) to ~ 0.01 Hartree with only 300 integration points per atom. Doubling the total number of points to 1200 reduces the spread at Γ_{3u}^+ to less than 0.002 Hartree while the other levels change by less than 0.001 Hartree. This illustrates that the degeneracy spread is an upper bound to the uncertainty in energy (aside from questions of convergence in basis); in this case, state Γ_{3g}^+ forms the bound. Computing time for this system was about 80 seconds on the IBM 360/91 using an unoptimized program. Recently, we have implemented a newer version which retains k -dependent quantities throughout the calculation; this has cut execution times by about a half.

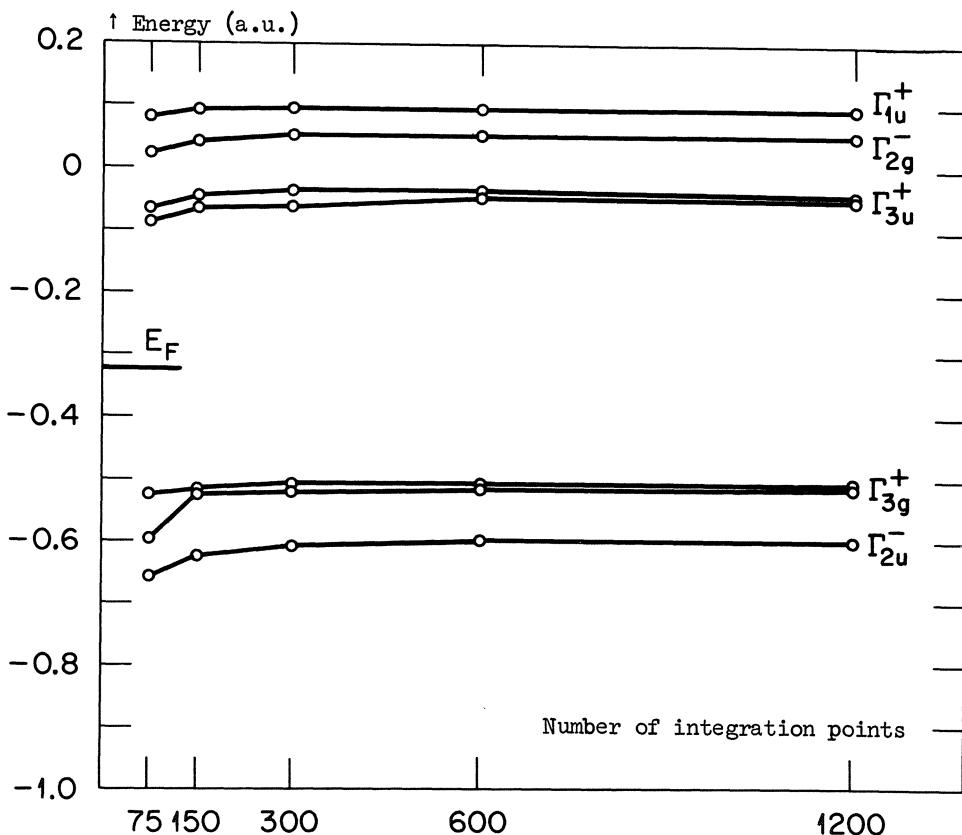


Figure 1. Convergence of energy bands in single layer graphite with number of integration points.

For the multi-layer graphite crystal, we have four atoms per unit cell and the energy bands depend upon the wavevector perpendicular to the layer planes. This dependence determines the complex nature of the Fermi surface and electrical properties of graphite. Our point convergence studies for this case indicate that we have determined the bands at the Fermi surface along the HPH edge of the Brillouin zone to about 0.001 Hartree. The band overlap along this direction is calculated to be about .03 eV, somewhat smaller than that deduced from recent experimental work (.036 to .04 eV), and predicts hole occupancy at P in contrast to recent experimental results^{11,12}. This problem can provide an interesting test of the adequacy of the one-electron model and whether the electron correlation must be included. This question can not be answered at the present stage of our work since the effect involves such slight changes in the bands (of magnitude .001 Hartree) as will certainly occur in proceeding to self-consistency, which could alter the present result.

D. Preliminary Results for Larger Systems

Work is nearing completion on the compound SiC and will be presented elsewhere. Calculations are also in progress on LiF, MgO and TiC, and we wish to indicate here how the DVM is working for these larger systems. In the case of SiC, we find that an integration grid of 600 points per atom converges the energy bands to the same accuracy as obtained for graphite. The computation time runs about one minute on the IBM 360/91 for each k point with about 27 Bloch functions, including 3d orbitals. The first indirect band gap ($\Gamma_{15}(v)$ - $X_1(c)$) obtained from the band structure is about 2.56 eV compared with the accepted experimental value of 2.39 eV. Initial work on TiC with neutral atom starting potential configurations including 3d and 4s atomic orbitals has given results which compare favorably with the APW results of Conklin and Silversmith¹³. A 600 point/atom grid is yielding convergence to better than .005 Hartree for the occupied states, with somewhat less accuracy for the excited states. Preliminary results for the optically interesting insulators LiF and MgO show that ionic potentials and diffuse negative ion wavefunctions present no special difficulties for the use of an LCAO basis. However, there are indications that self-consistency corrections to the crystal potential will cause significant shifts in the band structure.

In conclusion, we have found that the discrete variational method, as specialized to an LCAO basis with non-spherical one-electron potential, is an efficient and easily implemented method for treating the energy band problem. These techniques are not competitive with the APW or KKR methods for systems in which the muffin-tin approximation is adequate, but the DVM appears to be quite useful for cases in which non-spherical terms are significant.

REFERENCES

1. D. E. Ellis and G. S. Painter, preceding article.
2. D. E. Ellis and G. S. Painter (submitted for publication).
3. A. B. Kunz, Phys. Rev. 180, 934 (1969).
4. E. E. Lafon and C. C. Lin, Phys. Rev. 152, 579 (1966).
5. G. S. Painter and D. E. Ellis, Int. J. Quantum Chem. 3S, 801 (1969).
6. E. Clementi, IBM J. Res. Develop. Suppl. 9, 2 (1965).
7. E. Clementi and D. L. Raimondi, J. Chem. Phys. 38, 2686 (1963); 47, 1300 (1967).
8. F. S. Ham, Phys. Rev. 128, 82, 2524 (1962).
9. D. L. Greenaway, G. Harbeke, F. Bassani and E. Tosatti, Phys. Rev. 178, 1340 (1969), and references therein.
10. G. S. Painter and D. E. Ellis, Phys. Rev. (to appear June 1970).
11. P. R. Schroeder, M. S. Dresselhaus, and A. Javan, Phys. Rev. Letters 20, 1292 (1968).
12. The authors are grateful to J. W. McClure for a preprint of his review article, to be published in J. Phys. Chem. Solids.
13. J. B. Conklin, Jr. and D. J. Silversmith, Int. J. Quantum Chem. 2S, 243 (1968).

RECENT DEVELOPMENTS IN APPLYING AND EXTENDING THE METHOD OF TIGHT
BINDING (LCAO) TO ENERGY-BAND CALCULATIONS

Earl E. Lafon

Department of Physics, Oklahoma State University
Stillwater, Oklahoma

Roy C. Chaney and Chun C. Lin

Department of Physics, University of Wisconsin
Madison, Wisconsin

The purpose of this paper is to investigate the computational aspects of the method of tight binding and is roughly divided into two parts. In the first part we present the formulation and application of the tight-binding problem in its most rigorous and straight-forward form without introducing any numerical approximations and then examine the effects of the "nearest-neighbor approximation" which was frequently employed within this scheme to reduce the computational complexity. The second part is devoted to investigating various methods for increasing the variational freedom over that of the traditional minimal basis set and to examining the versatility and applicability of this method to a broad spectrum of crystalline materials.

I. FORMULATION

The method of tight binding as presented here is developed entirely within the scheme of linear variation of parameters using the Bloch sums $b_\alpha^i(\vec{k}, \vec{r})$ as basis. For this discussion, we write

$$b_\alpha^i(\vec{k}, \vec{r}) = (N)^{-\frac{1}{2}} \sum_v e^{i\vec{k} \cdot \vec{R}_v} \phi_\alpha[\vec{r} - (\vec{R}_v + \vec{t}_i)] \quad (1)$$

where N represents the number of unit cells in the crystal, \vec{k} defines the region of the Brillouin zone under consideration, and where vectors \vec{R}_v and \vec{t}_i label the positions of the unit cells and

of the various atoms contained within each unit cell respectively. The expression $\phi_\alpha[\vec{r} - (\vec{R}_v + \vec{t}_j)]$ represents some "atomic-like" function $\phi_\alpha(\vec{r})$ centered about the i -th atom of the v -th unit cell. Traditionally, these $\phi_\alpha(\vec{r})$ are chosen to be the Hartree-Fock orbitals of the free atoms (or some approximate versions of them) of which the crystal is composed. When this is done, and when care is taken to include those orbitals of the valence shell necessary to assure that the resultant trial wavefunction will possess the proper transformation properties under the operations of the group of the wave vector, we then call this a minimal basis set. Although historically the method of tight binding has been invariably employed in conjunction with such an atomic orbital basis, this restriction is not an integral part of the tight-binding scheme and in fact it will be shown that the variational freedom can be substantially improved by either replacing or augmenting the minimal basis set with Bloch sums of Gaussian-type (GTO) or Slater-type (STO) orbitals. We will begin our investigations using the minimal set and consider various extensions and augmentations in Sec. V.

With the trial wavefunction in the form of

$$\psi_n(\vec{k}, \vec{r}) = \sum_{\alpha} \sum_i a_{n,\alpha}^i(\vec{k}) b_{\alpha}^i(\vec{k}, \vec{r}), \quad (2)$$

the variation of the linear parameters $a_{n\alpha}^i(\vec{k})$ yields an energy minimum when

$$\det |H(\vec{k}) - ES(\vec{k})| = 0,$$

where the matrix elements of H and S are

$$\begin{aligned} H_{\alpha\beta}^{ij}(\vec{k}) &= \int b_{\alpha}^i(\vec{k}, \vec{r}) H b_{\beta}^j(\vec{k}, \vec{r}) d\tau \\ &= \sum_v e^{i\vec{k} \cdot \vec{R}_v} \int \phi_{\alpha}^*(\vec{r} - \vec{t}_i) H \phi_{\beta}[\vec{r} - (\vec{R}_v + \vec{t}_j)] d\tau, \end{aligned} \quad (3)$$

and

$$\begin{aligned} S_{\alpha\beta}^{ij}(\vec{k}) &= \int b_{\alpha}^i(\vec{k}, \vec{r}) b_{\beta}^j(\vec{k}, \vec{r}) d\tau \\ &= \sum_v e^{i\vec{k} \cdot \vec{R}_v} \int \phi_{\alpha}^*(\vec{r} - \vec{t}_i) \phi_{\beta}[\vec{r} - (\vec{R}_v + \vec{t}_j)] d\tau, \end{aligned}$$

respectively, and where H is the one electron Hamiltonian (in atomic units)

$$H = -\frac{1}{2}\nabla^2 + V_{\text{cry}}(\vec{r}). \quad (4)$$

Thus the problem involves the calculation of a number of multi-center integrals of the form

$$\int \phi_{\alpha}^*(\vec{r}_A) \phi_{\beta}(\vec{r}_B) d\tau, \quad (5a)$$

$$\int \phi_{\alpha}^*(\vec{r}_A) (-\frac{1}{2}\nabla^2) \phi_{\beta}(\vec{r}_B) d\tau, \quad (5b)$$

$$\int \phi_{\alpha}^*(\vec{r}_A) V_{\text{cry}}(\vec{r}) \phi_{\beta}(\vec{r}_B) d\tau, \quad (5c)$$

where \vec{r}_A and \vec{r}_B are as described in Fig. 1. Since these integrals are independent of \vec{k} , they must be calculated only once for each crystalline material.

II. EVALUATION OF MULTICENTER INTEGRALS

Integrals of the form of Eqs. (5a) or (5b) are a common occurrence in molecular work and present no problem. The major difficulty with the method of tight binding has always been the accurate evaluation of the potential energy integrals of Eq. (5c). The periodic $1/|\vec{r} - \vec{R}_V|$ singularities present a severe problem for numerical integration and an analytic or semi-analytic expression which removes this singularity problem is highly desirable. This can be accomplished by expressing the "atomic-like" functions $\phi_{\alpha}(\vec{r})$ as some linear combinations of STO

$$\phi_{\alpha}(\vec{r}) = \sum_i c_{\alpha,i} \chi_i^s(\vec{r}),$$

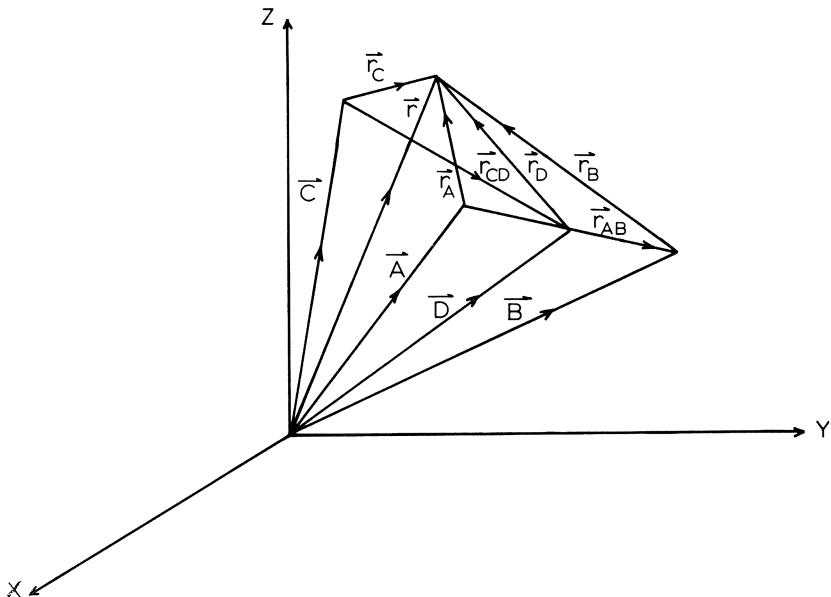


Figure 1. Relations between the various vectors occurring in the reduction of the Gaussian integrals.

or GTO

$$\phi_{\alpha}(\vec{r}) = \sum_i c_{\alpha,i} \chi_i^G(\vec{r}),$$

and by expanding the periodic crystalline potential $V_{\text{cry}}(\vec{r})$ in a Fourier series

$$V_{\text{cry}}(\vec{r}) = \sum_{\nu} V_1(\vec{k}_{\nu}) \cos \vec{k}_{\nu} \cdot \vec{r}_c + \sum_{\nu} V_2(\vec{k}_{\nu}) \sin \vec{k}_{\nu} \cdot \vec{r}_c, \quad (6)$$

where \vec{c} denotes the origin about which the potential is expanded and is shown in Fig. 1. The required integrals can thus be decomposed into a sum of integrals of the form

$$\begin{aligned} \int \chi^G(\vec{r}_A) V_{\text{cry}}(\vec{r}) \chi^G(\vec{r}_B) d\tau &= \sum_{\nu} V_1(\vec{k}_{\nu}) \int \chi^G(\vec{r}_A) \cos(\vec{k}_{\nu} \cdot \vec{r}_c) \chi^G(\vec{r}_B) d\tau \\ &+ \sum_{\nu} V_2(\vec{k}_{\nu}) \int \chi^G(\vec{r}_A) \sin(\vec{k}_{\nu} \cdot \vec{r}_c) \chi^G(\vec{r}_B) d\tau, \end{aligned} \quad (7)$$

for expansion in GTO and an analogous expression for expansion in STO. The resultant integrals of sines and cosines can be evaluated analytically for GTO and result in a semi-analytic form involving a one-dimensional integral which must be evaluated numerically for STO. Our investigations indicate the Gaussian expansion as generally preferable to the exponential, particularly for materials past the second row of the periodic table. For this reason only the development for Gaussians will be presented here, while a similar development for STO can be found in the literature.¹ For our problem, the basic integral is

$$C_{1s,1s}^G = \int \exp(-\alpha_1 r_A^2) \cos(\vec{k}_{\nu} \cdot \vec{r}_c) \exp(-\alpha_2 r_B^2) d\tau. \quad (8)$$

If this integral can be expressed in analytic form, then all other integrals involving higher orbitals can also be evaluated in closed form. Using the identity

$$\exp(-\alpha_1 r_A^2) \exp(-\alpha_2 r_B^2) = \exp[-\alpha_1 \alpha_2 r_{AB}^2 / (\alpha_1 + \alpha_2)] \exp[-(\alpha_1 + \alpha_2) r_D^2], \quad (9)$$

where the coordinates of the point D are related to those of A and B as

$$D_i = (\alpha_1 A_i + \alpha_2 B_i) / (\alpha_1 + \alpha_2), \quad i = x, y, z,$$

one may reduce Eq. (8) to the single-center integral

$$C_{1s,1s}^G = \exp[-\alpha_1 \alpha_2 r_{AB}^2 / (\alpha_1 + \alpha_2)] \int \exp[-(\alpha_1 + \alpha_2) r_D^2] \cos(\vec{k}_{\nu} \cdot \vec{r}_c) d\tau,$$

which is easily evaluated using the substitution

$$\vec{r}_c = \vec{r}_D + \vec{r}_{cD}, \quad (10)$$

yielding the simple analytic expression

$$C_{1s,1s} = [\pi/(\alpha_1 + \alpha_2)]^{3/2} \exp[-\alpha_1 \alpha_2 r_{AB}^2/(\alpha_1 + \alpha_2) - K_v^2/4(\alpha_1 + \alpha_2)] \\ \times \cos(\vec{k}_v \cdot \vec{r}_{CD}) \quad (11)$$

Integrals involving higher orbitals can be obtained from the above expression by simple differentiation. For example,

$$C_{1s,3s} = \int \exp(\alpha_1 r_A^2) \cos(\vec{k}_v \cdot \vec{r}_c) r_B^2 \exp(-\alpha_2 r_B^2) d\tau \\ = -(\partial/\partial \alpha_2) C_{1s,1s} \\ = [\pi/(\alpha_1 + \alpha_2)]^{3/2} \exp[-\alpha_1 \alpha_2 r_{AB}^2/(\alpha_1 + \alpha_2) - K_v^2/4(\alpha_1 + \alpha_2)] \\ \times \{ \frac{1}{4} [6(\alpha_1 + \alpha_2) + 4\alpha_1^2 r_{AB}^2 - K_v^2] (\alpha_1 + \alpha_2)^{-2} \cos(\vec{k}_v \cdot \vec{r}_{CD}) \\ + (\vec{k}_v \cdot \vec{r}_{AB}) \alpha_1 (\alpha_1 + \alpha_2)^{-2} \sin(\vec{k}_v \cdot \vec{r}_{CD}) \}. \quad (12)$$

In this manner, all necessary integrals can be obtained in analytic form. The absolute convergence of the sum over the reciprocal lattice in Eq. (7) is mainly dictated by

$$V_i(\vec{k}_v) \exp[-K_v^2/4(\alpha_1 + \alpha_2)], \quad i = 1, 2,$$

and therefore are very sensitive to the orbital exponents α_1 and α_2 . For most cases the convergence is rapid, often converging after only some twenty or thirty stars of the reciprocal lattice have been included in the summation. On the other hand, some integrals are always encountered for which $\alpha_1 + \alpha_2$ is "large". For these cases, convergence is usually slow and numerical techniques must be used to enhance the convergence. For this purpose, we use a method similar in motivation to the Ewald sum in that the poorly convergent $1/|\vec{r} - \vec{R}_v|$ singularities of $V_{cry}(\vec{r})$ are removed from the Fourier series and re-expressed in direct space as a sum over the crystal lattice. A detailed account of this convergence acceleration is described in the literature³ and will not be repeated here.

III. APPLICATION TO LITHIUM

As an example of the application of the method of tight binding, let us first examine the extreme case of lithium where the electrons participating in the conduction states cannot in any sense be characterized as being "tightly bound". For this example we employ the minimal basis set which, for lithium, is composed of the Hartree-Fock orbitals 1s, 2s, 2p_x, 2p_y, and 2p_z. Analytic GTO expressions for the 1s and 2s orbitals are obtained directly from Huzinaga's tables⁴ while an analytic GTO expansion of the 2p was obtained by a non-linear least squares fit to the tabular

functions of Fock and Petrashen.⁵ Without symmetrization, this leads to a simple 5x5 secular equation which, at a general point in the Brillouin zone, results in a trial wavefunction with only four variational parameters, while at points and lines of high symmetry, the variational freedom is even further reduced. Thus at the Γ -point of the conduction band there is only one degree of freedom, corresponding to the mixing of 1s and 2s, while at the H-point the conduction bond is pure 2p and has no variational freedom. In this investigation, all multicenter integrals are obtained without approximation by the method described, and the summation in Eq. (7) is carried to at least seven figures of convergence. In this calculation, we have employed a muffin-tin version of the Seitz potential⁶ for a body-centered cubic lattice of lattice constant $a = 6.65$ a.u.. It should be mentioned that the method described here for evaluating the potential energy integrals is not restricted to a muffin-tin form since Eq. (6) is quite general. This muffin-tin form was employed in the calculations of Ham⁷ by the Green's-function method and in the calculations of Rudge⁸ by the method of augmented plane waves (APW). In view of the high accuracy of the APW method, we shall use its results as a reference for comparison. The results of the tight-binding calculation using the minimal basis set and the comparison with APW and with the Green's-function method are presented in Table I. Considering the restricted variational freedom of the minimal basis set, we see that the agreement with the other two methods is very good with an average error of less than 0.01 rydberg. The errors at the H, N, and P points are seen to be respectively 0.003, 0.011, and 0.004 rydbergs which is of particular interest since the minimal basis provides no variational freedom at these points. This disagreement can be ascribed to lack of variational freedom in the trial tight-binding wavefunction and methods of improving this comparison by increasing this variational freedom is discussed in Sec. V.

IV. NEAREST-NEIGHBOR APPROXIMATION

Until a few years ago, the large majority of all the energy-band calculations by the method of tight binding employed the "nearest-neighbor approximation", i.e., neglecting all the multi-center integrals in Eqs. (3) except when the two atomic wavefunctions are situated at two nearest-neighbor sites of the crystal. In fact this approximation was used so frequently in tight-binding work that it is sometimes regarded as an integral part of the method itself. To analyze the validity of the nearest-neighbor approximation, we present in Fig. 2 the convergence of a typical matrix element as a function of truncation at single center, nearest neighbor, next nearest neighbor, etc. As can be seen, convergence is slow and not achieved until some 20 sets of equivalent neighbors have been included in the summation. A nearest-

Table I. Energy band structures of lithium (in rydbergs) calculated by the method of tight binding using the minimal basis set and using an augmented set and comparison with the results of the Green's-function method and the APW method. $a_0 = 6.65$ a.u.

$a_0 k_x / 2\pi$	Energies		Green's function	APW
	Tight-Binding minimal	augmented		
[100]				
0.000	-0.675	-0.678	-0.681	-0.679
0.250	-0.631		-0.640	-0.637
0.500	-0.502		-0.512	-0.510
0.625	-0.407		-0.414	-0.414
0.750	-0.291		-0.294	-0.295
1.000	-0.059	-0.062	-0.061	-0.062
[110]				
0.250	-0.588		-0.598	-0.596
0.375	-0.487		-0.497	-0.496
0.500	-0.400	-0.411	-0.411	-0.411
[111]				
0.125	-0.642		-0.651	-0.648
0.250	-0.547		-0.556	-0.555
0.375	-0.393		-0.400	-0.400
0.500	-0.189	-0.191	-0.191	-0.193

neighbor or even next-nearest-neighbor approximation for lithium using atomic orbitals could only be expected to give qualitative results at best. However, it has been shown that a more localized choice than that of Hartree-Fock orbitals can be made for the $\phi_\alpha(r)$, thus improving the convergence of the crystal-lattice summation of Eqs. (3). However, the fact remains that if the Hartree-Fock orbitals of the free atoms were used for $\phi_\alpha(r)$, a large number of neighbors generally need be included to achieve convergence of the lattice summation (except for special cases such as core states and the d-bands of certain transition elements), and any artificial truncation at nearest neighbors will destroy the quantitative aspects of this method.

V. VARIATIONAL FREEDOM

Let us now investigate the increase in accuracy which results from an increase in variational freedom over that of the minimal basis set. An obvious suggestion would be to augment the basis

set by including Bloch sums of the higher atomic orbitals. Direct application of this procedure has been found somewhat awkward and impractical due to the long-range nature of the higher orbitals. Fig. 3 shows the crystal wavefunction of lithium at the Γ point of the conduction band calculated by the minimal basis set. Included for comparison are the Bloch sums of $\vec{k} = 0$ formed by the 2s and 3s Hartree-Fock functions of the lithium atom.⁹ The constancy of the crystal wavefunction over the major part of the graph indicates that the method of tight binding, when carried to proper convergence, automatically yields some characteristics of free particles. Because of the strong resemblance between the 2s and 3s Bloch sums, addition of the latter to the minimal basis set essentially duplicates the 2s basis function and is not effective for increasing the variational freedom.

A more practical approach is to take each of the individual Gaussian-type orbitals $\exp(-\alpha r^2)$ which were employed to form the Hartree-Fock wavefunction of the free atom, i.e.,

$$\phi_{nl}(\vec{r}) = \sum_j \beta_j \exp(-\alpha_j r^2), \quad (13)$$

to form the single-Gaussian Bloch sums and use them as the basis functions. In other words, their mixing is to be determined by the variational principle rather than by the β 's obtained from⁴ the free-atom problem. For the 2s atomic function of lithium,

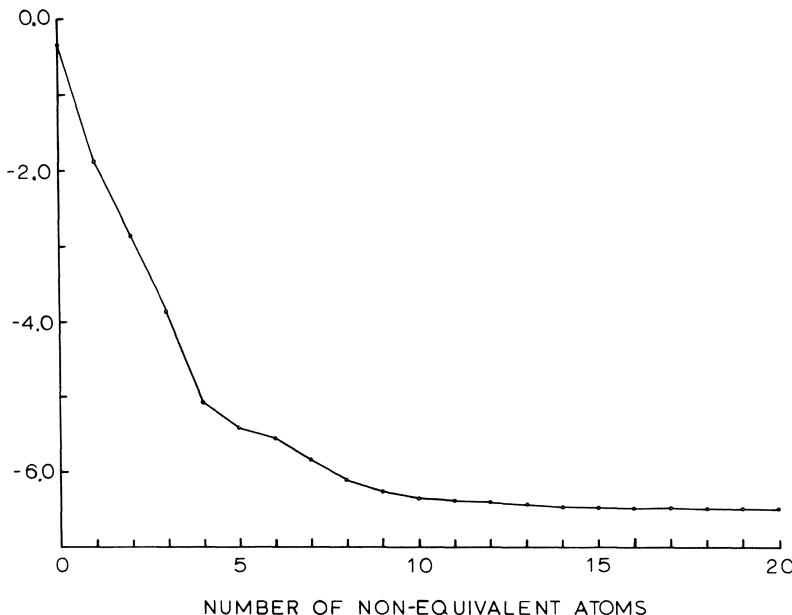


Figure 2. Convergence of the value of the matrix element $(b_{2s} | H | b_{2s})$ with respect to crystal-lattice summation.

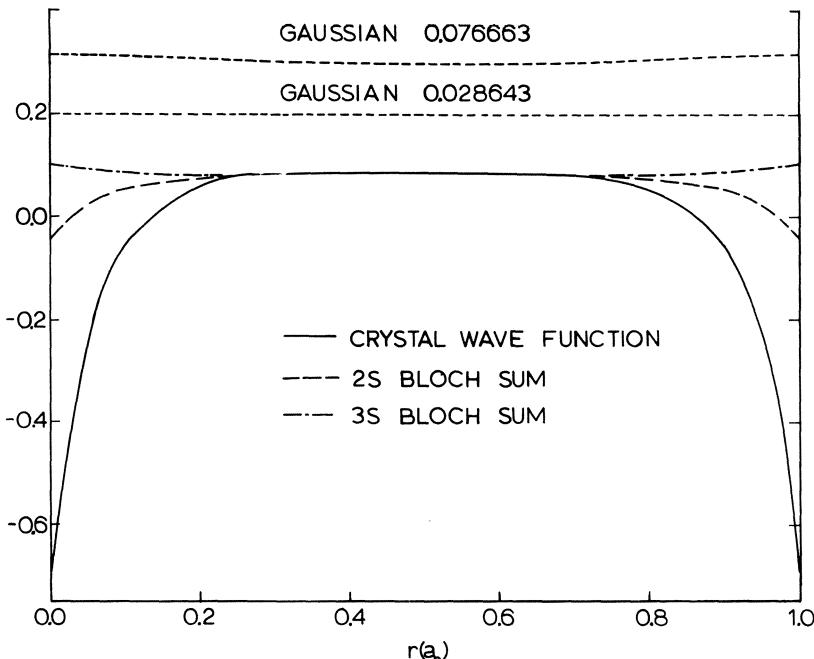


Figure 3. Crystal wavefunctions and Bloch sums of the Γ point along the [100] line of the lithium crystal. The two upper dashed lines are the single-Gaussian Bloch sums with $\alpha = 0.028643$ and $\alpha = 0.076663$ as labeled. In the lowest group the solid curve is the crystal wavefunction calculated by using the minimal basis set, and the two broken curves refer to the Bloch sums constructed by the 2s and by the 3s Hartree-Fock atomic orbitals. The two single-Gaussian Bloch sums are not normalized.

the values of α_i are 0.028643, 0.076663, 0.44462, 1.15685, 3.15789, 9.35329, 31.9415, 138.730, and 921.271. The two single-Gaussian Bloch sums formed by the two GTO of longest range ($\alpha_j = 0.028643$ and 0.076663) are displayed in Fig. 3. They differ only in the region very close to the atomic sites which is the domain of the short-range gaussians. This suggests that the GTO with $\alpha = 0.028643$ offers little functional variation over that already contributed by the other eight Bloch sums and hence can be omitted. This step is of special importance from the computational standpoint, for the long-range Gaussians are responsible for the slow convergence of Eqs. (3). The size of the secular equation is now dictated by the number of different Gaussian exponents chosen. In case it is desirable to reduce the size of the secular equation, the technique of contracted Gaussian may be recommended.³ Here

the GTO's are divided into several groups. Within each group, the relative weighting of the members is taken to be the appropriate β 's in Eq. (13), whereas the mixing between different groups is determined by the variational principle. This contraction procedure is not essential for the method of tight binding, and the readers are referred to Ref. 3 for detailed descriptions. By using the single-Gaussian Bloch sums as the basis functions and adopting the contraction procedure, we obtained 0.678 rydbergs for the energy of the Γ point of lithium.

We now turn our attention to the points H, N, and P for which the minimal basis produces no variational freedom. The use of single-Gaussian basis sets should be particularly helpful for improving the accuracy. The 2p atomic Hartree-Fock orbital can be fitted to Eq. (13) with $\alpha_j = 0.0229434, 0.0764918, 0.444620, 1.15685, 3.15789, 9.35329$, and 31.9415. The Gaussian component having the smallest value of α_j can be rejected for the reason explained in the preceding paragraph. The remaining GTO's generate six Bloch-sum basis functions which are then varied independently (no contraction procedure) to give the energies shown in Table I. This method of increasing the variational freedom is seen to be quite effective, having reduced the maximum discrepancy with APW from 0.011 rydbergs to 0.002 rydbergs by introducing only six degrees of variational freedom.

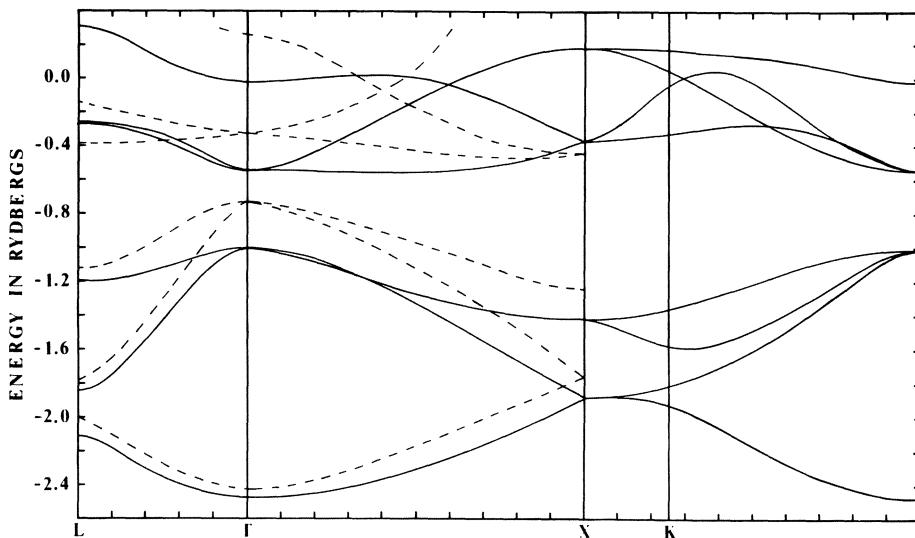


Figure 4. Energy band structure of diamond. The solid curves are the results of the tight-binding calculations (minimal basis set) and the dashed ones are those of OPW (Ref. 10).

VI. APPLICATION TO DIAMOND

To further demonstrate the general applicability of this method, we now examine the case of diamond. The minimal basis set for diamond is composed of the Hartree-Fock orbitals $1s$, $2s$, $2p_x$, $2p_y$, $2p_z$, for each of the two atoms in the unit cell and these orbitals are expressed analytically in terms of Gaussians using Huzinaga's tables.⁴ This results in a 10×10 secular equation. A lattice constant of 6.728 a.u. and a crystal potential formed from overlapping Hartree-Fock-Slater free-atom potentials are used in this calculation so that a direct comparison can be made between the method of tight binding and the method of orthogonalized plane waves (OPW) using some ninety plane waves.¹⁰ The results of this minimal basis set calculation (solid line) and the corresponding OPW calculation (dashed line) are shown in Fig. 4. It will be noticed that at almost every point in the Brillouin zone, the tight-binding results lie considerably below those of OPW. Since the variational procedure is used, and since the two methods are being applied to the same eigenvalue problem, we conclude that the method of tight binding using even a minimal basis set is considerably more accurate for diamond than some ninety OPW's. Again, we can improve the accuracy of our tight-binding procedure by increasing the variational freedom. For diamond, this is accomplished by retaining the original Hartree-Fock atomic orbitals $1s$, $2s$, $2p$, and by augmenting these by Bloch sums of two Gaussian $1s$ functions with orbital exponents 5.14773 and 0.49624 and by one Gaussian $2p$ functions with orbital exponent 0.35945. With this set of twenty basis functions, the energies of the various points of high symmetry are as follows: Γ_1 , -2.528; Γ_{25} , -1.013; Γ_{15} , -0.550; Γ_2 , -0.053; X_1 , -1.908; X_4 , -1.442; X_1 , -0.389; X_3 , 0.177; L_2 , -2.133; L_1 , -1.913; L_4 , -1.210; L_1 , -0.284; L_3 , -0.288; L_2 , 0.222 rydbergs. Even this small increase in variational freedom is seen to produce a significant increase in accuracy.

VII. APPLICATION TO LITHIUM FLUORIDE

As an example to illustrate the application of the method of tight binding to the binary-compound crystals, let us consider the energy-band calculation of lithium fluoride. The crystal potential is constructed by superposition of the potentials of the individual neutral lithium and fluorine atoms at the appropriate sites of the lattice. One may question the accuracy of a potential for an ionic crystal generated from the charge distribution corresponding to the neutral constituent atoms. However, in this paper we are mainly interested in the computational aspects of the problem and therefore will not be concerned with the criteria of selecting the most suitable crystal potential at this point. We adopt the minimal set which consists of the ten Bloch sums of the $1s$, $2s$,

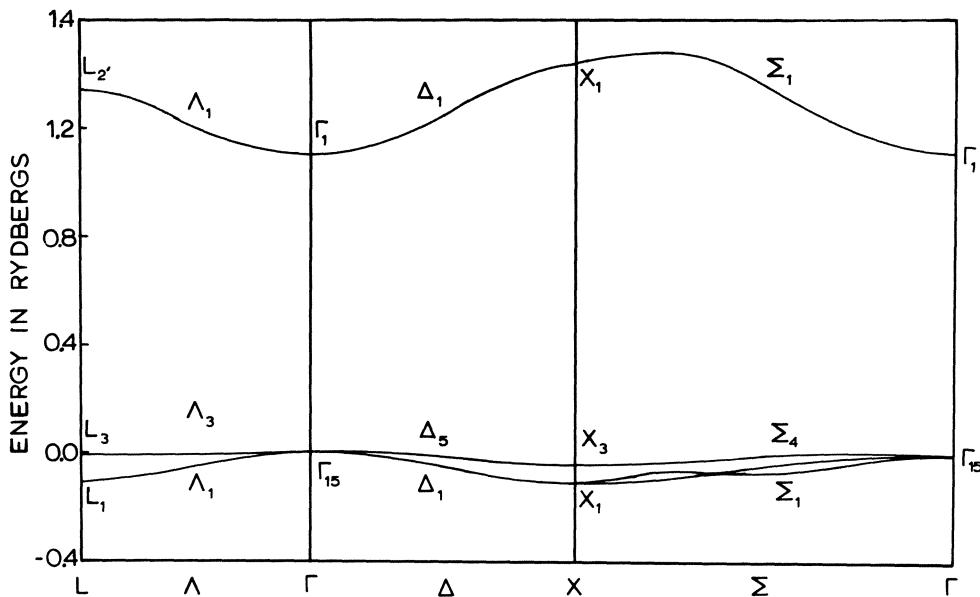


Figure 5. Energy band structure of lithium fluoride crystals.

$2p_x$, $2p_y$, $2p_z$ Hartree-Fock orbitals of the lithium atoms and of the fluorine atoms with a lattice constant of 7.594 a.u. Fig. 5 shows the valence and conduction band resulting from this minimal set. The calculated band gap is 15.0 eV which compares favorably to the experimental value of 13.6 eV.¹¹ Since only the 1s, 2s, and 2p states of the constituent atoms were included in the basis set, the excited states corresponding to the d-orbitals do not appear in Fig. 5.

REFERENCES

1. E. E. Lafon and C. C. Lin, Phys. Rev. 152, 579 (1966).
2. I. Shavitt, in Methods in Computational Physics, edited by B. Alder, S. Fernbach, and M. Rotenberg (Academic Press, Inc., New York, 1963), Vol. 2, p. 1.
3. R. C. Chaney, T. K. Tung, C. C. Lin, and E. E. Lafon, J. Chem. Phys. 52, 361 (1970).
4. S. Huzinaga, J. Chem. Phys. 42, 1293 (1965).
5. V. Fock and M. J. Petrashen, Physik. Z. Sowjetunion 8, 547 (1935).
6. W. Kohn and J. Rostocker, Phys. Rev. 94, 1111 (1954).
7. F. S. Ham, Phys. Rev. 128, 82 (1962).
8. W. E. Rudge, private communication.
9. A. W. Weiss, Astrophys. J. 138, 1262 (1963).
10. F. Bassani and M. Yoshimine, Phys. Rev. 130, 20 (1963).
11. D. M. Roessler and W. C. Walker, J. Phys. Chem. Solids 28, 1507 (1967).

ENERGY BANDS BY THE LCAO CELLULAR METHOD

P. D. DeCicco

Department of Physics and

Center for Materials Science and Engineering

M. I. T., Cambridge, Massachusetts

Conceptually, the most simple method for solving the energy band problem is the cellular method in which one solves the Schrödinger equation within the unit cell and imposes boundary conditions on the solution at the surface of the unit cell thereby introducing the wave vector dependence of the energy. Computationally, the implementation of this method and the subsequent use of the resulting wave functions suffer from complications associated with the shape of the unit cell. In terms of compactness and the functional form of the wave functions, the LCAO method as employed by Lafon and Lin is the most simple. Although the LCAO method is only an approximate method, these authors have shown that the conduction band of lithium can be reproduced using only the Bloch sums of atomic 1s, 2s and 2p functions to an accuracy of about 0.01 rydbergs¹. Furthermore, the atomic functions used in this calculation were each linear combinations of two Slater orbitals with coefficients and orbital exponents chosen so as to approximate the numerical Hartree-Fock orbitals. It is reasonable to expect that considerably greater accuracy could be obtained without increasing the number of Slater orbitals used to represent the Bloch function at a given point in the Brillouin zone by varying the linear coefficients and orbital exponents at each point to minimize the one-electron energy. The computational difficulty with the LCAO method arises in the calculation of the matrix elements of the one-electron Hamiltonian between orbitals centered on different atomic sites. This difficulty is so severe that it is found profitable to replace the Slater basis by the less compact Gaussian basis^{2, 3}. The method proposed here makes use of the assumption, supported by the success of Lafon et. al., that a few Bloch sums of Slater orbitals can provide an accurate representation of the wave function in the energy band problem. The remaining ingredient in this method is the comput-

ationally simple part of the cellular method, namely the solution of the Schrodinger equation for a trial value of the energy.

LCAO-Cellular Method

Let us consider a periodic crystal with one or more atoms per unit cell with position vectors \vec{q}_j relative to the lattice point associated with a given cell. Let us assume that the wave function of a Bloch state can be represented with sufficient accuracy by Eq. (1)

$$\Psi_{\vec{k}}(\vec{r}) \cong \sum_{\vec{R}} e^{i\vec{k} \cdot \vec{R}} \varphi(\vec{r} - \vec{R}) \quad (1)$$

where

$$\varphi(\vec{r}) = \sum_{j=1}^M A_j |\vec{r} - \vec{q}_j|^{-n_j} e^{-\alpha_j |\vec{r} - \vec{q}_j|} Y_{\ell_j m_j} \left(\frac{\vec{r} - \vec{q}_j}{|\vec{r} - \vec{q}_j|} \right) \quad (2)$$

Thus to find the (unnormalized) wave function we are left with the problem of determining $2M+1$ parameters. Let us suppose that we know the solution of the Schrodinger equation for a trial value of energy E , within some spherical regions surrounding the atoms in the unit cell. In the case of a crystal potential which is spherical within spheres surrounding the atoms, the solution would be obtained from the radial Schrodinger equation while in the case of an arbitrary form of potential within the spheres the solution might be obtained by solving coupled equations for the radial functions in a partial wave expansion. If we now require that the approximate solution of Eq. (1) agree with the exact solution in $2M+1$ respects, we can then determine the $2M$ parameters in Eq. (1) as well as $E(k)$. In principle, of course, the knowledge of the exact solution over a finite region of space gives us infinitely many conditions to impose on our approximate wave function. Since our approximate wave function contains only a finite number of variable parameters, we can only count upon having a reasonable representation of the important features of the exact solution and therefore we should select a set of $2M+1$ important conditions to impose. Let us now proceed to consider various possible choices.

Augmented LCAO Method

One way of choosing the conditions which we impose upon the LCAO approximate wave function is quite analogous to the augmented plane wave (APW) method. In the APW method the wave function in the interstitial region is assumed to be a linear combination of plane waves with the appropriate translational symmetry. The coefficients multiplying these plane waves are determined in principle by the boundary conditions on the wave function and its grad-

ient at the surfaces of the APW spheres. In practice, however, only a finite number of terms are included in the plane wave expansion so that the boundary conditions cannot be satisfied exactly and the variational principle giving rise to a secular equation is used to determine the coefficients. If instead we assume the wave function in the interstitial region to be a linear combination of a few Bloch summed atomic orbitals, we can then employ the Löwdin alpha-function expansion⁴ to determine the logarithmic derivatives of the important partial waves of the interstitial function at the surfaces of the APW spheres. The logarithmic derivatives of the radial functions in the interior solution can be calculated as in the APW method and the condition that the logarithmic derivatives match will determine the energy.

As a particular example, let us consider the conduction band of Li. At $k = 0$, the s-wave is the only important partial wave and the energy is determined by the condition that the s-wave part of the solution inside the sphere have the same logarithmic derivative as the Bloch summed atomic s-function. It should be noted here that the number of conditions imposed on the crystal wave function in this example is greater than one since the known atomic orbital was used. Along the 100 direction we expect the p function to mix in with the s function in our Bloch sum. Thus we will have the mixing coefficient and the energy as unknowns to be determined by conditions on the s-wave and p-wave logarithmic derivatives.

LCAO-Taylor Series Method

About any atomic center, the exact solution to the energy band problem can be expanded in the form

$$\Psi_{\vec{k}}(\vec{r}) = \sum_{\ell m} u_{\ell m}(r) Y_{\ell m}(\theta \varphi) \quad (3)$$

and the radial functions $u_{\ell m}(r)$ have a Taylor series expansion

$$u_{\ell m}(r) = r^{\ell} [1 + a_1 r + a_2 r^2 + a_3 r^3 + \dots] \quad (4)$$

where

$$\begin{aligned} a_1 &= -Z/\ell + 1 \\ a_2 &= -(2Za_1 + E - V_0)/(4\ell + 6) \\ a_3 &= -[2Za_2 + (E - V_0)a_1 - V_1]/(6\ell + 12) \end{aligned}$$

and

$$V(\vec{r}) = \frac{-2Z}{r} + V_0 + V_1 r + \dots \quad (5)$$

It is also a simple matter to expand Bloch sums of Slater orbitals in Taylor series about the atomic center and thus obtain conditions on the approximate LCAO expression for the wave function by comparing coefficients in the two Taylor series expansions.

Before looking at a specific example of the LCAC-Taylor series method it is appropriate to go into some general considerations. First, the a_1 coefficient is simply a measure of the nuclear charge and is therefore common to all self-consistent fields. Secondly, the V_1 term when inserted into Poisson's equation implies an infinite exchange charge density at the nucleus so that in fact $V_1=0$ in the Hartree and Hartree-Fock approximations. The statistical exchange potentials proportional to the $1/3$ power of the charge density do have a non-zero value of V_1 arising from the cusp in the charge density at the atomic center. This feature of the statistical exchange would seem to be spurious. A third point of interest is that if we look only at the Taylor coefficients displayed above, it would appear that the difference between a local self-consistent field and the non-local Hartree-Fock potential lies in the quantity V_0 which in the HF approximation is merely a \vec{k} -dependent shift to be applied to the one-electron energy.

In a case where the Taylor coefficients considered above provide enough conditions to determine the wave function and the energy, we would seem to have a contradiction of the well known fact that energy bands do depend upon the self-consistent field. However, if we impose upon a basis containing so few variable parameters the condition that it be able to reproduce the appropriate atomic orbitals, we will have then built into our calculation a good deal of information about the atomic self-consistent field. A very successful example of a calculation in which the self-consistent field for the atom is brought into the crystal by means of the functional form of an atomic orbital is the calculation of the energy bands of copper using the Chodorow potential.⁵ To obtain a generalized Chodorow potential we can take a Bloch sum of a linear combination of Slater d orbitals and define the crystal potential by requiring that this Bloch sum be a solution of Schrodinger's equation. The set of Slater orbitals used should be adequate for the representation of the atomic 3d orbital and the linear combination should contain two variable parameters so that the first few Taylor coefficients of the Bloch function could be made to agree with those determined from Schrodinger's equation as indicated above.

LCAO-Cusp Method

The origin of the rather general class of methods presented above was an attempt to construct a very simple representation of the wave functions corresponding to the valence band of LiH. LiH is an insulator with an NaCl crystal structure and has two singly degenerate occupied bands; the Li 1s band and the valence band

which may be roughly designated the hydride 1s band. Apart from the free-electron model which would seem inappropriate for an insulator, the most simple form of valence Bloch function would seem to be a Bloch sum of Slater 1s orbitals centered on the H sites. As a check on the accuracy of this wave function we might consider the cusp condition (the a_1 coefficient of the Taylor series) at the hydrogen site.

$$\begin{aligned} \Psi(\vec{r}) = & \sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} e^{-\alpha|\vec{R}-\vec{r}|} = 1 - \alpha r + \frac{\alpha^2 r^2}{2} + \dots \\ & + \sum'_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} e^{-\alpha|\vec{R}|} [1 + \sum_{i=1}^3 \frac{r_i R_i \alpha}{R} + \frac{1}{2} \sum_{i=1}^3 \sum_{j=1}^3 r_i r_j \\ & (\frac{-\alpha \delta_{ij}}{R} + \frac{R_i R_j \alpha^2}{R^2} + \frac{R_i R_j \alpha^3}{R^3}) + \dots] \end{aligned} \quad (6)$$

of which the s-wave part is

$$\begin{aligned} \Psi_s(r) = & 1 - \alpha r + \frac{\alpha^2 r^2}{2} + \sum'_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} e^{-\alpha|\vec{R}|} \\ & + \sum'_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} e^{-\alpha|\vec{R}|} r^2 (\frac{\alpha^2}{6} - \frac{\alpha}{3R}) \end{aligned} \quad (7)$$

with the primed summation running over the fcc lattice of H centers excluding the point at the origin. From Schrodinger's equation we know that

$$\frac{\Psi_s(r)}{\Psi_s(0)} = 1 - r - (E(\vec{k}) - V_0 - 2)r^2 + \dots \quad (8)$$

and by equating the coefficients of the linear terms in Eqs. (7) and (8) we obtain a transcendental equation which determines α :

$$\alpha = 1 + \sum'_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} e^{-\alpha|\vec{R}|} \quad (9)$$

From the coefficients of the quadratic terms we have

$$E(\vec{k}) = V_0 + 2 - \alpha [3 + \sum'_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} e^{-\alpha|\vec{R}|} (1 - \frac{2}{\alpha R})] \quad (10)$$

which determines the band energy $E(\vec{k})$ to within the additive constant V_0 . From the solutions of Eqs. (9) and (10) at $\vec{k} = 0$ and $\vec{k} = 2\pi/a(100)$ a valence band width $E(X) - E(T) = 0.21$ rydbergs is found. This result is in nearly perfect agreement with the corresponding result of an earlier cellular calculation⁶.

Future Developments

Particularly in view of the uncertainty involved in the prescription for the one-electron potential, it is fair to say that methods in band theory already widely in use are adequate to solve the energy band problem. For applications involving charge densities, momentum densities and matrix elements, however, the most simple compact representation of the Bloch functions may be very advantageous. If only to enhance the feasibility of carrying out problems involving several atoms per unit cell, impurities, or the electron phonon interaction, it will be desirable to arrive at the most simple successful form of the LCAO method. Since it has already been shown that energy band methods have a contribution to make to the theory of the electronic spectra of molecules⁷, it is to be expected that some form of LCAO-cellular method can also contribute.

If it should develop that the LCAO-Taylor series method is successful for a wide variety of examples, we will then be able to side-step the question of how to construct a one-electron potential. Our choice of basis functions based on the atomic self-consistent field results would then specify the potential implicitly in terms of a few parameters which in turn would be determined from Schrodinger's equation for the one-electron problem as illustrated above.

If, on the other hand, it should be found desirable to use another form of the LCAO-Cellular method, it will be necessary to have a prescription for the exchange potential. If we employ the somewhat artificial device of restricting an exchange potential formula to an atomic cell, it is fairly easy to devise an approximate exchange or exchange-correlation potential. In the spirit of the potential suggested by Liberman⁸ which is exact for the hydrogen atom, we might consider an exchange potential which is exact for both the hydrogen atom and the Hartree-Fock ground state of the helium atom. This potential is clearly the negative of the Coulomb potential produced by a 1s electron. It has the advantages of having a finite exchange charge density at the nucleus and the asymptotic form at large r in a free atom needed to produce the series of Rydberg excited states. In the crystal, a smoothed version of this exchange potential could be obtained by multiplying the exchange potential for each atom by a Fermi function of r which, for example, might be equal to one half at the Wigner-Seitz sphere radius, and close to unity at the atomic sphere radius. The exchange potential for the crystal could then be a lattice sum of atomic exchange potentials.

References

1. E. E. Lafon and C. C. Lin, "Energy Band Structure of Lithium by the Tight-Binding Method", Phys. Rev. 152, 579 (1966).
2. E. E. Lafon and C. C. Lin, "Recent Developments in Applying and Extending the Method of Tight Binding (LCAO) to Energy Band Calculations", Proceedings of this Conference.
3. R. C. Chaney, T. K. Tung and C. C. Lin and E. E. Lafon, "Application of the Gaussian-Type Orbitals for Calculating Energy Band Structures of Solids by the Method of Tight Binding", J. Chem. Phys. 52, 361 (1970).
4. P.-O. Löwdin, "Quantum Theory of Cohesive Properties in Solids", Adv. in Phys. 5, 96 (1956).
5. M. I. Chodorow, "The Band Structure of Metallic Copper", Phys. Rev. 171, 675 (1939).
6. D. S. Ewing and F. Seitz, "On the Electronic Constitution of Crystals: LiF and LiH", Phys. Rev. 50, 760 (1936).
7. K. H. Johnson, "Bands Bonds and Boundaries", Proceedings of this Conference.
8. D. A. Liberman, "A Potential Function for Band Structure Calculations", Proceedings of this Conference.

4. INTERPOLATION AND INTEGRATION IN k -SPACE

INTERPOLATION AND k-SPACE INTEGRATION: A REVIEW*

F. M. Mueller

Argonne National Laboratory

Argonne, Illinois 60439

ABSTRACT

The principal techniques used to form representations of integrals over the Brillouin zone which include Dirac δ -functions in their integrands are reviewed. The problem was first solved using Monte-Carlo procedures applied directly to the Hamiltonian. Recent interest has centered on two methods which utilize microscopic interpolation. The singular integration is carried out either through Monte-Carlo procedures or through a procedure which uses a further linear expansion and numerical integration. Both of these procedures represent their results in terms of histograms. Techniques for including matrix elements in the integrand are considered and new results presented. A new technique is given which uses high order Hermite functions to numerically integrate the principal-value kernel.

*Work performed under the auspices of the United States Atomic Energy Commission.

The properties of real materials found in the laboratory may be viewed as a response to some externally applied probe. The calculation of such a response leads, in a natural way, to a thermodynamic average over the system. In the case of a classical system this average is given by a phase space integral, whereas in a quantum mechanical system a weighted statistical average over the states is used. For a periodic system such as a pure metal or insulator such a thermodynamic average over states may be replaced by the familiar integral over the Brillouin zone. Even when the relatively simple approximation of zero temperature is imposed, such averages over the Brillouin zone cannot, in general, be evaluated in closed form. One must resort to numerical methods.^{1,2}

For concreteness a model process $P(\underline{r}, t)$ can be considered in which the imaginary or absorptive part is calculated at a point in the system \underline{r} at time t . The real part may be found as Hilbert transform. The Fourier transform, in momentum \underline{q} and energy w , of the model periodic disturbance is given by

$$P(\underline{q}, w) = \frac{1}{(2\pi)^3} \sum_{if} \int |M_{if}(\underline{k}, \underline{q})|^2 (n_i - n_f) d\underline{k} \delta(E_i(\underline{k}) - E_f(\underline{k} + \underline{q}) - w) \quad (1)$$

where M is the matrix element between an initial Block state i at \underline{k} and a final state f at $\underline{k} + \underline{q}$; n_i is the thermodynamic distribution function; and the Dirac delta function conserves energy in the model process, which we assume represents elastic scattering.

In considering the scattering of elementary particles, in which the initial and final states are well-represented by plane waves, the matrix elements M^2 of eq. (1) are generally approximated either as a local interaction $M^2(\underline{k} - \underline{q})$ or as an average $\langle M^2 \rangle$. One then focuses on the energy structure as provided by the Dirac delta function. In solids, this plane wave approximation, which ignores matrix elements, is valid in a surprisingly large class of materials. Essentially the approximation is valid for materials to which the pseudopotential³ method may be applied.

However, Fig. 1 shows a case where the plane wave approximation does not apply. The figure shows the strong \underline{k} -dependence of the interband oscillator strength of copper as calculated by Mueller and Phillips.⁴ The electronic bands of copper are made up of localized (in \underline{r}) d electrons and broad plane wave electrons, and the mixture of these two basic types changes due to the strong \underline{k} -dependence of the hybridization throughout the Brillouin zone. As a result, the interband oscillator strength shows at least as much structure as the energy bands themselves.

A first example of eq. (1) is provided by a case in which $M^2 = 1$ is strictly valid. The initial state has an N -fold degeneracy at

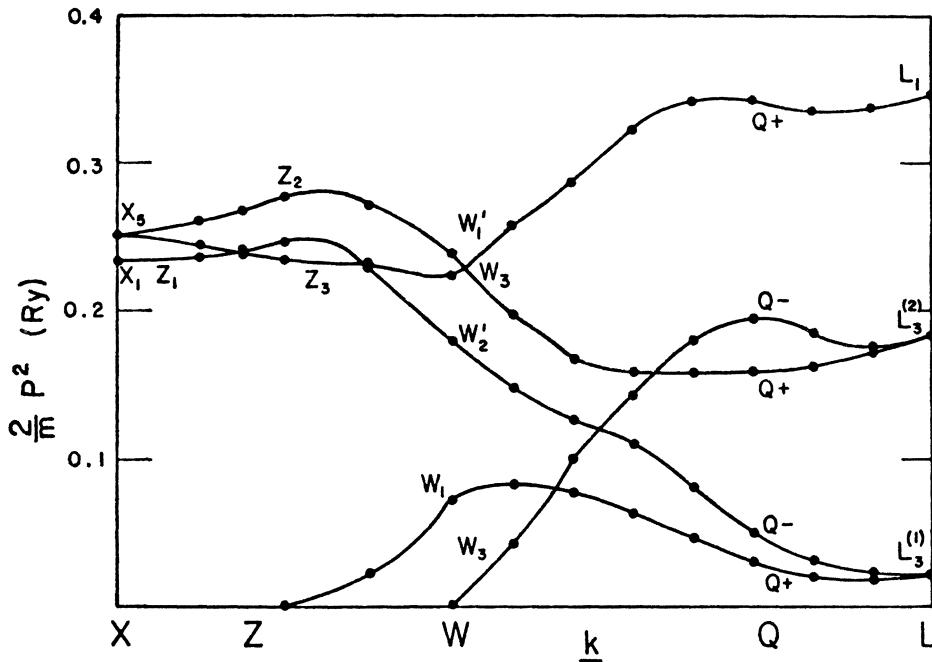


Fig. 1. The interband momentum matrix elements P_{ij}^2 scaled by $2/m$ are shown in Ry plotted against k around the outside of the Brillouin zone from X to W to L . Here i ranges over the first 5 d bands, while j is fixed to be the lowest conduction band. Thus the oscillator strengths at X correspond to $X_5 \rightarrow X_4$, and $X_5 \rightarrow X_1$ transitions. The upper L_3 state is denoted by $L_3^{(2)}$. The heavy-mass responsible for the 2-eV absorption edge corresponds to $X_5 Z_2 W_1, Q+ L_3^{(2)}$.

energy of $0-\epsilon$, the Fermi energy is $E_F = 0$, and the final states are created uniformly in \underline{r} (or $\underline{q} \rightarrow 0$). Equation (1) then reduces to simply the density of states g at energy w or to

$$g(w) = \frac{1}{(2\pi)^3} \sum_f \int d\underline{k} \delta(E_f(\underline{k}) - w). \quad (2)$$

If the dispersion relation $E_f(\underline{k})$ is particularly simple, then the three dimensional integral in eq. (2) may be reduced to the two dimensional form:

$$g(w) = \frac{1}{(2\pi)^3} \sum_f \int \frac{dS}{|\nabla_{\underline{k}} E_f(\underline{k})|} |E_f(\underline{k})| = w \quad (3)$$

and the surface integral in eq. (3) may be calculated analytically. The possibility of a zero denominator in eq. (3) exposes the singular nature of $g(w)$.

The analytic treatments of van Hove⁵ and of Phillips⁶ show that there exist four types of critical points corresponding to the four possible combinations of signs which the expansion coefficient for a quadratic form may take in three dimensions. Moreover, Phillips has shown from group theoretic and topological⁷ considerations alone, that the number of each of the four types of critical points for a given band is restricted both in their differences and in their total number. The restriction is given by a sum rule which Phillips calls the "minimal" set.

In the literature, critical structure in the density of states or in optical spectra is frequently attributed to single critical points, often symmetry points. However, simple counting considerations alone show that any really large peak must be drawn from equally large segments of the Brillouin zone. Moreover, because of the hyperboloidal dependence of the dispersion relation near critical points, the structure tends to be extended non-spherically. Such wide \underline{k} -dispersal drastically modifies the validity of the $\langle M^2 \rangle = \text{constant}$ approximation. Furthermore, because of such wide \underline{k} -dispersal, the examination of energy structure along only single lines of the Brillouin zone -- perhaps along symmetry lines -- presents a distorted picture of corresponding critical point structure. One simply must examine all points globally in the Brillouin zone. Because such global examination is physically impossible, one resorts to some form of three dimensional interpolation. Thus the problem of calculating quantities of physical interest has two distinct aspects: (1) interpolation of the \underline{k} -structure, and (2) the integration of singular functions based on this structure. This dual aspect is discussed further below.

The analytic treatments of van Hove and Phillips conveniently represented, or evaluated, $g(w)$ as a piece-wise continuous function.

For example, $E^{1/2}$ was used for free-electrons. Generally the complications of the realistic dispersion relations force the use of less compact forms. Three different techniques appear in the literature for representations: (1) point functions usually defined at a sequence of uniformly spaced values in w ; (2) histogram representations in which $g(w)$ is averaged between a sequence of values w_1 and $w_1 + \Delta$; and (3) moment-trace methods in which $g(w)$ is treated as a distribution function represented by its n th order moments $\mu_n = \int dw w^n g(w)$. In effect, the latter technique expands $g(w)$ in a complete set of functions ψ_m . If the distribution is bounded by some highest value w_h , then the complete set of Legendre polynomials P_m may be used as

$$g(w) = \sum_m a_m P_m(w/w_h) \quad (4)$$

as was first noted by Montroll.⁸ As discussed below, a representation in terms of the complete Hermite function $\psi_m(w)$ deserves special attention because they have simple Hilbert transform properties

$$g(w) = \sum_m C_m \psi_m(w). \quad (5)$$

As is true of any ordered polynomial set, the expansion in terms of Hermites preserves a simple relation between the expansion coefficients C_m and the moment μ_n of $g(w)$. Although any of the three techniques may be used as a valid representation of $g(w)$, the Hermite expansion (5) is likely to be especially useful. A more complete discussion of Hermite functions is given in the Appendix.

Historically, the numerical calculation of densities of states can be separated into two periods; one before and one after the advent of large-scale computers. The calculations of Wohlfarth and coworkers⁹ are representative of the earlier era. Rather modest (5x5) tight-binding secular equations were solved at a few (85) points in the Brillouin zone. Estimates of the critical-point structure were used to introduce the proper analytic singularities into the resulting density of states curves. A resurgence in popularity of localized methods, some of which are discussed elsewhere in this volume, has caused a re-examination of this technique. In the recent calculations, a density of states histogram is built out of eigenvalues found at many random points. Since the Hamiltonian itself forms the interpolation scheme, band crossings and eigenvectors are treated in a completely natural manner and are continuous functions. Thus the resulting programs are extremely simple to code and debug. However, the necessarily large number of diagonalizations makes for a slow and tedious process compared to functional evaluation. The disadvantage probably outweighs the advantage, but, as always, one's budget must be balanced against

the desired accuracy. As each new generation of computers becomes faster, objections based on inefficiency will become less serious.

After large-scale computers came into general use, histogram densities of states were first calculated using a wide variety of band structure methods. Again the Hamiltonian itself was used as the interpolation function. In these large-scale calculations, the basis set was made so large that all bands were well-converged--even in the case of transition metals. Of course, modern tight-binding calculations discussed above also contain this convergence feature.

The major objection to this group of calculation is the use of a special (regular) sample-point distribution function. Also the relative strengths of the eigenvalues are reweighted by the group degeneracy of the sample points. Both of these aspects are numerically incorrect. Best results are achieved if equally weighted, random points are used. A priori this seems paradoxical since an integration using regular sampling, in one dimension converges as $1/N$ where N is the sampling frequency, whereas Monte-Carlo integration converges as $1/\sqrt{N}$. However, when regular sampling is used in band-structure evaluations, a problem arises because of the confluence of constant energy contours of a given energy and the regular sampling sequence. The sampling becomes "unfair" in the energy space in the same way that Moiré patterns are produced by the correlations of overlapping, similar structures. Thus, although densities of states produced by random points have a slower rate of convergence than those using regular points, they cannot suffer from unwarranted correlations. As Mueller *et al.* have discussed,¹⁰ the correlation problem may be substantially reduced by sampling on a net whose frequency is a prime number.

The idea behind the variable reweighting scheme is that the sample point is to be "smeared out" to fill a cubic volume. The reweighting scheme simulates the truncation of the volume by using the wedge planes to form the irreducible portion of the Brillouin zone. However, as the number of sampling points is increased or decreased the ratio of reweighted to standard weight points varies as a surface-to-volume ratio. In the limit of infinitely fine subdivisions all sampling points would be weighted by the full group degeneracy factor of 48 for cubic groups. The reweighting thus exaggerates the accuracy of the calculation by a factor of $\sqrt{48}$. The argument can be condensed as follows: reweighting of special symmetry points and lines are aspects of sets of measure zero and hence cannot enter in a discussion of a volume property such as the density of states.

Then came the "modern methods" in which a functional representation is used to speed up the interpolation problem. The methods differ among themselves principally in the method used to perform

the singular integrals. Two methods appear extensively in the literature. The most popular method, the "Gilat-Raubenheimer" method, is based on an exact integration technique¹¹ in which a local, linear approximation to the energy structure is used. Since detailed presentations of this method appear in other articles in this volume, the method needs but brief attention here. Since the integration is exact, the only problems encountered are in the validity of the linear approximation. As Janak, for example, discusses in the next article,¹² the validity of the linear approximation may be substantially improved by simply using integration cells of extremely fine subdivisions. However, higher order terms are always present. The effect of such terms, although difficult to accurately quantify, is to qualitatively shift the energy of the constant energy contour away from the "proper" value by a small amount. The consequences of the error in cross-sectional area are much smaller. In effect, then, the energy grid, and the resulting calculation, is blurred to this order. Because of this blurring effect the word "histogram" is used in the abstract to this paper.

The second method which uses local interpolation is generally known as the "QUAD" scheme.^{10,13} Its primary difference from the Gilat-Raubenheimer method is in the final integration. That integration is performed using a Monte-Carlo technique, where the energy of the n th band at a given random k is found from a quadratic interpolation of a set of grid points. The errors in the procedure are twofold: a quadratic interpolation function is only accurate to second order; and the Monte-Carlo integration procedure introduces a statistical fractional error which is proportional to $1/\sqrt{N_i}$, where N_i is the occupation number of the i th histogram element. For a basic accuracy of the order of 1% (roughly comparable to the best experiments), approximately 10,000 pieces of statistical information are needed per histogram box. If about five bands contribute in such an energy region, the statistical accuracy will be met by sampling of the order of 1,000,000 individual points. Such a sampling process for a typical transition metal takes on the order of 10 minutes of computer time on the IBM 360/50/75 at Argonne National Laboratory. The more difficult error to estimate is the accuracy of the quadratic interpolation procedure. The 10 full-quadratic expansion coefficients are derived by a least-squares fitting procedure to 27 cubically ordered expansion points. Table I presents two estimates of the interpolation error. The estimates are derived from energy band calculations of platinum produced by the combined interpolation scheme. Column D gives the total rms deviation between the original grid of energies and the quadratic approximation. In order to "resolve" structure the deviation should be less than about 1/3 of the histogram width.

More important than the simple energy deviation, which is a "local" fitting error, is the global effect of the quadratic approximation on the density of states. In order to estimate this

TABLE I. Specification of MESH parameters.

$\Delta = 0.001 \text{ Ry}$		NTOT = 1,000,000		NB ϕ X = 600	
		Eigenvalues		Eigenvectors	
MESH	NDIAG	D(MESH) (m Ry)	ξ (MESH) (%)	A(MESH) (%)	C(MESH) (%)
1	6	60.47	--	52.71	--
2	20	20.14	27.51	42.71	51.24
3	46	12.28	12.04	22.63	21.43
4	89	5.73	4.37	12.47	9.71
5	152	2.72	3.69	7.19	6.37
6	240	1.47	2.17	5.05	5.61
7	356	0.79	1.82	3.83	4.93
8	505	0.51	1.49	3.27	4.61
9	690	0.40	1.32	2.70	4.38
10	916	0.31	1.13	2.49	4.12
11	1186	0.27	1.07	2.32	4.01
12	1505	0.22	1.00	2.19	3.95
133	1876	0.19	0.97	1.99	3.76
14	2304	0.16	0.94	1.95	3.72
15	2792	0.14	0.90	1.79	3.68
∞	∞	0.00	0.77	0.00	0.77

effect, a series of calculations of the density of states of platinum were performed. The calculations differed only in the fineness of the MESH parameter M . The correlation function

$$\xi^2(\text{MESH}) = \frac{\sum_n (H_M(n) - H_{M-1}(n))^2}{\sum_n (H_M(n) - H_{M-1}(n))^2} \quad (6)$$

calculated to determine the effect of the MESH parameter M is also given in Table I, where $H_M(n)$ is the n th element of a histogram H at mesh M . Note that, because of the Monte-Carlo integration error, ξ as defined above is never zero. The error due to the quadratic approximation alone is roughly proportional to $\sqrt{\xi^2(M) - \xi^2(\infty)}$. Figure 2 presents the density of states in the d -band region for platinum as produced by a MESH parameter of 13, a histogram width of 0.001 and 1,000,000 Monte-Carlo points.

A distinctive feature of the "QUAD" scheme is the technique used to include the effects of the strongly k -dependent matrix elements. The simplicity of the Monte-Carlo integration procedure allows the k -dependence of M to be included at essentially no extra cost in either programming, or in running time. Since M transforms as a scalar, additional quadratic expansions are available to represent M and the elements (determined by energy consideration) added to the histogram may be weighted by $M(k)$. The procedure allows for separate energy and matrix element variations. The error of such a calculation is made up of three components. The Monte-Carlo and energy errors occur as in the case of constant matrix elements. A new source of error is the inaccuracy of a quadratic approximation for M . The next most complicated function after the density of states, namely the basis function density of states, was used to systematically test for the latter error. The function is given by

$$g_s(w) = \frac{2}{(\pi)^3} \sum_n \int dk |S|_k|^2 \delta(E_n(k)-w) \quad (7)$$

where here S is the lowest plane wave basis function, and $|k\rangle$ is the general Bloch wave for the n th band. A sum of eq. (7) over all basis functions would, of course, yield the total density of states, Fig. 2. The result of such a projection is illustrated in Fig. 3.

Using the projected density of states shown in Fig. 3, one may define a rms error, $A(M)$ and a correlation function $C(M)$, analogous to $D(M)$ and $\xi(M)$, respectively. Table I illustrates that quantities of physical interest may be calculated to the order of 4% accuracy with rather modest numbers of grid points. Since a full diagonal-

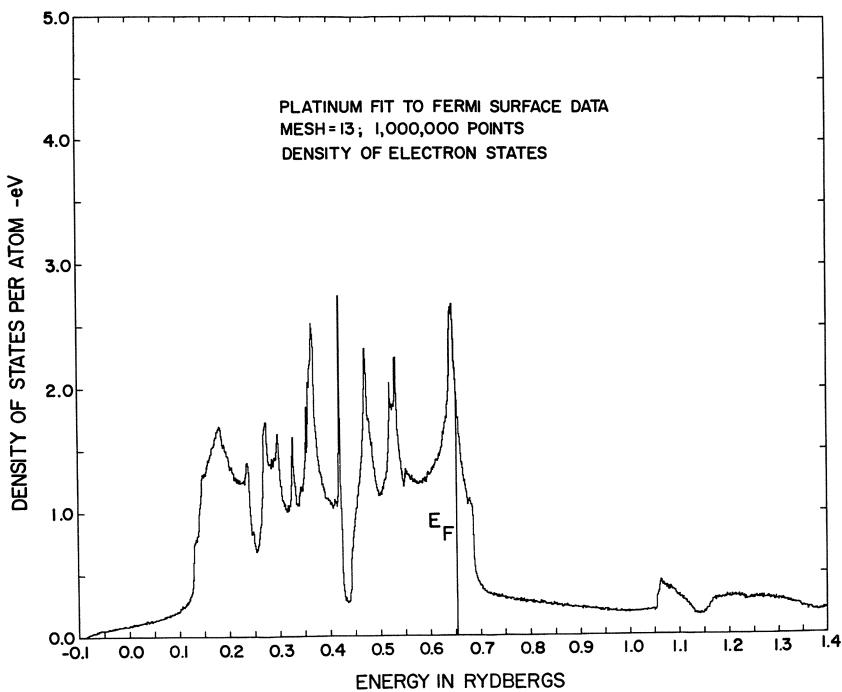


Fig. 2. The density of states of platinum.

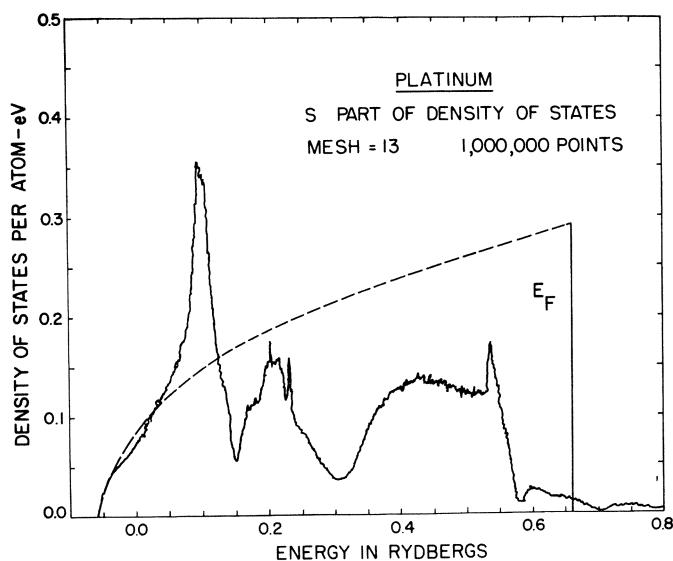


Fig. 3. The S part of the density of states of platinum.

ization of the secular equation, including eigenvectors, must be performed at each individual k point the number of grid points used must be modest. The larger error in the eigenfunction expansion, as compared to the eigenvalues, is due to the fact that eigenfunctions are discontinuous at band crossings whereas the eigenvalues are piecewise continuous. A sequence of parabolic segments defined on a fine grid approximates such discontinuities rather poorly. This defect is present in all functional interpolation methods, but is completely absent in the Hamiltonian schemes. However, the error in the QUAD scheme is no worse than that given in Table I, and this amount is tolerable. The accuracy of the QUAD scheme might be improved by analytically integrating the quadratic expansion coefficients. Computationally analytical integration leads to the problem of the evaluation of $\text{MESH}^{**} 3 * \text{NW} * \text{NB}$ incomplete elliptic integrals, where NW is the number of histogram boxes and NB is the number of bands. The time required for transition metals would probably be of the order of three to five hours on an IBM 360/50/75, depending on the techniques employed. As things stand now, however, the more expensive Hamiltonian schemes must be used to attain accuracy substantially greater than is provided by QUAD.

Essentially, then, the methods used to deal with integrals such as eq. (2) divide into three types: those, such as the root-sampling method, which exploit a small secular equation and perform the integration of eq. (2) via Monte-Carlo sampling; those, such as the Gilat-Raubenheimer method, which exploit the local linear dependence of the energy bands to perform the integration analytically; and those, such as the QUAD scheme, which combine a local quadratic dependence with Monte-Carlo integration. The choice between these three methods should be based on minimizing both the computational cost and the real time needed to exploit a given method. The root-sampling method has the advantage of complete generality, but tends to be slow and not very accurate for statistical reasons. The linear expansion methods have great accuracy, but are not very general in the sense that a linear expansion can be inadequate, particularly for eigenvectors. The QUAD method is clearly a compromise method since it has good generality and speed, but is not as accurate (because of the Monte-Carlo sampling) as the linear methods.

The generality of the root-sampling method is most appropriate to explorations of structure in the early stages of an investigation. In later stages one of the other two methods should be employed to obtain a complete picture. At the moment, based on a criterion of accuracy per unit calculational time, the Gilat-Raubenheimer and the QUAD methods are roughly comparable.¹⁴ The Gilat-Raubenheimer may have a slight edge. On the other hand, the Monte-Carlo codes tend to be simpler to code and debug and therefore may make better use of the scientist's time. A fair comparison should be performed.¹⁴ In terms of practical use, these methods should now allow the

accurate calculation of a whole host of experimentally important quantities. Only the future will determine how this new ability to make accurate comparisons will aid the dialog between experiment and theory.

ACKNOWLEDGMENTS

I would like to thank R. Afshar, K. H. Bennemann, M. H. Cohen, J. W. Garland, J. F. Janak, J. C. Shaffer, and A. R. Williams for helpful discussions, the Computation Center at Argonne National Laboratory for excellent service, and Miss M. Katzin for programming assistance.

REFERENCES

1. L. D. Landau and E. M. Lifshitz, Statistical Physics (Addison-Wesley Publishing Co., Reading, Mass., 1969) second edition.
2. A. A. Maradudin, E. W. Montroll, and G. H. Weiss, Theory of Lattice Dynamics in the Harmonic Approximation (Academic Press, New York, 1963).
3. W. A. Harrison, The Pseudopotential Method (W. A. Benjamin, Inc., New York, 1966) and references therein.
4. F. M. Mueller and J. C. Phillips, Phys. Rev. 157, 600 (1967).
5. L. van Hove, Phys. Rev. 89, 1189 (1953).
6. J. C. Phillips, Phys. Rev. 104, 1263 (1956).
7. M. Morse, Functional Topology and Abstract Variational Theory (Gauthiers-Villars, Paris, 1939).
8. A. A. Maradudin, op. cit., p. 65 and references therein.
9. G. D. Fletcher and E. P. Wohlfarth, Phil. Mag. 42, 106 (1951), Proc. Phys. Soc. (London) A65, 192 (1952); J. F. Cornwell and E. P. Wohlfarth, J. Phys. Soc. Japan 17, (Suppl. B-1), 32 (1962).
10. F. M. Mueller, J. W. Garland, M. H. Cohen and K. H. Bennemann, to be published.
11. G. Gilat and L. J. Raubenheimer, Phys. Rev. 144, 390 (1966); a similar method has been implemented by L. R. Saravia and D. Brust, Phys. Rev. 170, 683 (1968); see also other papers of this conference.
12. J. F. Janak, to be published.
13. F. M. Mueller, J. W. Garland, M. H. Cohen and K. H. Bennemann, Argonne Report ANL-7556, unpublished. This report contains all of the Fortran programs used in the QUAD scheme.
14. J. F. Janak and F. M. Mueller, private communication.
15. R. Afshar, F. M. Mueller and J. C. Shaffer, ANL Report.
16. R. Afshar, F. M. Mueller and J. C. Shaffer, to be published.
17. A. R. Williams and F. M. Mueller, private communication.

APPENDIX: Some Properties of Hermite Functions

The complete set of Hermite functions is relatively familiar to solid state physicists since they are the eigenfunctions of the simple harmonic oscillator:

$$(p^2 + kx^2)\psi_n = E_n \psi_n \quad (A1)$$

where the eigenvalues are

$$E_n = (n + 1/2)\hbar w_c, \quad w_c = \sqrt{\frac{k}{m}} \quad (A2)$$

and the eigenfunctions are

$$\psi_n = N_n H_n(x) e^{-x^2/2} \quad (A3)$$

where

$$N_n^2 = (2^n n! \sqrt{\pi})^{-1} \quad (A4)$$

and H_n is the n th order Hermite polynomial.

The utility of these functions stems from their orthonormality over the whole real line. Moreover, through the use of Hermite integration (the analog of Gaussian integration over the whole, real line), and a knowledge of the zeros of high order Hermite polynomials, the numerical problems of expanding functions of interest to solid state physicists in Hermite functions are made minimal. A separate article gives the zeros¹⁴ and weight function for Hermite integration of order $n = 20, 26, 50, 76, 150$, and 300. Also available are those of the 500th order, which is probably high enough for any problem of practical interest. The Hermite function Fortran programs will be published as an Argonne report,¹⁵ and are available on request. Because there are no special problems in generating these functions, the simple, upward recursion relation was used:

$$H_{n+1}(x) = 2x H_n(x) - 2n H_{n-1}(x). \quad (A5)$$

The only problem is the factor of $n!$ in the Hermite polynomials, which was treated by extending the exponential part of the double precision work of an IBM 360/50/75 to include arbitrarily high powers of 16. To avoid overflow/underflow problems the functions are replaced by zero if they are smaller than $16^{**} (-17)$.

The integral of an arbitrary function $g(x)$ is approximated by

$$\int_{-\infty}^{\infty} g(x) dx \approx \sum_{i=1}^n W_n'(x_i^n) g(x_i^n) \quad (A6)$$

where x_i^n is the i th zero of the n th Hermite polynomial, and

$$W_n' (S_i^n) = (n \psi_{n-1}^2 (S_i^n))^{-1} \quad (A7)$$

is the weight function.

Of special interest is the expansion of two related functions: the first is the Hilbert transform operator. The antisymmetrical kernel can be expanded in a pair of Hermite functions as:

$$\pi \frac{P}{(w-w')} = \sum_{nm} S_{nm} \psi_n(w) \psi_m(w') \quad (A8)$$

where P means that the principle value is to be taken of any integrals on the lhs of (A8). Analysis shows that S_{nm} is an anti-unitary matrix and is given by

$$\begin{aligned} S_{nm} &= (-i)(i)^n (-i)^m \int_{-\infty}^{\infty} dt \operatorname{SGN}(t) \psi_n(t) \psi_m(t), \\ &= (i)^{n-m-1} S'_{nm} \end{aligned} \quad (A9)$$

and the integral S' has a (symmetric) series representation as

$$S'_{nm} = \sqrt{\frac{2n! m!}{\pi}} (-1)^{n+h} \times \sum_{r=0}^n \frac{(-1)^r (2n+2h-2r-1)!!}{r! (n+2h-r+1)! (n-r)!} \quad (A10)$$

where one assumes $n < m = n+2h+1$. The latter relation, although correct in principle to all orders, becomes inaccurate above $n, m \geq 35$. Therefore, S_{nm} was found from (A9) by simply integrating using the zeros of the 500th order Hermite functions. The latter technique is extremely fast and up to order $n+m = 500$, is exact.

The kernel (A8) simplifies matters. If one expands a function in Hermite functions as:

$$g(w) = \sum_n c_n \psi_n(w) \quad (A11)$$

then the Hilbert transform is simply given as

$$F(w) = \sum_m d_m \psi_m(w) \quad (A12)$$

where

$$d_m = \sum_n S_{nm} c_n. \quad (A13)$$

An example of this technique is the Hilbert transform of the triple cosine density of states¹⁶

$$g(w) = \frac{2}{(2\pi)^3} \int dk \delta(E(k)-w) \quad (A14)$$

where

$$E(k) = 0 - \cos(k_x a) - \cos(k_y a) - \cos(k_z a) \quad (A15)$$

and where the Hilbert transform is

$$F(w) = \frac{P}{\pi} \int \frac{dw'}{w-w'} g(w') . \quad (A16)$$

These have the analytic representations:

$$\left\{ \begin{array}{l} g(w) \\ F(w) \end{array} \right\} = \frac{2}{\pi} \int_0^\infty dt J_0^3(t) \left\{ \begin{array}{l} \cos(wt) \\ \sin(wt) \end{array} \right\} . \quad (A17)$$

Figure 4 gives the result of integrating (A17) at the 250 positive zeros of the 500th order Hermite function.

Although the relations (A11) and (A12) are exact in principle, the convergence of the expansion in Hermite functions may be substantially improved by removing the low order moments from $g(w)$ and transforming these analytically. Then a Laurent expansion of $F(w)$ will begin with order $m+1$ as

$$\frac{P}{w-w'} = \frac{P}{w} \sum_{m=0}^{\infty} \left(\frac{w'}{w} \right)^m , \quad (A18)$$

so that

$$F(w) = \frac{1}{\pi} \frac{P}{w} \sum_{m=0}^{\infty} w^{-m} \int dw' w'^m g(w') . \quad (A19)$$

In practice good results are achieved if only the zeroth moment is removed. From $g(w)$ in (A17) the triangle-like function $g_s(w)$ is subtracted where

$$g_s(w) = \frac{2}{3} [1 - \text{SGN}(w)w/3] . \quad (A20)$$

This function is also plotted in Fig. 4 ($\int g_s(w)dw = 2$). The analytic transform of (A20) is given by

$$F_s(w) = -\frac{2}{3\pi} \left[\ln \left| \frac{w-3}{w+3} \right| - \frac{w}{3} \ln \left| \frac{(w-3)(w+3)}{w^2} \right| \right] \quad (A21)$$

This is also plotted in Fig. 4. Note that to first order, the log poles in (A21) are subtracted out. The residuum of (A14) less (A20) has been expanded using the first 250 Hermite functions and a scale factor of $\alpha = 6/\sqrt{2N+1}$ where $N = 250$ so that, in effect, the highest

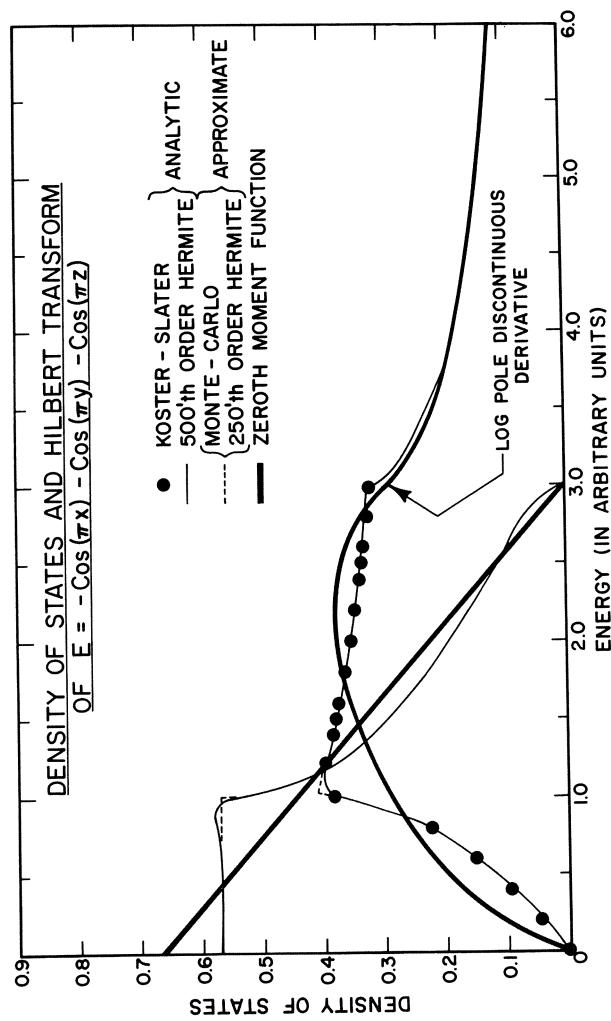


Fig. 4. The density of states and Hilbert transform of the triple cosine density of states.

TABLE II. The first few Hermite function expansion coefficients of the triple cosine density of states. As explained in the text, the scale factor $\alpha = 6/\sqrt{2N+1}$ was used where here $N = 150$. The coefficient for C_n (D_n) are non-zero for even (odd) order because they were derived from a Monte-Carlo function which was accurate to about 1%. Note that the convergence of D is much slower than C . Removing the zeroth moment from $g(w)$, as explained in the text, greatly improves this rate so that the expansion of the residual of D is essentially zero after $N = 75$.

n	C_n	D_n
0	1.0758	0.0010
1	-0.0019	0.2436
2	0.7517	-0.0008
3	-0.0009	0.3469
4	0.6029	0.0002
5	-0.0042	0.4376
6	0.4306	-0.0053
7	0.0007	0.4434
8	0.2842	-0.0045
9	0.0037	0.3696
10	0.2355	-0.0015
20	0.1263	($n+1$) 0.2815
40	0.0048	0.1926
60	0.0054	0.1232
80	-0.0002	0.0977
100	0.0001	0.0796
120	0.0008	0.0730
140	-0.0005	0.0693

functions just span the range from 0 to 6. This scaling avoids any "edge effects" due to a confluence of the upper level of the range of integration and the Hermite function expansion set. Table II lists the first 20 expansion coefficients.

As a speculation for the future, another Hermite expansion may be considered¹⁷-- here the expansion of the Dirac delta function is

$$\delta(E_n(\underline{k}) - w) = \alpha \sum_m \psi_m(\alpha E_n(\underline{k})) \psi_m(\alpha w) \quad (A22)$$

where α is an arbitrary scaling parameter. When the expansion is truncated at some high N , all structure on the lhs of (A16) will be resolved to the order of the last expansion coefficient N . Similar to the moment functions given above is the expansion coefficient of the density of states

$$C_m = \int_{-\infty}^{\infty} dw \psi_m(\alpha w) g(w) \quad (A23)$$

or, more directly,

$$C_m = \alpha \cdot \sum_n \frac{2}{(2\pi)^3} \int d\underline{k} \psi_m(\alpha E_n(\underline{k})). \quad (A24)$$

Finding the expansion coefficient through direct integration of (A18) should be especially valuable if the $E_n(\underline{k})$'s are available on a sufficiently fine grid. This method represents the answer, in effect, as an n th order Gilat-Rabenheimer technique. Moreover, it should be extremely simple to implement. Through a combination of (A24) and (A8), the susceptibility could be represented in (numerically) closed form, and without great complications.

GILAT-RAUBENHEIMER METHODS FOR k-SPACE INTEGRATION

J. F. Janak

IBM Thomas J. Watson Research Center
Yorktown Heights, New York

ABSTRACT

Application of the Gilat-Raubenheimer method to the density of states and its energy derivatives and to the photoemission distribution is discussed. Some estimates of the errors involved in the method are supplied, and a discussion of a polynomial interpolation scheme is given.

I. INTRODUCTION

The calculation of the density of states and other electronic properties of solids has traditionally been regarded as a more or less separate subject from the calculation of energy bands. Although the energy-band problem is a difficult one, it is important to recognize that the use to which band calculations are put has some bearing on how the calculations are done. For example, a large number of energy eigenvalues is required to produce an accurate density of states histogram.¹ The Gilat-Raubenheimer method² affords a means of obtaining the density of states to the same accuracy using fewer energy eigenvalues, but also requires the gradients of $E(k)$. The overall time needed to produce a density of states is reduced if the additional time required to construct the gradients is more than compensated by the reduction in the number of points for which the calculations must be done. The band calculator thus needs some idea of the requirements and relative merits of different schemes for computing observable quantities. This paper supplies that information for the Gilat-Raubenheimer method as applied to the density of states and its energy derivatives and for the photoemission distribution. Some properties of

polynomial interpolation schemes are also discussed.

II. THE GILAT-RAUBENHEIMER METHOD

To calculate the density of states

$$N(E) = \int d^3k \delta(E - E_n(\vec{k})) \quad (1)$$

One first divides the irreducible wedge (1/48 of the Brillouin zone for cubic crystals) into a large number of cubes. The histogram method can be thought of as follows; let the energy in each cube be constant, equal to the value $E_n(\vec{k}_0)$ at the cube center. To find the density of states at energy E , simply count all the cubes with energies $E_n(\vec{k}_0)$ in a range ΔE about E .

The Gilat-Raubenheimer (GR)² method starts by transforming the integral in eq. (1) to an integral over surfaces of constant energy:

$$N(E) = \int \frac{dS}{|\nabla_k E|} \quad (2)$$

$$E_n(\vec{k}) = E$$

In each of the small cubes, we make a linear expansion of $E_n(\vec{k})$ about the cube center:

$$E = E_n(\vec{k}_0) + (\nabla_k E_n)_{\vec{k}_0} \cdot (\vec{k} - \vec{k}_0) \quad (3)$$

The actual energy surface in the cube is replaced by the plane defined by eq. (3), and the integral in eq. (2) is replaced by a sum over cubes

$$N(E) \approx \sum_{\vec{k}_0} \frac{S(E, \vec{k}_0)}{|\nabla_k E_n|_{\vec{k}_0}} \quad (4)$$

Here $S(E, \vec{k}_0)$ is the area of the portion of the plane defined by eq. (3) within the cube and $|\nabla_k E_n|_{\vec{k}_0}$ is the magnitude of the gradient at the cube center. Once the function $S(E, \vec{k}_0)$ is known [Gilat and Raubenheimer^{2,3} give algebraic expressions for S in terms of $E - E_n(\vec{k}_0)$ and $(\nabla_k E_n)_{\vec{k}_0}$], one simply performs the sum over all cubes to get the density of states. In practice, of course, the density of states is to be computed at a fairly large number of energies; in this case, it is best to go through the

cubes only once, using eq. (3) to find all the energies to which a given cube contributes [if the cube edge is $2b$, eq. (3) implies that

$$|E - E_n(\vec{k}_0)| \leq b \left[\left| \frac{\partial E_n}{\partial k_x} \right|_{\vec{k}_0} + \left| \frac{\partial E_n}{\partial k_y} \right|_{\vec{k}_0} + \left| \frac{\partial E_n}{\partial k_z} \right|_{\vec{k}_0} \right] \quad (5)$$

and adding the resulting contributions to $N(E)$ to the appropriate components of a vector.

The basic idea is that, by making a first-order fit to the energy surfaces in each cube, instead of a zeroth-order fit, as in the histogram scheme, fewer cubes will be required (for the same accuracy in the final density of states) than in the histogram method. If this reduction in the number of cubes more than compensates for the additional computations required in each cube [supplying the gradient $(\nabla_k E_n)_{\vec{k}_0}$ and evaluating the area $S(E, \vec{k}_0)$], the algorithm will be faster (for any particular accuracy criterion) than the histogram method.*

It is very difficult to give reliable estimates of the error involved in the method. There are, however, two sources of error; the first is an assignment error arising from defining the energies by eq. (3). That is, the contribution of a certain plane to the density of states will be assigned by the method to an energy which is incorrect by an amount of order $\frac{1}{2} b^2 |\partial^2 E_n / \partial k^2|_{\vec{k}_0}$, where $2b$ is the cube edge. This energy error, which is just the missing quadratic term in eq. (3), can be thought of as arising because the plane defined by eq. (3), which is used to approximate the actual energy surface in a given cube, is not necessarily the tangent plane to the energy surface; the gradient in eq. (3) is evaluated at the cube center rather than on the energy surface.

* There is a whole hierarchy of algorithms, depending on the order of approximation used in each cube. For example, one could make a second-order fit in each cube, which would require still fewer cubes, but at the expense of supplying the gradients and the second derivatives of the energy. It does not appear possible to give an algebraic representation of the result in this case, so that there is the additional expense of a complicated numerical calculation, rather than a relatively simple function evaluation, in each cube. Nevertheless, the total number of cubes required would probably be small enough that the required energies and derivatives could be supplied directly from "first-principles" calculations (assuming that accurate and efficient methods exist for finding the derivatives) without any intermediate interpolation.

The second error is an area error; the area of the approximating plane need not be the same as the area of the actual energy surface in the cube. We can examine this error by projecting the actual surface (here assumed locally quadratic) onto the approximating plane:

$$\frac{dS}{|\nabla_k E_n|} = \frac{dS_{\text{plane}}}{(\vec{n} \cdot \vec{n}_{\text{plane}}) |\nabla_k E_n|} = \frac{dS_{\text{plane}}}{|\nabla_k E_n|_{k_0} \left[1 + (k_0 - k_0)_i \frac{\partial E_n}{\partial k_1} \frac{\partial E_n}{\partial k_2} \right]^{1/2}}, \quad (6)$$

so that

$$\int_{\text{Cube}} \frac{dS}{|\nabla_k E_n|} \approx \frac{S(E, \vec{k}_0)}{|\nabla_k E_n|_{k_0}} \left[1 - \sum_i \frac{\left(\frac{\partial E_n}{\partial k_1} \right)_{k_0} \left(\frac{\partial E_n}{\partial k_2} \right)_{k_0}}{|\nabla_k E_n|_{k_0}^2} \right] \quad (7)$$

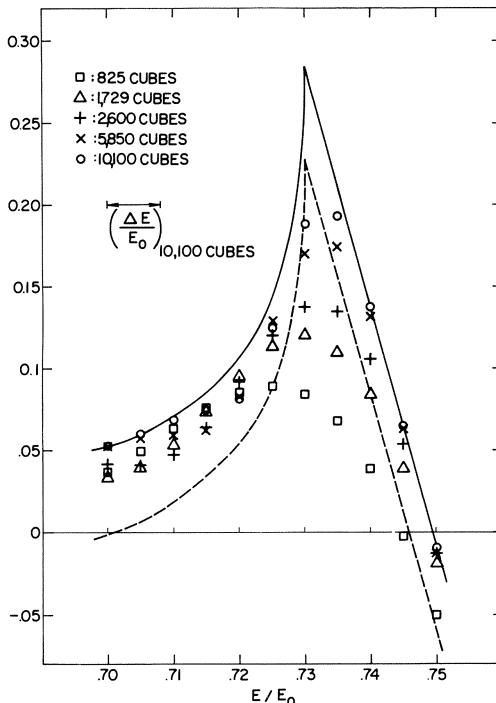


Figure 1. Test of GR method using 2-OPW band structure at L. Solid curve is exact $N(E)$ with free-electron density of states subtracted out; dashed curve represents 1% error in $N(E)$; points are computed by GR method. Energy units are such that $(\hbar^2/2m)(2\pi/a)^2 = 1$; V_{111} in these units is 0.02.

where $\vec{\xi}_c$ is the vector from the origin to the center of area of the approximating plane.⁴ Now $\vec{\xi}_c$ is $O(b)$, so the relative error in the density of states from a particular cube is $O(b)$. However, this error will tend to cancel from cube to cube because $\vec{\xi}_c$ will have different signs in different cubes, so the error in the total density of states will be more like $O(b^2)$ [It could, however, be as bad as $O(b)$]. To sum up, then, we expect the error in the GR method to go as $O(b^2)$ in general, with the possibility of less rapid convergence at isolated energies.

A realistic test of the method is illustrated in Fig. 1, which shows the density of states of the lower band arising from two OPW's degenerate at L in the face-centered cubic structure. The parameters in the 2×2 secular equation are appropriate to the splitting of the lowest aluminum bands at L .⁵ The solid curve in the figure is the exact density of states $N(E)$ for this problem, computed analytically, with the density of states of the free-electron sphere, $N_s(E)$, subtracted out. The dashed curve is $0.99N(E) - N_s(E)$, and any point below this curve will be in error by an amount exceeding 1% of the total density of states. The points are computed by the GR method using different total numbers of cubes in the irreducible wedge (the method of cube placement, and the relation between the cube edge $2b$ and the total number of cubes, are discussed in the appendix), and the figure thus gives a fairly reliable estimate of the number of cubes which is required to obtain the density of states of aluminum to within 1% by the GR method. One sees immediately that at least 5,000 cubes are required for high accuracy, and a significant improvement is obtained by going to 10,100 cubes. Even with 10,100 cubes, however, the error in the density of states of aluminum still exceeds 1% if one gets too close to the critical points.

A straightforward way to estimate the number of cubes required without making numerical tests is through the analysis of the assignment error given above. The maximum value of this error for a given band is

$$\Delta E \sim \frac{1}{2} b^2 \left| \frac{\partial^2 E_n}{\partial k_i \partial k_j} \right|_{\max} \quad (8)$$

To accurately plot the density of states of a band, the spacing between energy points will have to be a small fraction f of the band width W , and one wants to ensure that the assignment error is no larger (preferably considerably smaller) than the spacing between energy points. Thus

$$b \lesssim \left[2fW / \left| \frac{\partial^2 E_n}{\partial k_i \partial k_j} \right|_{\max} \right]^{1/2} \quad (9)$$

The value of ΔE obtained from eq. (8) for 10,100 cubes for the

2-OPW band just considered is shown in Fig. 1. It is almost twice the distance between plotted energy points, showing that the criterion given in eq. (9) is not particularly meaningful, and of course is relevant only to those energies where the second derivative is near its maximum value. Nevertheless, the criterion shows that the greatest difficulty will arise in transition-metal d-bands, where the bandwidth W is much smaller than for aluminum, but the curvature can be quite large near symmetry axes (such as Δ) where two d-bands are degenerate. In fact the Δ_5 states are degenerate all along Δ , so the difficulty will persist over the whole range of energies covered by the Δ_5 states; this can be a sizable fraction of W . All this suggests that a very large number of cubes will be needed to plot the density of states of the d-bands with high precision. On the other hand, the results (see Section III) for 10,100 cubes are quite good. Thus, although a large number of cubes is required for very high precision results, the convergence of the method is sufficiently uniform that good results can be obtained with a fairly small number of cubes, an effect which is apparent in Fig. 1.

III. INTERPOLATION

Use of the GR method requires knowing the energies $E_n(\vec{k})$ and gradients $\nabla_{\vec{k}} E_n(\vec{k})$ at $\sim 10^4$ points in the irreducible wedge for high-precision density of states calculations. Since most computers are not fast enough to supply this many energies (let alone the gradients) from a "first-principles" band calculation in a reasonable time, the procedure has generally been to supply the energies from a band calculation on a coarse mesh, and then to obtain the energies and gradients on the finer mesh required by the GR method by some form of interpolation. Note that this interpolation step is really independent of the GR method, and will eventually be eliminated by the development of faster computers and faster band-calculating algorithms. At present, however, it is a necessary step in any accurate density of states calculation. With the caveat that the GR method involves a completely different linear interpolation, we shall call the problem of supplying the energies and gradients at a sufficient number of points the "interpolation problem."

A short digression will point out some of the difficulties in the interpolation of energy bands. Imagine that two bands cross with different slopes at a point $(k, 0, 0)$ on the Δ -axis (Fig. 2a). Slightly off the Δ -axis, say along $(k, \delta, 0)$ with δ constant and small, the degeneracy will be lifted, and the bands will appear as in Fig. 2b. If one tries to label the bands so that they are smooth along Δ (i.e., so that each of the two straight lines in Fig. 2a is a band), the resulting bands will be discontinuous

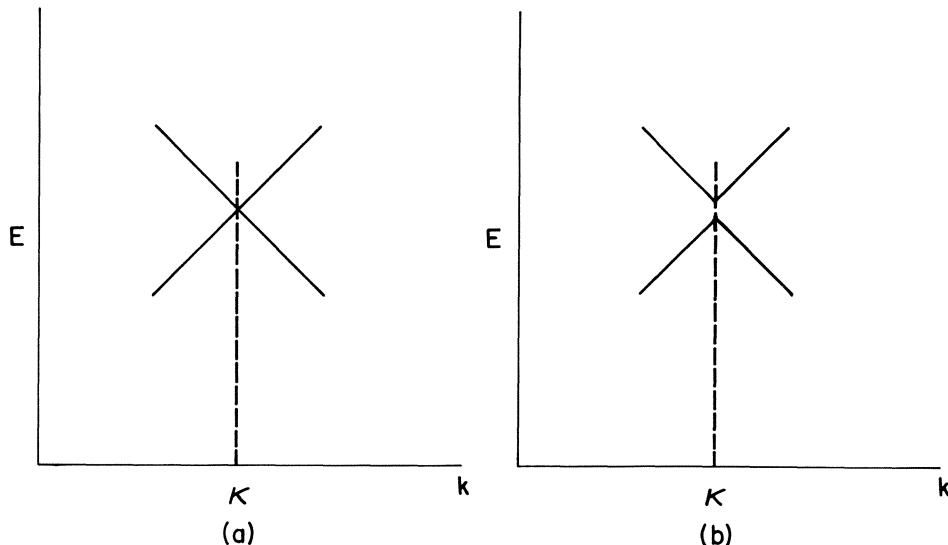


Figure 2. Band structure near a degeneracy. (a) E vs. k along the line $(k, 0, 0)$. (b) E vs. k along the line $(k, \delta, 0)$ where δ is constant and small.

slightly off Δ (either band will have to jump across the gap at the point K in Fig. 2b). If, on the other hand, one labels the bands at each point in reciprocal space in order of increasing energy, the resulting bands will always be continuous, but a forfeit must be paid in the form of bands which are not differentiable at degeneracies (the bands so labeled will have corners at the point K in Fig. 2a). The problem arises in all materials: although the best-known examples are the Δ_5 -states in the d-bands of transition metals, band crossings always arise at high enough energies in any material.

The problem of corners and band crossings is the root of the difficulty with interpolation. Determinantal interpolation schemes^{6,7} will represent these corners adequately, but finding the parameters in the determinant usually involves a nonlinear minimization procedure which must be done over again for each material (an exception is the pseudopotential description of nearly-free-electron metals). Determinant evaluation, furthermore, requires a relatively large amount of computer time. Polynomial interpolation schemes, on the other hand, are problem-independent and relatively fast, but can be made to accurately represent band corners only by interpolating from a fairly fine mesh of energies.

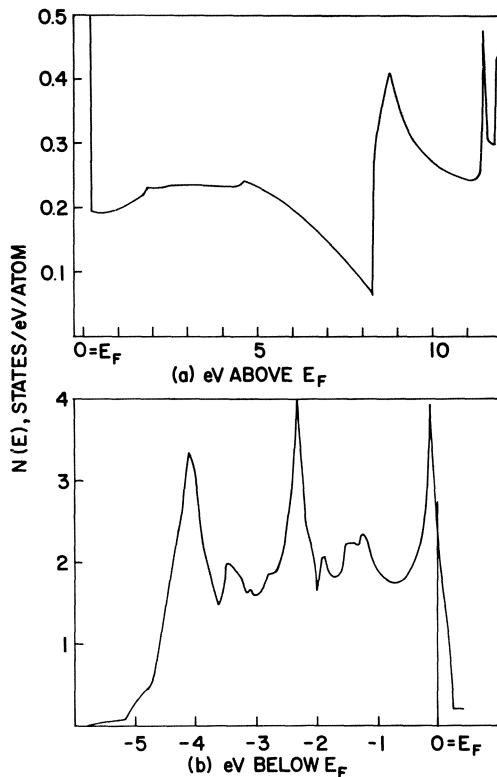


Figure 3. Density of states (both spins) of palladium. (a) $N(E)$ above E_F ; (b) $N(E)$ below E_F . $N(E_F) = 2.28$ states/eV/atom.

We have used polynomial interpolation to obtain the density of states of palladium^{8,9} (see Fig. 3). Energies on a simple cubic coarse mesh were obtained via the nonrelativistic KKR method¹⁰ from the Hartree-Fock-Slater potential for 4d¹⁰ palladium.¹¹ To interpolate to a point on the fine mesh, the closest point on the coarse mesh is first located, and the Lagrangian interpolation polynomial which fits the computed energies at this point and its 26 closest neighbors is then constructed and evaluated. The analytic derivative of the Lagrangian interpolation function is used to obtain the components of the gradient. This 27-point Lagrangian polynomial is the direct product of quadratic interpolations in each of the three cubic directions. If Δ is the spacing between points on the coarse mesh, one would expect the error in this form of interpolation to go as Δ^3 for analytic energy bands. Unfortunately, the presence of the corners in the d-bands of palladium leads to an error which decreases only as Δ .

To study the convergence of polynomial interpolation, the KKR calculations were run for various values of Δ down to $\Delta=1/32$ (corresponding to 3345 KKR points in the irreducible wedge). The energies and gradients were interpolated in all cases to the fine mesh of 10,100 points in the irreducible wedge used for the GR calculation of the density of states, and the results for various Δ 's were compared to the results for $\Delta=1/32$. (This procedure should give reliable values of the error for values of Δ not too close to 1/32, and the error near $\Delta=1/32$ can be estimated by extrapolation). In this way, it was found that, for $\Delta=1/32$, the rms error in the energies was $\sim 5 \times 10^{-4}$ Ry, the maximum error in the energy (which is the significant quantity for testing the interpolation procedure) was $\sim 5 \times 10^{-3}$ Ry, and the rms error in the magnitude of the gradient was $\sim 5 \times 10^{-2}$ Ry/($2\pi/a$). The maximum energy error was found to increase by a factor of about 3 for each doubling of Δ . Thus polynomial interpolation is a slowly-convergent procedure for d-bands: to get the maximum error in energy down to 0.01 Ry (which is still a quite large error), one would have to supply $\sim 1,000$ "first-principles" energies in the irreducible wedge. Another source of difficulty is the error in the gradients, which is naturally always larger than the error in the energies. The gradients enter the GR calculation both in the computation of the area, and in the denominator in eq. (4).

The errors induced in the density of states by errors in the area due to errors in the gradients is difficult to estimate; but the errors arising from the denominator of eq. (4) will tend to cancel, and the total error in the density of states due to this source is approximately proportional to the average error in the magnitude of the gradient. This is much smaller than the rms error quoted above.

As a test of determinantal interpolation methods, the density of states of aluminum following from Snow's energy bands⁵ (see Fig. 4) was computed using an interpolation scheme based on a non-local pseudopotential model.¹² A 4-OPW pseudopotential secular equation, in which all off-diagonal elements were taken to be constant, was set up and solved at the 89 points in the irreducible wedge for which energies are given by Snow. The solutions of this 4x4 determinant at the 89 points are then subtracted from Snow's energies. The resulting differences (different for each band) can be made very smooth functions of \vec{k} (constant to two figures) by choosing the constant parameters in the 4x4 secular equation to fit Snow's energies at the point W. To obtain an energy on the fine mesh (again comprising 10,100 points in the irreducible wedge), the difference function was interpolated to the fine-mesh point by 27-point Lagrangian interpolation, and then added to the solution of the 4x4 determinant at that point. The gradients were obtained by analytically differentiating the 4x4 determinant, a

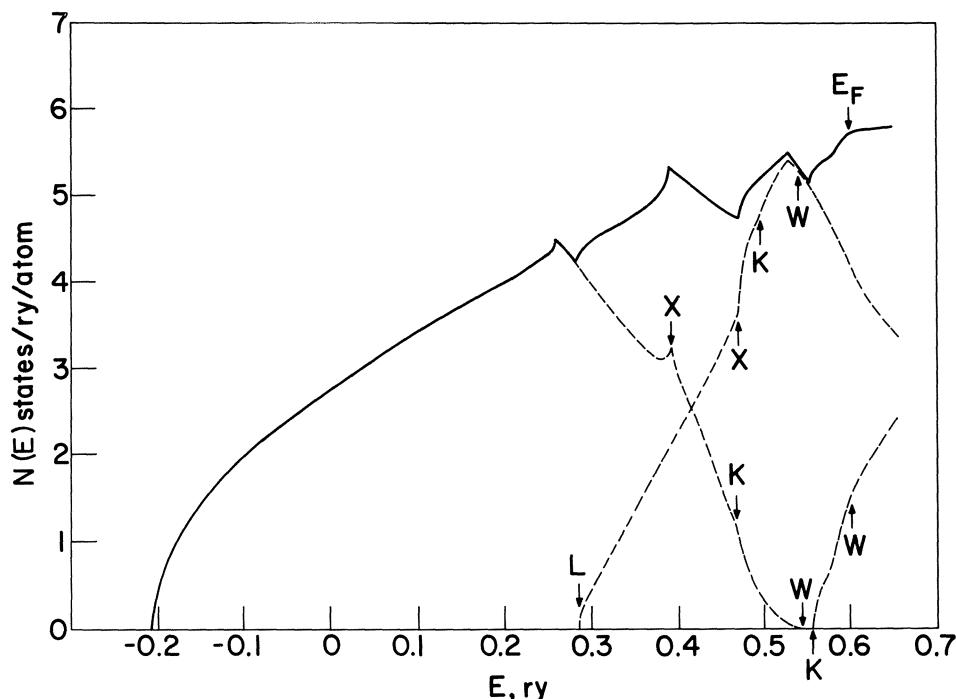


Figure 4. Density of states following from Snow's energy bands for aluminum.

relatively simple procedure in this case since only the diagonal elements are k -dependent. This interpolation scheme effectively subtracts from the given bands a known function (which is not interpolated) such that the difference (which is interpolated) is smooth. Since the difference turns out to be constant to two figures, it seems safe to assert that this scheme produces interpolated energies for aluminum which are accurate to at least three significant figures; the density of states shown in Fig. 4 is thus a high-accuracy result following from a self-consistent APW calculation; it is truly representative of the present state of the art.

IV. GILAT-RAUBENHEIMER METHOD FOR PHOTOEMISSION

If effects at the crystal surface are neglected and the electrons are assumed to have infinite lifetimes, the photoemission energy distribution, in the direct-transition model, is given by

an integral of the form

$$D(E, \omega) = \sum_{nn'} f(E - \omega) [1 - f(E)] \int d^3k |<n|\vec{p}|n'>|^2 \delta(E - E_n(\vec{k})) \delta(E - \omega - E_{n'}(\vec{k})) \quad (10)$$

where $f(E)$ is the Fermi function, and $<n|\vec{p}|n'\rangle$ the momentum matrix element. This is a very crude approximation to the actual photo-emission process, but it demonstrates the essential feature of a direct-transition photoemission calculation: the integral contains the product of two δ -functions. Just as one δ -function reduces an integral over the Brillouin zone to an integral over a surface, the integrand in eq. (10) is nonzero only when both δ -functions are satisfied, and the integral can be rewritten as a line integral along the curve of intersection of the two constant-energy surfaces defined by the δ -functions:^{13,14}

$$D(E, \omega) = \sum_{nn'} f(E - \omega) [1 - f(E)] \int dl \frac{|<n|\vec{p}|n'\rangle|^2}{|\nabla_k E_n \times \nabla_k E_{n'}|} \quad (11)$$

$E_n(\vec{k}) = E$
 $E_{n'}(\vec{k}) = E - \omega$

A straightforward generalization of the GR method can be used to compute the line integral. Again divide the Brillouin zone up into a large number of cubes, and replace the energy surfaces in each cube by planes. If two constant-energy planes intersect in a given cube, replace the integral along the curve of intersection of the two energy surfaces in that cube by an integral along the straight line of intersection of the two planes. Taking the matrix elements and the cross product of gradients equal to their values at the cube center, the integral in eq. (11) thus becomes the sum over cubes of the length of the line of intersection, weighted by $|<n|\vec{p}|n'\rangle|^2 / |\nabla_k E_n \times \nabla_k E_{n'}|$ evaluated at the cube center.

The method depends on having a simple algorithm for finding the length of that portion of an arbitrary line within a cube. This can be obtained⁹ as follows: let the line be given by

$$\vec{k} = \vec{d} + \vec{v}t \quad (\vec{d} \cdot \vec{v} = 0; v^2 = 1) \quad (12)$$

with respect to the cube center (Here, \vec{v} is a unit vector along the line, and \vec{d} is the shortest vector from the cube center to the line. It is not hard to find expressions for \vec{d} and \vec{v} in terms of the equations of the two constant-energy planes). If, as before, $2b$ is the cube side, a portion of the line will lie in the cube only if there is a range of values of t for which the three inequalities

$$-b \leq d_i + v_i t \leq b, \quad i = 1, 2, 3 \quad (13)$$

are simultaneously satisfied. Each of these inequalities defines

a lower limit t_i and an upper limit u_i on t . Define

$$t_1 = \max(t_1, t_2, t_3); \quad t_2 = \min(u_1, u_2, u_3). \quad (14)$$

Then, by examining how the three allowed ranges of t overlap, it is easy to see that a portion of the line lies within the cube if, and only if, $t_2 > t_1$; furthermore, the length of this portion is just $t_2 - t_1$. Thus we have a straightforward way of both finding if the two constant-energy planes intersect within the cube, and finding the length of the line of intersection if they do.

The algorithm described above has been used to calculate the photoemission spectrum of palladium in the direct-transition model (assuming constant momentum matrix elements)^{8,9}; discussion of a test of the algorithm appears in reference (9).

V. GILAT-RAUBENHEIMER METHOD FOR DERIVATIVES OF THE DENSITY OF STATES

The energy derivatives of the density of states at the Fermi level are of use in calculating the temperature dependence of properties of metals. We have seen that, even using the GR method with 10^4 cubes in the irreducible wedge, it is difficult to get the density of states to within 1% in the neighborhood of critical points. If the Fermi level happens to fall near a critical point, it would thus be dangerous to try to obtain derivatives of the density of states by numerical differentiation, although this procedure would probably be satisfactory in regions where the density of states is smooth and can be calculated to high accuracy with relatively few cubes. Many interesting cases, unfortunately, will be precisely those in which the Fermi energy is near a critical point, since it is only these cases in which anomalous behavior will occur.

An alternate procedure for obtaining derivatives of the density of states begins by using Gauss' theorem to transform eq. (2) to a volume integral:

$$N(\epsilon) = \int \frac{dS}{|\nabla_k \epsilon_n|} = \int dV \nabla_k \cdot \left[\frac{\nabla_k \epsilon_n}{|\nabla_k \epsilon_n|^2} \right] \quad (15)$$

The volume integral extends over the volume of reciprocal space both inside the energy surface and inside the Brillouin zone. (Since Gauss' theorem applies only to closed surfaces, there should

be extra terms involving integrals over the Brillouin-zone surfaces for those energy surfaces which intersect the Brillouin-zone boundaries. However, these extra integrals, when taken over opposite Brillouin-zone surfaces, can be shown to exactly cancel by symmetry). The energy derivative of the volume integral in eq. (15) is again a surface integral (the difference in the volume integrals for energy E and energy $E+dE$ is an integral over the shell between the two energy surfaces):

$$N'(E) = \int_{E_n(\vec{k})=E} \frac{dS}{|\nabla_k E_n|} \nabla_k \cdot \left[\frac{\nabla_k E_n}{|\nabla_k E_n|^2} \right] \quad (16)$$

Higher derivatives of $N(E)$ can be obtained by repeating this procedure; for example,

$$N''(E) = \int_{E_n(\vec{k})=E} \frac{dS}{|\nabla_k E_n|} \nabla_k \cdot \left\{ \left[\nabla_k \cdot \left(\frac{\nabla_k E_n}{|\nabla_k E_n|^2} \right) \right] \frac{\nabla_k E_n}{|\nabla_k E_n|^2} \right\} \quad (17)$$

If one has some way of supplying the higher derivatives of $E_n(\vec{k})$, these integrals can be done by a generalization of the GR method suitable for handling integrals of functions, rather than unity, over an energy surface.

The simplest procedure is to use the GR algorithm directly, replacing the divergence term by its value at the cube center and using this value as a weight factor for the contribution of each cube. To obtain the derivative of the density of states from eq. (16) in this way, the second derivatives of $E_n(\vec{k})$ are required at the center of each cube. Although trying to obtain these from 27-point Lagrangian interpolation would be pushing this form of interpolation too far, they can probably be obtained with fair accuracy from determinantal interpolation schemes. It is also possible to obtain them, through $k \cdot p$ theory, directly from band calculations if the matrix elements of momentum have been computed.

At any rate, Fig. 5 shows the results obtained for the derivative of the density of states of the tight-binding s-band in the face-centered cubic structure. The divergence term in eq. (16) is replaced by its value at the center of the cube, as discussed above, and all derivatives of $E_n(\vec{k})$ are obtained analytically. The results are accurate except just beyond the critical point at L, where the errors become very large. Going to 77,124 cubes in the irreducible wedge (halving each cube edge) reduces the error only by a factor of 2; thus the errors in the algorithm are $O(b)$. The source of this error is easily understood: there are regions of reciprocal space for this band in which the divergence term changes very rapidly. But the variation in the integrand in eq. (16) over each cube has been neglected (automatically leading to an error of the

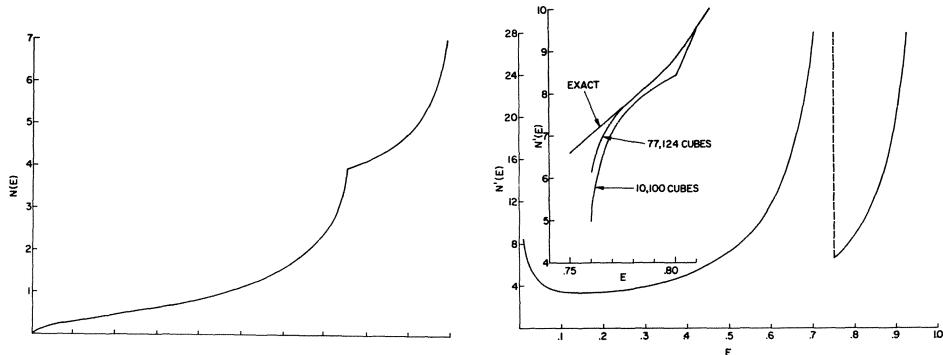


Figure 5. Derivative of $N(E)$ for tight-binding s-band in fcc structure. Algorithm reproduces the derivative except near $E=0.75$; the errors in this region are shown in the inset. Error at $E=0.76$ is 20% for 10,100 cubes, and 10% for 77,124 cubes in the irreducible wedge.

order of b times the gradient of the integrand), and this approximation is simply not good enough in those regions.

One way to improve the convergence to $O(b^2)$ is to retain both zeroth- and first-order terms in the expansion of the divergence term about the cube center, and then to use the formulas given by Gilat and Kam⁴ to handle the \vec{k} -dependent terms. However, this procedure requires the third derivatives of $E_n(\vec{k})$, commodities which may be difficult to get if the energies are being generated by an interpolation scheme.

Another possibility is suggested by the fact that the divergence appearing in eq. (16) is large whenever the Gaussian curvature is large. One could rank the cubes according to decreasing values of the integrand, and subdivide those cubes in which the integrand is largest. In this way, one would end up with very small cubes clustered around the regions where they are needed, with larger cubes in the remainder of the irreducible wedge. (The energies and derivatives could still be supplied by interpolation schemes.) Of course, this would be a time-consuming procedure if carried out for a large number of energies, but one is interested in the derivative of the density of states only at the Fermi energy.

Still another possibility, the one that is most attractive from the standpoint of interpolation schemes, is to simply compute extremely accurate values of the density of states, for a few energies near the Fermi energy, by subdividing the cubes according to curvature as above, and then to obtain the derivatives of the

density of states by numerical differentiation. None of the above procedures have been implemented as yet, and so no information is available as to which is the most reasonable to use.

Acknowledgement

I am indebted to A. R. Williams and D. E. Eastman for extensive discussions and assistance during the development of much of this work.

References

1. F. M. Mueller, J. W. Garland, M. H. Cohen, and K. H. Bennemann (to be published).
2. G. Gilat and L. J. Raubenheimer: Phys. Rev. 144, 390 (1966).
3. L. J. Raubenheimer and G. Gilat: Phy. Rev. 157, 586 (1967).
4. G. Gilat and Z. Kam: Phys. Rev. Letters 22, 715 (1969).
5. E. C. Snow: Phys. Rev. 158, 683 (1967).
6. F. M. Mueller: Phys. Rev. 153, 659 (1967).
7. L. Hodges, H. Ehrenreich, and N. D. Lang: Phys. Rev. 152, 505 (1966).
8. J. F. Janak, D. E. Eastman, and A. R. Williams: Sol. State Comm. 8, 271 (1970).
9. J. F. Janak, D. E. Eastman, and A. R. Williams: Proceedings of Electronic Density of States Symposium, NBS, Gaithersburg, Md., Nov. 1969 (to be published)
10. Calculations performed by A. R. Williams
11. J. O. Dimmock and A. J. Freeman, private communication.
12. Work performed in collaboration with A. R. Williams.
13. D. Brust, Phys. Rev. 139, A489 (1965).
14. E. O. Kane, Phys. Rev. 175, 1039 (1968).

Appendix: Placement and number of cubes for GR method

Difficulties in the GR method associated with points where the energy gradient must vanish by symmetry are avoided by placing the cube centers on an offset mesh. If $2b$ is the cube edge, the cube centers are located at the points defined by

$$\vec{k} = (2i+1)b\hat{i}_x + (2j+1)b\hat{i}_y + (2k+1)b\hat{i}_z \quad (A1)$$

where i , j , and k are integers, and \hat{i}_x , \hat{i}_y , \hat{i}_z are unit vectors along the cubic axes. The ranges of i , j , and k are fixed by the definition of the irreducible wedge, which we discuss separately for the sc, bcc, and fcc structures.

A. Simple cubic structure. Let all distances be measured in units of π/a . The irreducible wedge can then be defined as the region

$$1 \geq k_x \geq k_y \geq k_z \geq 0 \quad (A2)$$

We place cubes with centers defined by eq. (A1) in such a way that an integral number (N , say) of cubes will fit along the k_x axis. This requires

$$b = 1/2N, \quad N-1 \geq i \geq j \geq k \geq 0 \quad (A3)$$

The total number of cubes in the irreducible wedge is

$$C(N) = \frac{N(N+1)(N+2)}{6} \quad (A4)$$

Note that some of these cubes are exactly bisected by the (110) planes bounding the irreducible wedge. Since the gradient of $E_n(k)$ must lie in the (110) plane if k does, it follows that the planes of constant energy in these cubes are perpendicular to the (110) plane, and the area of the portion of a constant energy plane inside the irreducible wedge is exactly half the total area of the constant energy plane in the cube. Thus, the contributions of cubes bisected by a (110) plane should be divided by 2, cubes cut by 2 (110) planes should be divided by 6, etc. The cubic symmetry, in other words, is used to correctly treat those cubes which have portions sticking out of the irreducible wedge.

B. Body-centered cubic structure. Let all distances be measured in units of $(2\pi/a)$. The irreducible wedge is then defined as the region

$$k_x \geq k_y \geq k_z \geq 0; \quad k_x + k_y \leq 1 \quad (A5)$$

Again place the cubes with centers defined by eq. (A1), and choose b so that an integral number N of cubes will fit along the k_x -axis. Then

$$b = 1/2N; \quad N-1 \geq i \geq j \geq k \geq 0; \quad i+j \leq N-1 \quad (A6)$$

Those cubes with portions sticking out of the (110) bounding planes have their contributions divided by 2, as follows from the symmetry argument given above.

The total number of cubes is given by

$$C(N) = \begin{cases} N(N+2)(N+4)/24, & N \text{ even;} \\ (N+1)(N+2)(N+3)/24, & N \text{ odd.} \end{cases} \quad (\text{A7})$$

C. Face-centered cubic structure. Let all distances be measured in units of $(2\pi/a)$. Then the irreducible wedge can be defined as the region

$$i \geq k_x \geq k_y \geq k_z \geq 0; \quad k_x + k_y + k_z \leq 3/2 \quad (\text{A8})$$

The (111) face of the irreducible wedge (defined by the condition $k_x + k_y + k_z \leq 3/2$) introduces a complication not present in the other two structures. There will be cubes with centers inside the irreducible wedge with portions sticking out through the (111) face; but now there will also be cubes with centers just outside the (111) face with portions sticking into the irreducible wedge. These cubes have to be included so that the entire irreducible wedge is covered, and the condition defining the (111) face has to be changed slightly when translated into a condition defining allowed cube centers. The contributions of cubes cut by the (111) face can still be handled by symmetry: consider two irreducible wedges on opposite sides of the zone. Cubes with centers just inside the (111) face of one wedge will be transformed, by the (111) translation, into cubes with centers just outside the (111) face of the other wedge, and vice versa. Since the contribution of the two irreducible wedges to any quantity computed in this paper is twice the contribution of one, it follows that the contributions of cubes cut by the (111) face should be divided by 2 (cubes cut by (110) planes also have their contributions reduced, as discussed above).

Once again, choose cubes with centers defined by eq. (A1), and choose b so that an integral number N of cubes fit along the k_x -axis. Also modify the condition defining the (111) face to include cubes with centers outside this face, but with portions inside the irreducible wedge. Then the allowed values of i , j , and k are

$$b = 1/2N; \quad n-1 \geq i \geq j \geq k \geq 0; \quad 2(i+j+k) \leq 3N-1. \quad (\text{A9})$$

The total number of cubes obtained in this way is given by the following expression:

$$C(N) = \begin{cases} N(N+2)(2N+5)/24, & N \text{ even;} \\ (N+1)(4N^2+17N+9)/48, & N = 2k+1 \text{ (k odd);} \\ (N+3)(4N^2+9N-1)/48, & N = 2k+1 \text{ (k even).} \end{cases} \quad (\text{A10})$$

THE CALCULATION OF BRILLOUIN ZONE INTEGRALS BY
INTERPOLATION TECHNIQUES

R.L. Jacobs and D. Lipton

Mathematics Department, Imperial College
London, S.W.7, England

In this paper we present a quasi-analytic technique for calculating Brillouin zone integrals of the form

$$R\chi_0(\underline{q}, \omega) = N^{-1} P \sum_{\underline{k}} (f(\underline{k}) - f(\underline{k} + \underline{q})) / (\epsilon_{\underline{k}} - \epsilon_{\underline{k} + \underline{q}} - \hbar\omega) . \quad (1)$$

We shall discuss this and other techniques for evaluating these integrals with particular attention to numerical accuracy. The method can be applied to the density of states (in which case it is almost equivalent to that of Gilat and Raubenheimer¹) and to the calculation of magnetic susceptibility functions and we shall present some of these calculations later. Full details of the technique are given in other publications^{2,3}.

Our technique for calculating the Brillouin zone integral of equation (1) is a generalisation of that of Gilat and Raubenheimer¹ for calculating the density of states, i.e. the energy bands are replaced by a locally linear approximation which enables the integral to be calculated essentially analytically. In order to see the necessity for this procedure it is worthwhile examining simpler procedures in some detail.

The simplest method for evaluating the integral in equation (1) is to choose some collection of points in \underline{k} -space, to evaluate the integrand at these points and to add up these suitably weighted values. When $\omega \neq 0$ it can easily be shown that on account of the

principle parts nature of the integrand the expected variance of the estimate is infinite. It is less obvious but also true that when $\underline{q} \neq 0$ and $\omega = 0$ the expected variance of an estimate obtained by this method is also infinite. In figure (1) we present curves of $\chi_0(\underline{q}, 0)$ evaluated for a single tight-binding s-band in a face-centred cubic lattice⁴. It can easily be seen from figure (1) that two consequences of the infinite variance are that the estimated value of the integral does not approach a limit as we increase N from 4000 to 32000 and that wild fluctuations in the value occur e.g. when $N=16384$. We therefore can conclude that no simple point-sampling technique can ever lead to accurate values of the integral in equation (1). The low \underline{q} limit of $\chi_0(\underline{q}, 0)$ is just the density of states, $N(E)$, and point-sampling provides a convergent procedure for evaluating, $N(E)$, in a histogram representation.¹ There is still a random noise proportional to $N^{\frac{1}{2}}$ in $N(E)$ and this makes it impossible to calculate $N'(E)$ and $N''(E)$, which are needed for calculations of the magnetic properties of metals. For these reasons we have set up a new technique for calculating these integrals.

Our technique is based on two devices which together enable us to evaluate the integral quasi-analytically. The first is the locally linear approximation to the energy bands; the energy bands are calculated on a cubic mesh in \underline{k} -space and approximated by a linear least-mean-squares fit to the values at the corners of the cube. The second is a transformation of equation (1) to a form analogous to the surface integral representation of the density of states. The volume integral in equation (1) becomes the product of a surface integral and an integral over an energy parameter, i.e.

$$R\chi_0(\underline{q}, \omega) = (NZ(\underline{q}))^{-1} P \sum_{\text{star of } \underline{q}} \int_{\Delta_{\min}}^{\Delta_{\max}} R(g_{\underline{q}}, \Delta) \frac{d\Delta}{\Delta - h\omega} \quad (2)$$

where

$$R(g_{\underline{q}}, \Delta) = Z(\underline{k}_c) \int'' (f(\underline{k} + g_{\underline{q}}^{-1}\underline{q}) - f(\underline{k})) |\text{grad}_{\underline{k}} \Delta_{\underline{k}, \underline{q}}|^{-1} dS_{\underline{k}}, \quad (3)$$

$$\Delta_{\underline{k}, \underline{q}} = \epsilon_{\underline{k} + \underline{q}} - \epsilon_{\underline{k}} \quad (4)$$

and $\int'' dS$ denotes an integral over the surface $\Delta_{\underline{k}, \underline{q}} = \Delta$ contained in a single cube, \underline{k}_c is the centre of the cube, Δ_{\max} and Δ_{\min} are the maximum and minimum values of $\Delta_{\underline{k}, \underline{q}}$ restricted to the cube, Σ'' refers to a sum over all cubes contained in the irreducible part of

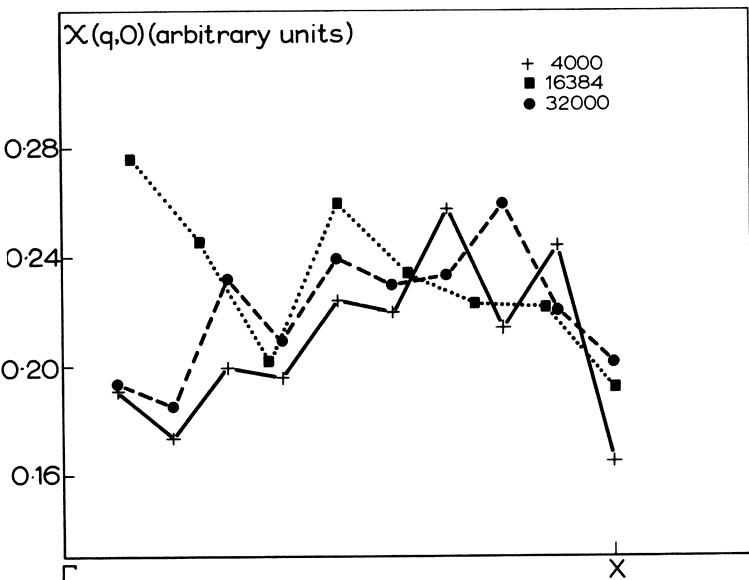


Figure 1: $\chi_{\text{q},0}$ in a one-band model calculated by a point-sampling technique.

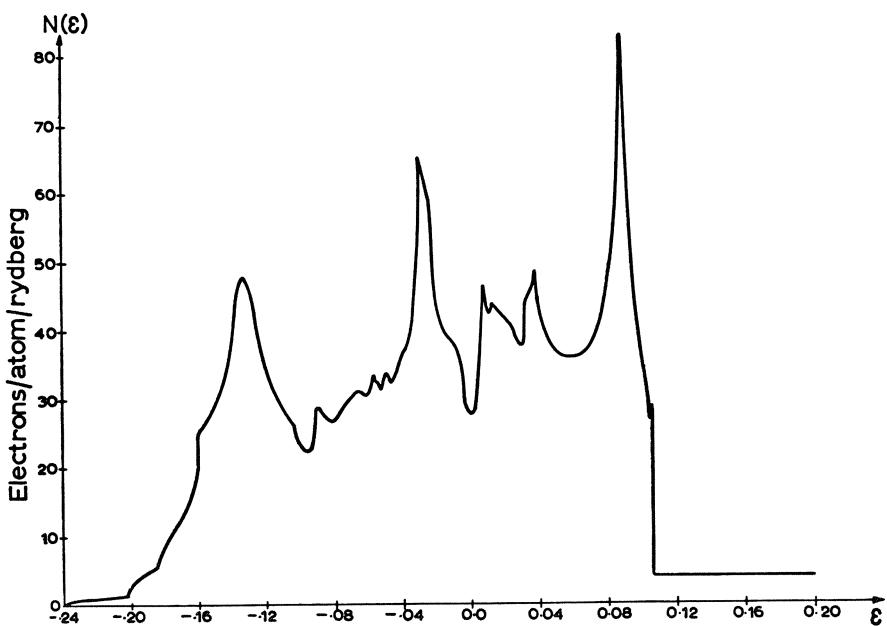


Figure 2: The density of states, $N(E)$, for a representative model Hamiltonian transition metal band structure.

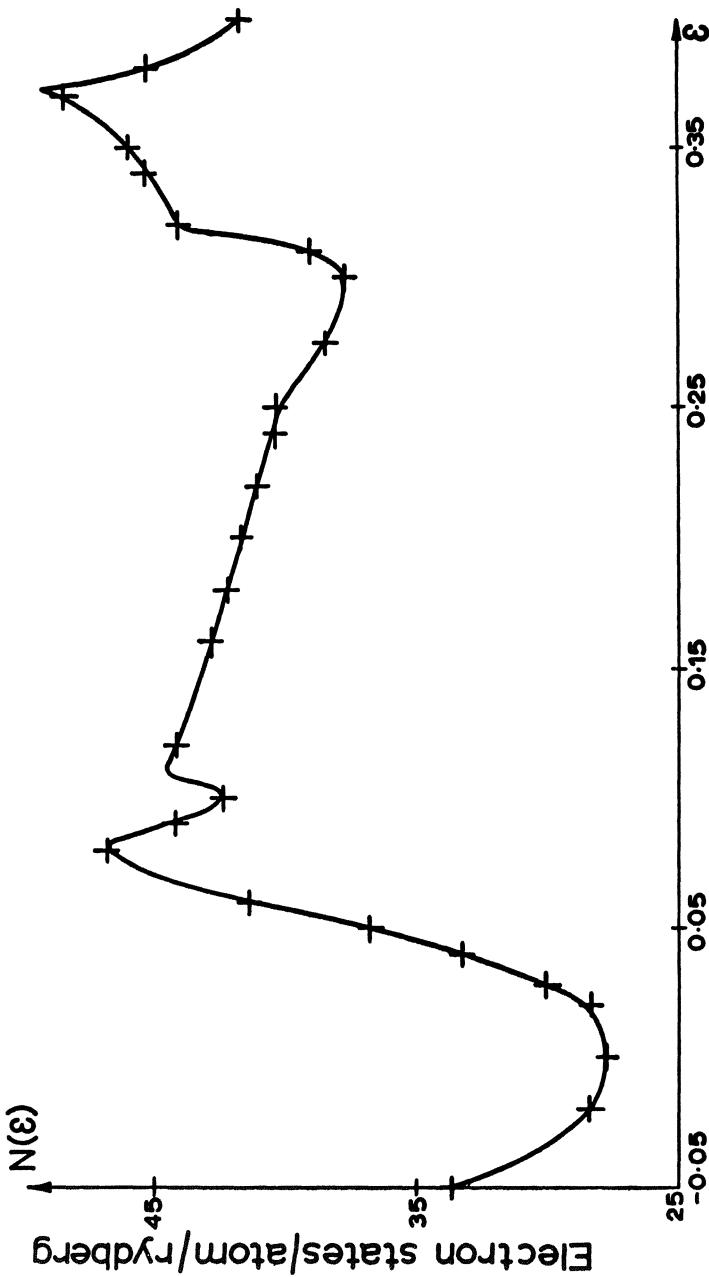


Figure 3: A blow-up of $N(E)$ in a narrow energy region. (N.B. Lower scale is rydbergs $\times 10.$)

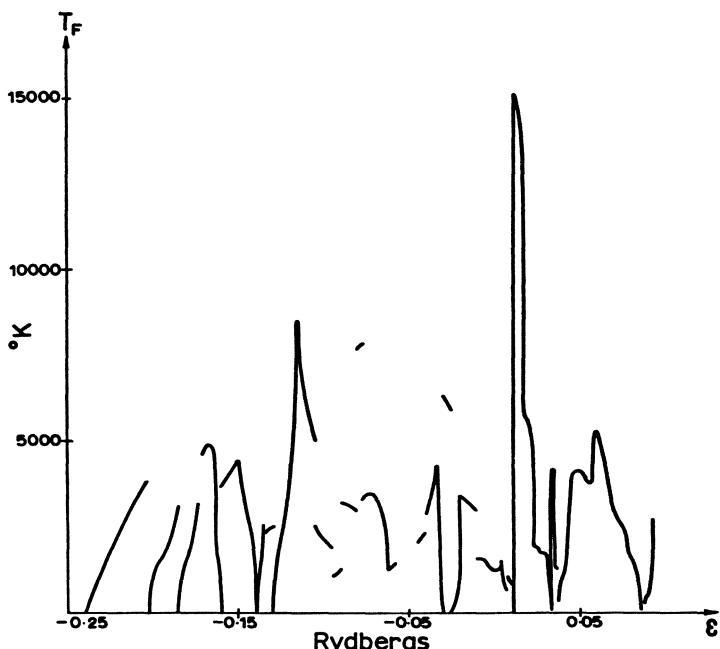


Figure 4: The local degeneracy temperature, $T_F(E)$.

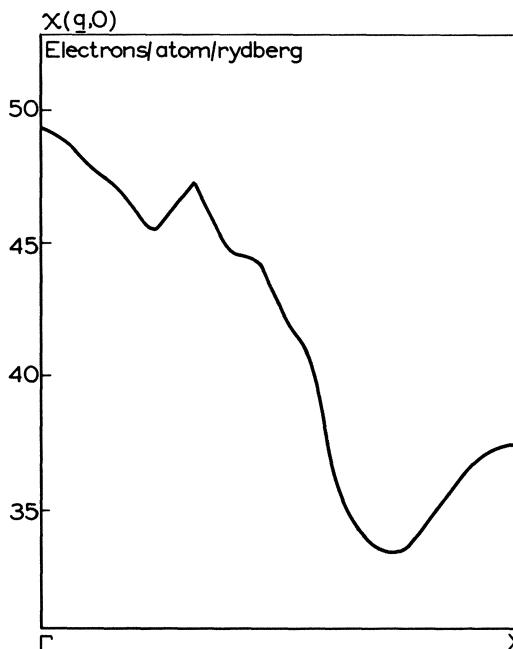


Figure 5: The sum of the intraband contributions to $\chi_{\infty}(q,0)$ with constant matrix elements for a palladium-like transition metal.

the zone, $g_{\underline{q}}$ in the rotation operator in the star of \underline{q} and $Z(\underline{k}_c)$, \underline{q} and $Z(\underline{q})$ are the number of vectors in the star of \underline{k}_c and \underline{q} respectively.

The function $R(g_{\underline{q}}, \Delta)$ in equation (3) is evaluated for given Δ by splitting it into two terms each depending on one Fermi factor. The surfaces in the locally linear approximation are planes in each cube and $|\text{grad}_{\underline{k}} \Delta_{\underline{k}, \underline{q}}|$ is constant on each plane. Two cases then arise: (i) the Fermi factor is constant over a given plane, (ii) the Fermi factor is 1 over a portion of the plane and zero over the remainder. In the first case the area function $R(g_{\underline{q}}, \Delta)$ if non-zero is evaluated using formulae given by Gilat and Raubenheimer¹; in the second case $R(g_{\underline{q}}, \Delta)$ is evaluated by trapezoidal integration of the portion of the area on which the Fermi factor is unity. This is the hardest part of the program and it is at this point that we would insert matrix elements if we were using them. $R(g_{\underline{q}}, \Delta)$ is evaluated for several values of Δ within each cube and approximated by a piecewise linear function between these values. The known analytic approximation to $R(g_{\underline{q}}, \Delta)$ is then used in equation (2), to evaluate $\text{Re}x_o(\underline{q}, \omega)$.

We have thus set up a convergent numerical procedure for calculating $\text{Re}x_o(\underline{q}, \omega)$. The computer program based on this procedure is quite difficult. The difficulty arises mainly from the fact that while there are only five distinct ways in which a plane can cut a cube, there are 90 ways in which the resulting cross-sections can be truncated by a Fermi factor. The program has been tested by calculating $x_o(\underline{q}, 0)$ for a single tight-binding s-band in a face-centred cubic lattice⁴. The results which are presented in reference (2) show that the change in the estimate of $x_o(\underline{q}, 0)$ in going from a grid containing 916 points is at worst 5%. We may conclude that the error using the finer grid is only ~1% and is systematic rather than random. The time taken is ~10 secs. per point on the CED Run compiler system.

We have used our program to calculate other quantities. In figures (2) and (3) we present the low \underline{q} limit of $x_o(\underline{q}, 0)$ for a parametrised transition metal band structure. The smoothness and high resolution of these curves should be noted. In figure (4) we present $T_F(E)$, the local degeneracy temperature, for the same band structure which depends on $N(E)$, $N'(E)$ and $N''(E)$ as follows:

$$T_F(E) = (6/\pi^2 k_B^2 a(E))^{\frac{1}{2}} \quad (5)$$

$$a(E) = N''(E)/N(E) - (N'(E)/N(E))^2 \quad (6)$$

This curve was obtained by differentiating the density of states in figure (2) numerically. The error is of the order of 10% in favourable regions. Finally in figure (5) we present the intraband contribution to $\chi_0(q, \omega)$ for a palladium-like transition metal. The matrix elements are taken to be constant and q is in the (100) direction. The large peak should be noted. A similar peak exists in the (110) direction but is even more prominent. This may account for the anomalies observed in the phonon spectrum of palladium by Brockhouse⁵.

References

- ¹ Gilat, G. and Raubenheimer, L.J., Phys. Rev., 144, 390 (1966).
- ² Lipton, D. and Jacobs, R.L., J.Phys.C., [2], 3 June (1970).
- ³ Lipton, D. and Jacobs, R.L., J.Phys.C., to be published.
- ⁴ Callaway, J., Phys. Rev., 170. 576 (1968).
- ⁵ Miller, A.P. and Brockhouse, B.N., Phys. Rev. Letts., 20, 798 (1968).

Acknowledgments

The authors would like to thank Professor E. P. Wohlfarth for suggesting this problem. One of us (R.L.J.) would like to thank the Mathematics Department of Imperial College, the Royal Society and the I.B.M. Corporation for financial aid which made it possible for him to attend this conference.

COMPUTATIONAL METHOD FOR GENERALIZED SUSCEPTIBILITY

Joshua B. Diamond

Physics Department

University of Pennsylvania, Philadelphia, Pennsylvania

DESCRIPTION OF METHOD

A computer program has been developed to calculate the real part of the non-interacting generalized susceptibility function at zero temperature, as given by equation (1) below.

$$\text{Re } \chi_0(\vec{q}, \omega) \propto \sum_{nn'} \int_{\text{BZ}} d^3k \frac{|M_{kn, k+\vec{q}+\vec{K}n'}|^2 (f(\epsilon_{kn}) - f(\epsilon_{k+\vec{q}+\vec{K}n'}))}{\epsilon_{k+\vec{q}+\vec{K}n'} - \epsilon_{kn} - \omega} \quad (1)$$

$$f(\epsilon_{kn}) = \begin{cases} 1 & \epsilon_{kn} < \epsilon_F \\ 0 & \epsilon_{kn} > \epsilon_F \end{cases}$$

In this equation ϵ_{kn} is the energy of a Bloch state with wave vector k and band index n , the $T = 0$ Fermi function $f(\epsilon_{kn})$ is defined as usual, and $M_{kn, k+qn'}$ is a matrix element assumed slowly varying with k . The principal value integration runs over the first Brillouin zone, the sums over n and n' go over all bands, and \vec{K} is the reciprocal lattice vector which reduces $k+q$ to the first Brillouin zone.

In this method the band energies and gradients are determined on a discrete mesh of points in k -space. Each mesh point is sur-

rounded by a cube such that all the cubes together fill up the first Brillouin zone, with each mesh point lying at the center of its cube. In practice, each cube must be weighted appropriately to give the correct effective volume of integration. The integration over the zone is performed on one cube at a time. The energies are approximated linearly through the volume of the cube using the known values of the energy and gradient at the cube center, or by some alternative method of linear interpolation.¹ The matrix element is assumed to be constant within each cube. We then arrive at equation (2), which gives the approximation to the susceptibility which is actually computed.

$$\text{Re } \chi_0(\vec{q}, \omega) \propto \sum_{n \neq n'} \sum_c |M_{\vec{k}_c n, \vec{k}_c + \vec{q} + \vec{k}_{n'}}|^2 P \int_{\text{cube}} d^3 k \frac{f(\tilde{\epsilon}_{\vec{k}n}^{(c)}) - f(\tilde{\epsilon}_{\vec{k}+q+\vec{k}n'}^{(c)})}{\tilde{\epsilon}_{\vec{k}+q+\vec{k}n'}^{(c)} - \tilde{\epsilon}_{\vec{k}n}^{(c)} - \omega}$$

$$\tilde{\epsilon}_{\vec{k}n}^{(c)} = \epsilon_{\vec{k}n} + \nabla_{\vec{k}} \epsilon_{\vec{k}n} \cdot \vec{k}$$
(2)

$$\vec{k} = \vec{k} - \vec{k}_c$$

$\tilde{\epsilon}_{\vec{k}n}^{(c)}$ is the linearly interpolated energy in cube c , \vec{k}_c is the coordinate of the cube center, and the sum over c runs over all cubes.

Let us look at the integral over a single cube. If the Fermi factors are constant inside the cube, then either the integral is zero or the effective volume of integration is the whole cube. Assume now that the first Fermi factor $f(\tilde{\epsilon}_{\vec{k}n}^{(c)})$ is not constant over the cube. The surfaces of constant energy within the cube are planes. The plane $\tilde{\epsilon}_{\vec{k}n}^{(c)} = \epsilon_F$ (which is the local Fermi surface) is assumed to intersect the cube. This plane divides the cube into two regions. In one region the Fermi function is zero; in the other it is equal to unity. The effective volume of integration is thus a truncated portion of the cube, shown schematically as the shaded part of the illustration below.



¹ With some investment in additional analytic integrations, the program can be extended to include a linear variation of the matrix element factor within each cube.

The situation is similar for the second Fermi factor $f(\tilde{\epsilon}_{\vec{k}+\vec{q}+\vec{K}}^{(c)})$.

When one looks at the integrals explicitly, one finds that the contribution from an arbitrary cube can be written as a sum of integrals of the following form:

$$\int_{x_0}^{x_1} dx \int_{y_0}^{mx+n} dy \int_{z_0}^{px+qy+r} dz \frac{1}{ax+by+cz+d}$$

These are easily evaluated. The chief complexity in the method comes in determining correctly the contributions which arise from the various possible ways that a plane surface can intersect a cube. Following the analysis of Gilat and Raubenheimer² these cases have all been treated and the necessary integrals have been worked out analytically and stored in the computer program. The program selects the appropriate contribution for each cube and adds them up to get the total susceptibility. Once having made a linear interpolation for the energies and having assumed locally constant matrix elements, the integration is in principle performed exactly, limited only by the finite precision of the computer arithmetic.

The principal value nature of the integral is handled analytically, and does not introduce spurious noise, which might be expected from other numerical integration methods when dealing with vanishing energy denominators.

In this method the actual Fermi surface is replaced by a segment of a plane, within each small cube. So, in order to obtain accurate results one must at least use a mesh fine enough so that the linear dimensions of the important features of the Fermi surface are several times larger than the length of a side of a mesh cube.

APPLICATIONS OF METHOD

Several calculations using this method, all done for zero frequency, will now be presented. The first two employ idealized band structures and the numerical results can be compared with exact analytic calculations.

² G. Gilat and L. J. Raubenheimer, Phys. Rev. 144, 390 (1966).

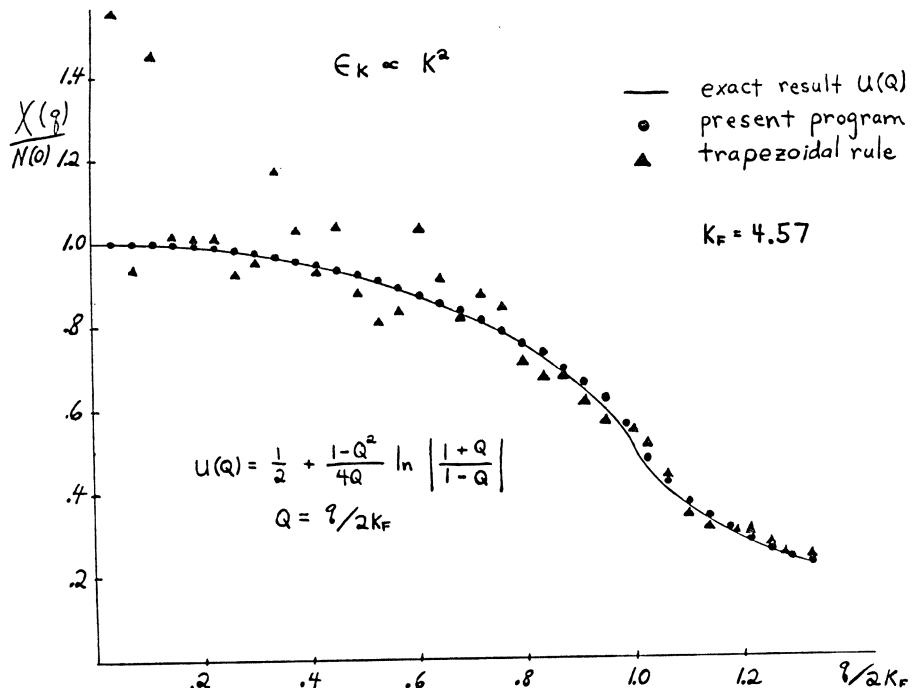


Figure 1. Free electron susceptibility

In figure 1. the susceptibility for free electrons is shown calculated in several ways. The solid line gives the exact result, obtainable by direct integration, and is just the familiar Lindhard function $U(Q)$, defined in the figure. The round points give the results of the linear interpolation program described in the first part of this paper. In units of the distance between adjacent mesh points (or the side of an integration cube) the Fermi wave vector K_F is approximately 4.5, so that this is a rather coarse mesh. The program is in excellent agreement with the analytic result, with a maximum percentage deviation of 3%, occurring near the logarithmic singularity in the derivative at $q = 2K_F$; the average absolute deviation is 1.2% for the curve.

By contrast the results of applying the three-dimensional version of the trapezoidal rule are plotted in figure 1. with triangular points. With this integration scheme each cube in K -space contributes an amount equal to the integrand evaluated at the cube center, weighted by the volume of the cube. The mesh used is the same as before. Here, the noise is quite large, especially for small q , while with the linear interpolation method

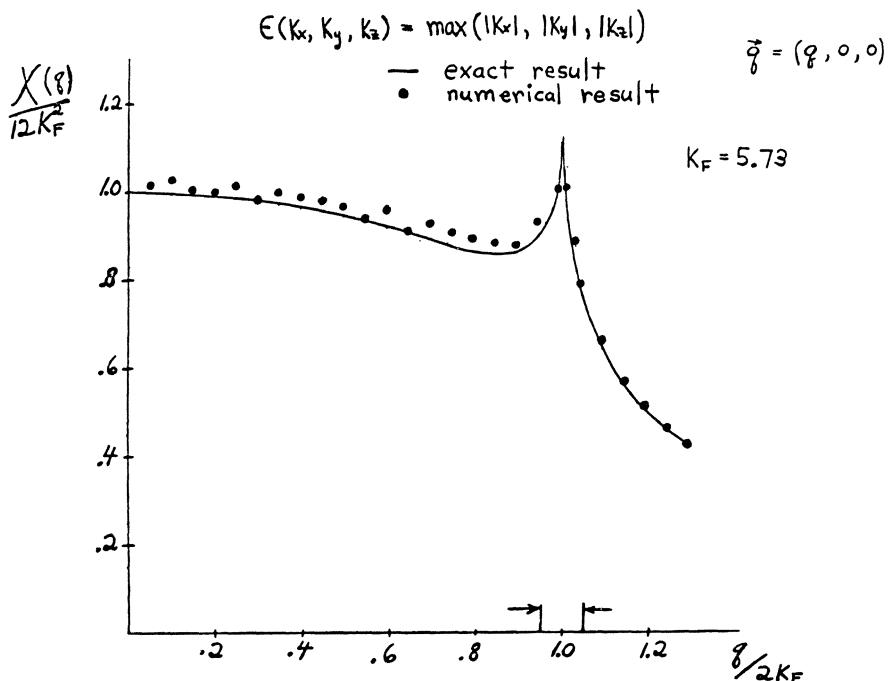


Figure 2. Cubic Fermi surface susceptibility

the results are highly accurate at the smallest q 's considered. The difficulties with the trapezoidal rule scheme on a coarse mesh have been pointed out before³ - I use it here only to indicate the potential for noise in these calculations. With a finer mesh the linear interpolation method yields results which approach even more closely the analytic Lindhard Function.

The nesting of Fermi surface areas is significant in determining the structure of the wave vector dependent susceptibility. If two parallel planar pieces of Fermi surface exist, with appropriate velocities, a logarithmic singularity in the susceptibility appears at the nesting wave vector. This singularity can be derived from a locally linear expansion of the energy; one expects therefore that this linear interpolation program will be sensitive to the nesting phenomenon, as indeed it is.

The rather artificial band structure defined by the equation
 $E \vec{k} = \max(|k_x|, |k_y|, |k_z|)$

³ W. E. Evenson and S. H. Liu, Phys. Rev. 178, 783 (1969).

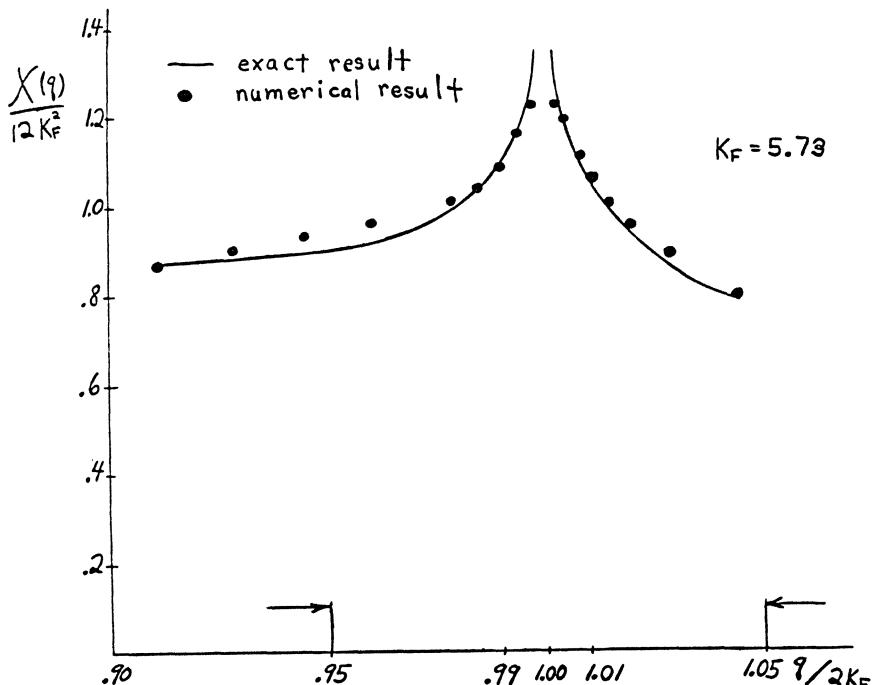


Figure 3. Nesting region in cubic Fermi surface susceptibility

yields a Fermi surface with the shape of a cube. K_F is defined to be $\frac{1}{2}$ the length of a side of this cube. The susceptibility is easily calculated analytically for q along a 100 direction and, as expected, exhibits a log singularity at $q = 2K_F$. The solid line in figure 2. is the analytic result for this susceptibility. For the numerical calculation a mesh was used such that $K_F = 5.73$ in units of the distance between mesh points. The points plotted in figure 2. are the results of the linear interpolation program and show good agreement with the exact result. The small deviations for $q < 2K_F$ are due to the artificial discontinuities in the energy gradients across 110 planes. These are necessary to yield a cubic Fermi surface; however, the program cannot simulate them accurately. For q near to and greater than the nesting value $2K_F$ the regions containing the discontinuities give a lesser contribution to the susceptibility, and the program comes into excellent agreement with the analytic result. The region around the singularity, indicated in figure 2. by the two arrows, is plotted on an expanded horizontal scale in figure 3.

For the cubic Fermi surface results shown, the maximum percentage deviation from the analytic result is $4\frac{1}{2}\%$. The average

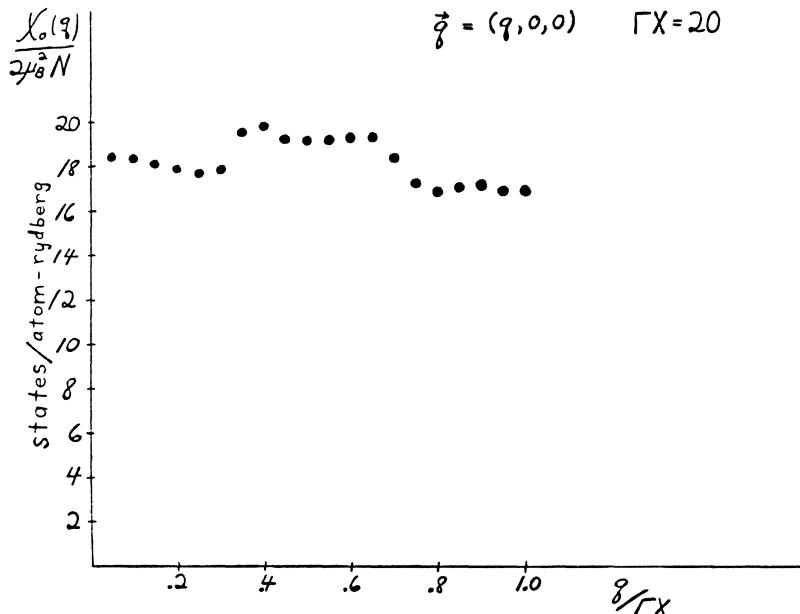


Figure 4. Tight binding d-band susceptibility

absolute deviation is $2\frac{1}{2}\%$. This computation has been repeated for a second mesh with similar results.

As a final example, a computation using the linear interpolation program in a more realistic setting will be discussed. The susceptibility of a single tight binding d band in a fcc lattice is shown in Figure 4.; q is taken along a 100 direction. The integration is effectively over the fcc Brillouin zone. In units of the distance between adjacent mesh points the distance from Γ to the X point in the zone is equal to 20; i.e. there are twenty cubes between Γ and X . The matrix element in the susceptibility formula has been taken equal to 1. For readers familiar with palladium, this tight binding band is approximately fit to the heavy hole d band in Pd.

Details and possible interpretation of this curve will not be discussed here, except to point out that as $q \rightarrow 0$, the curve is smooth, and that it in fact extrapolates to within 2% of the band density of states at the Fermi energy, calculated on a much finer mesh (effectively 56 cubes from Γ to X) by a histogram technique. For the susceptibility, using 100 symmetry, the average integration time per q point on an IBM 360 Model 75 was about 10 seconds.

It seems probable that this numerical integration method will prove increasingly useful in quantitative calculations of electronic properties in metals. The author gratefully acknowledges many helpful conversations with Drs. O. K. Andersen, W. E. Evenson, and J. R. Schrieffer, at the University of Pennsylvania. Further details of the method and a FORTRAN listing of the program are available from the author.

This work was supported by the National Science Foundation and the Advanced Research Projects Agency.

LCAO INTERPOLATION METHOD WITH NONORTHOGONAL ORBITALS

L. F. Mattheiss

Bell Telephone Laboratories, Incorporated

Murray Hill, New Jersey

Most applications of the Slater-Koster tight-binding interpolation method¹ utilize basis functions formed from orthogonalized atomic orbitals. We shall refer to this scheme as the LCOAO method or the linear-combination-of-orthogonalized-atomic-orbitals method. In such applications, the independent tight-binding energy integrals that occur in the Hamiltonian matrix are treated as disposable parameters which are chosen to fit the results of accurate band calculations at symmetry points in the Brillouin zone.

In fitting the band structures of transition-metal compounds, the author has found that it is advantageous to apply this tight-binding interpolation method using Bloch functions formed from nonorthogonal atomic-like orbitals. We refer to this approach as the LCAO method. It is anticipated that this method will provide a useful basis in any system where directional bonding effects are important. The LCAO method has several advantages over the LCOAO approach in fitting the band structures of transition-metal compounds.

First, the LCAO method treats crystal field effects in a very simple and physical way, one which is analogous to the treatment of crystal field splittings in an isolated transition-metal complex via the molecular-orbital method.² In the molecular orbital and LCAO methods, crystal field splittings originate from overlap and covalency between the d orbitals of the transition-metal ion and the s and p orbitals of the neighboring ligands. We shall show later that the LCOAO method introduces these crystal field effects in a much more artificial way.

The second advantage of the LCAO method is that it leads to more realistic tight-binding parameters. To see how this occurs, let us consider a typical transition-metal compound where the metal-ligand separation is small compared to the metal-metal or ligand-ligand distances. Large wave function overlap will occur between the metal and ligand orbitals. As a result, the orthogonalized transition-metal d orbitals will contain a significant ligand s and p admixture. A parameter that corresponds to a tight-binding d-d interaction between orthogonalized d orbitals will therefore contain appreciable ligand s-d and p-d components and these can often dominate the direct d-d contribution, changing both the sign and magnitude of the final parameter.

In applying the LCAO method,³ the independent tight-binding integrals that occur in the Hamiltonian as well as the overlap matrices are treated as disposable parameters in the fitting procedure. As we have mentioned, the largest overlap effects occur between orbitals localized on the transition-metal ion and its neighboring ligands so that metal-metal and ligand-ligand overlap effects can be neglected.

Typical interactions between transition-metal ions and their neighboring ligands are illustrated in Fig. 1. To the left, we have a σ -bond formed by a ligand s orbital and the transition-metal eg orbital of $3z^2-r^2$ symmetry. The Hamiltonian integral involving these orbitals is denoted by ($s\sigma$) and there is a

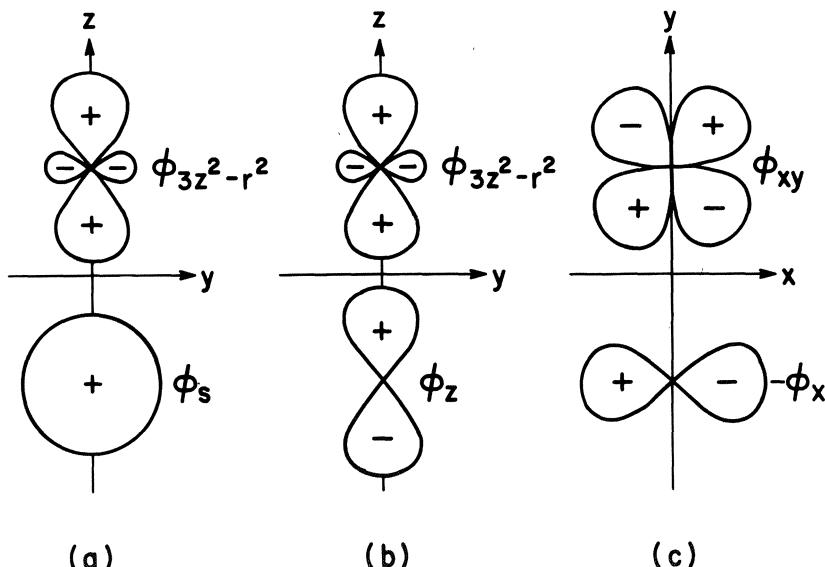


Figure 1. Schematic representation of metal-ligand bonds.

corresponding overlap integral. In the center we again show a σ -bond, but this time it is formed by a ligand p and metal $3z^2-r^2$ orbital. The energy and overlap integrals resulting from these orbitals represent a second set of interpolation parameters. Finally, to the right we have a π -bond formed from a ligand p and a transition-metal t_{2g} orbital. The energy and overlap integrals involving these orbitals represent the final set of metal-ligand interaction parameters.

To illustrate the LCAO method, we consider its application to the APW band structure of rhenium trioxide.⁴ The ReO_3 unit cell is shown in Fig. 2. The rhenium atom is darkly shaded and located at the center of the cubic unit cell. It is surrounded by six oxygen atoms, each of which are located at the face centers so that there are a total of three oxygen atoms per unit cell. The simple cubic Brillouin zone is shown to the right, with the standard notation identifying symmetry points and lines.

Fig. 3 contains the ReO_3 band profiles along symmetry lines of the Brillouin zone. The solid lines correspond to the LCAO results, the open circles represent the APW results used in the fitting procedure. The overall fit is accurate to better than 0.02 Ry, and in most cases the agreement is better than a few thousandths of a rydberg.

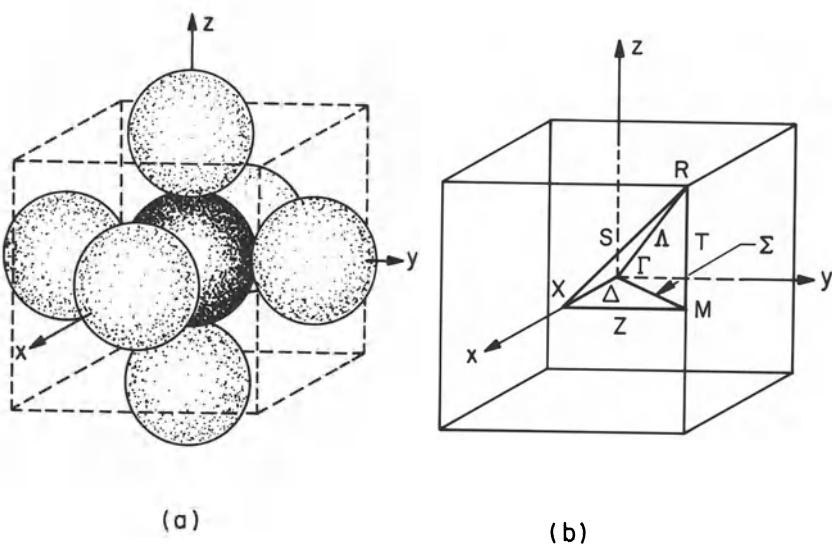


Figure 2. ReO_3 unit cell and Brillouin zone.

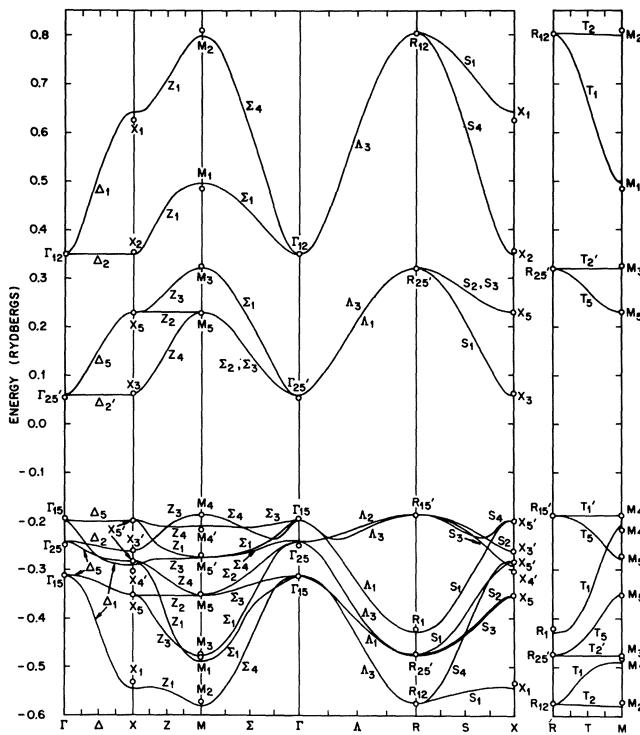


Figure 3. LCAO energy bands for ReO_3 .

On this energy scale, the oxygen 2s bands (which are not shown) are located at about -1.4 Ry. The oxygen 2p bands are shown in the lower portion of the figure and they extend from about -0.2 to -0.6 Ry. These are separated by a small energy gap from the rhenium t_{2g} bands. Finally, near the top of the figure, we have the rhenium e_g bands which are completely separated in energy from the t_{2g} bands by the crystal field effects. In fitting the ReO_3 band structure, we have used LCAO basis functions which correspond to oxygen 2s and 2p orbitals at each of the three oxygen sites in the unit cell and five rhenium 5d orbitals. This leads to a 17×17 secular equation. By including nearest-neighbor d-d and p-p interactions, we end up with a total of 23 LCAO parameters.

There is a noticeable asymmetry in the widths of the rhenium 5d and oxygen 2p bands. A detailed analysis of these results shows that the 5d and 2p bands each have a natural width of about 0.1 Ry which is due to d-d and p-p interactions, respectively. The remaining width is due to overlap and covalency effects between the rhenium 5d and oxygen 2s and 2p orbitals. These

interactions increase the 5d bandwidth to about 0.75 Ry while the oxygen 2p bandwidth is increased to only 0.4 Ry. In order to introduce this asymmetry into the LCOAO method, it is necessary to rigidly shift the e_g bands relative to the t_{2g} bands by about 0.3 Ry.⁴ This shift is contained in a parameter that is equivalent to a one-center 5d integral in the tight-binding approximation. Such a shift is required in the LCOAO method because the rhenium-oxygen interactions must preserve the center-of-gravity of the noninteracting rhenium 5d and oxygen 2s and 2p bands.

This asymmetry arises in a far more natural way in the LCAO method. In this approach, the e_g and t_{2g} orbitals have about the same energy and the asymmetry between the conduction and valence bandwidths results from cancellation between the rhenium-oxygen overlap and covalency contributions to the bandwidth. In the antibonding rhenium 5d bands, the phase of the overlap and covalency contributions is such that the two contributions add. In the oxygen 2s and 2p bands, the two contributions enter with opposite signs and cancellation occurs.

To illustrate this fundamental difference between the LCOAO and LCAO methods, we consider the results shown in Fig. 4. In Fig. 4(a), we have the APW results for ReO_3 plotted along the [100] direction. In (b), we have the LCOAO results, and in (c), the LCAO bands. Both interpolation methods yield an accurate representation of the APW results. However, we regard the success of the LCOAO method as fortuitous, resulting from the relatively simple nature of the ReO_3 band structure. More recent attempts to apply the LCOAO method to the rutile structure have been far less successful.

Notice that we have not included the rhenium 6s-6p conduction bands in the interpolation calculation. The bands connecting Γ_1 and X_4' correspond to the lower portions of these bands in the APW results shown in (a), and these bands are omitted in the interpolation results shown to the right.

Now, let us consider the results shown in the lower portion of the figure. The results shown in (d) were obtained by an APW calculation in which the potential within the oxygen APW spheres was set equal to zero. Roughly speaking, these results correspond to a band calculation for simple cubic rhenium with an enormous lattice constant. Notice that several changes have taken place. First, the d bands are quite flat, as one would expect from the large rhenium-rhenium separation. Second, the energy of the e_g bands is lowered so that they actually lie below the t_{2g} bands. And finally, the energy of the rhenium 6s-6p conduction bands has dropped by about 0.7 Ry so that the bottom of this band (Γ_1) is slightly below the d bands. This is a common feature of transition-metal band structures.

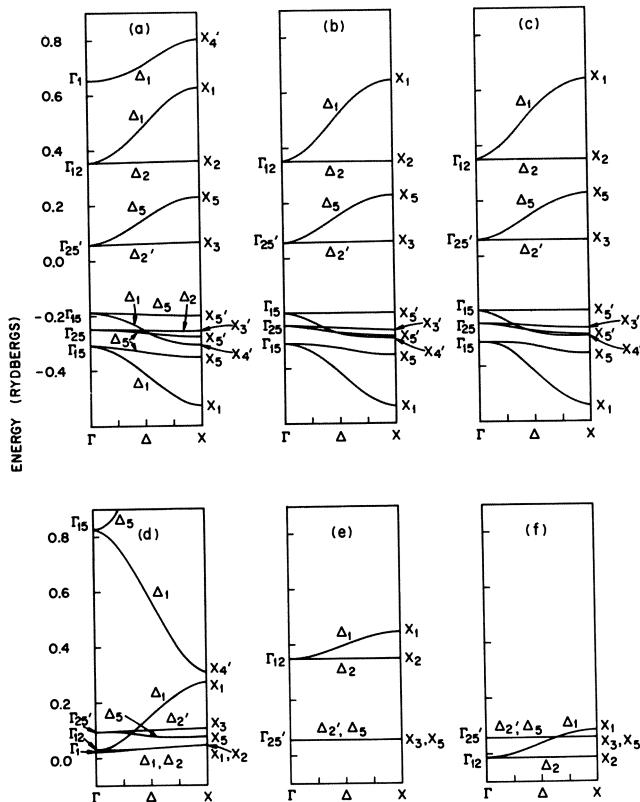


Figure 4. APW, LCOAO, and LCAO results for ReO_3 along [100].

The results shown to the right represent an attempt to simulate these APW results in the LCOAO and LCAO calculations, respectively. In obtaining the LCOAO results shown in (e), all rhenium-oxygen covalency parameters were set equal to zero. Notice that the crystal field splitting between the e_g and t_{2g} manifolds remains since it is represented by a one-center d integral in this scheme. To the right in (f), we have the LCAO results that are obtained by setting the rhenium-oxygen covalency and overlap parameters equal to zero. These results closely simulate the APW results shown to the left. The small differences in the Δ_1 and Δ_5 bands are easily attributed to the s-d and p-d hybridization effects, which are automatically included in the APW calculation but not in the LCAO method. From the APW results shown at the left, it is quite clear that metal-ligand overlap and covalency is responsible not only for crystal field effects in the d bands, but that analogous effects are responsible for raising the energy of the transition-metal s-p bands above that of the d band states.

1. J. C. Slater and G. F. Koster: Simplified LCAO Method for the Periodic Potential Problem, Phys. Rev. 94, 1498 (1954).
2. J. Owen and J. H. M. Thornley: Covalent Bonding and Magnetic Properties of Transition Metal Ions, Reports on Progress in Physics, edited by A. C. Stickland (Stonebridge Press, Bristol, England, 1966).
3. L. F. Mattheiss: Crystal Field Effects in the Tight-Binding Approximation: ReO_3 and Perovskite Structures, (to be published).
4. L. F. Mattheiss: Band Structure and Fermi Surface of ReO_3 , Phys. Rev. 181, 987 (1969).

INTERPOLATION IN k -SPACE WITH FUNCTIONS OF ARBITRARY SMOOTHNESS

Donn G. Shankland

Air Force Institute of Technology

Wright-Patterson AFB, Ohio

One can generally make first-principles calculations of crystalline properties only at a very few points of high symmetry in the Brillouin zone. It is then extremely important to be able to interpolate the relevant functions throughout the zone with smooth, well behaved functions possessing the appropriate symmetry of the crystal zone. However, a straightforward attempt to fit the computed data with some of a set of symmetry-adapted functions may well lead to unacceptably wild results. This is a special case of a very common problem of interpolation, where periodic or other boundary conditions, and/or the dimensionality of the domain, combine to render the usually obvious basis set of functions unsuitable for fitting to a limited number of data points. For example, a simple one-dimensional set of data is shown in Fig. 1, together with Fig. 2, the interpolating function obtained by fitting the data with the first 20 Fourier components. Clearly, the straightforward approach is inadequate. It is the purpose of this paper to explain how smooth interpolating functions can be: (a) constructed from a basis set possessing the proper symmetry properties, and (b) fitted to a set of data with little more numerical work than is involved in the unsatisfactory approach shown. The technique can be adapted to a wide variety of problems, lattices, dimensions, and to various asymptotic requirements. Different smoothness requirements can be imposed, and their effects on the interpolating function readily ascertained. The method has been extensively tested and has been found to yield superior results in all cases. The original development of this technique was done at the Boeing Company, and the details are available from them.¹ Here we will first discuss the usual approach, to find out why it gives bad results on occasion, and then reformulate the

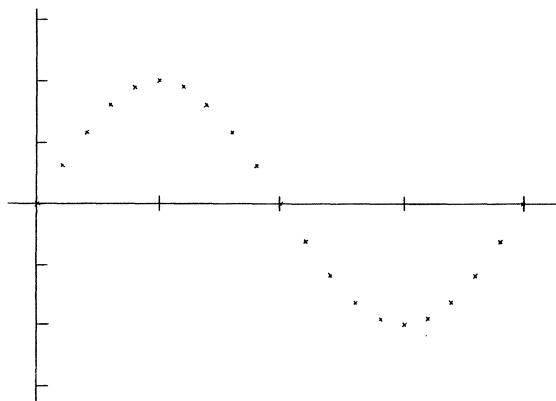


Fig. 1. Raw Data-20 Points

problem so as to obtain the improved results possible. The equations involved will then be derived to illustrate the numerical simplicity of the technique.

In the usual approach we have a set of N data points to be fit with some of a set of basis functions which possess the appropriate symmetry properties of, e.g., the Brillouin zone. The basis set is, theoretically, complete, and possesses far more than N functions. So the basic interpolation problem is under-determined. However, one wishes the interpolating function to be smooth in some sense, so he generally selects N functions which possess the longest periods from the set in order to avoid rapid wiggles in the resulting curves. This choice then renders the problem determinate, and the remainder of the problem consists in

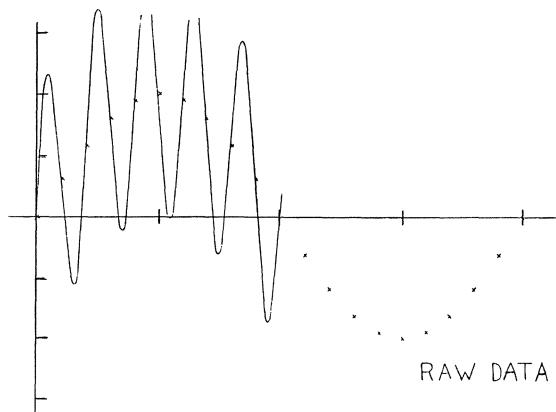


Fig. 2. Interpolant Using 20 Terms

assuming that some linear combination of these chosen functions will pass through the data points, thereby determining the coefficients of the linear combination by solving a system of N linear equations in N unknowns. In many cases the results are acceptable. However, when they are not, as seems to happen precisely when the results are needed most, one usually resorts to some compromise in order to improve the smoothness. Three such compromises are: truncating the linear combination obtained, so as to reduce the effect of high-frequency wiggles; starting the problem over with a smaller number of functions than the number of data points, and requiring only a least-squares fit to the data; or abandoning the periodicity requirements and fitting the data with a polynomial of a few terms. All of these compromises throw away some information in the data computed, and none of them are necessary.

To see just what went wrong, we first need some quantitative measure of the roughness of the interpolating function. I would submit that a simple and satisfactory definition of roughness might be the integrated square of the deviation of the function from its mean, plus the integrated square of the derivative of the function, plus perhaps the integrated square of the second derivative, and even including higher terms of the same sort, if desired. The various derivative contributions need not be equally weighted, at your pleasure. The result is a roughness functional R given by

$$R = \frac{1}{2} \int_0^{2\pi} dx \{ c_0 f^2(x) + c_1 f'^2(x) + c_2 f''^2(x) + \dots \} \quad (1)$$

When the linear combination of N functions passing through the data points is substituted in this functional, one obtains a number, and that is the end of it. However, if one had assumed a linear combination of a larger number of basis functions than there were data points, one would have, consistent with the requirement that the data points be met, some freedom to adjust the linear combination so as to reduce the roughness R , where

$$f(x) = \sum_{r=-M}^M f_r e^{irx} \quad (2)$$

$$R = \frac{\pi}{2} \sum_{-M}^M |f_r|^2 (c_0 + c_1 r^2 + c_2 r^4 + \dots) \equiv \sum_{-M}^M |f_r|^2 c(r^2) \quad (3)$$

This, then, is the key to the interpolation problem. After suitably defining roughness, one chooses a set of basis functions larger than the number of data points to be fit, and requires that the roughness be minimized, subject to the constraints that the interpolating function pass through the given data. This is a determinate problem regardless of how large a set is used, and one can even use the entire complete set of functions with the specified periodicity, if one has analytic ways of summing certain series. If not, one can choose a set sufficiently larger than the number of data points to produce significant improvements in smoothness, and yet small enough that the interpolating function can be reconstructed in the computer by directly summing the terms. Dramatic improvements in smoothness are usually obtained with just a few extra terms.

The interpolation problem as now formulated is the minimization of the roughness functional, subject to a set of linear constraints, and so is best handled through the method of Lagrangian multipliers. One forms a new roughness functional R^* by adding to R the constraints, multiplied by parameters λ_i

$$R^* = R + \sum_{j=1}^N \lambda_j (y_j - f(x_j)) \quad (4)$$

and then requires the R^* to be a minimum with respect to free variations of the coefficients f_r and the λ_i ,

$$\frac{\delta R^*}{\delta f_r} = f_r c(r^2) - \sum_{j=1}^N e^{-irx_j} \lambda_j = 0 \quad (5)$$

$$\frac{\delta R^*}{\delta \lambda_i} = y_i - \sum_{r=-M}^M f_r e^{-irx_i} = 0 \quad (6)$$

One obtains a set of equations giving the f_r in terms of the λ 's; and the constraint equations. Substituting the expression for the f_r in the constraint equations, one obtains a system of N linear equations in the N unknown parameters λ_i ,

$$y_i = \sum_{i=1}^N \sum_{r=-M}^M \frac{e^{ir(x_i - x_j)}}{c(r^2)} \lambda_j \quad \lambda_j = \sum_{i=1}^N H_{ij} \lambda_j \quad (7)$$

The matrix of coefficients of this set of equations, H , where

$$H_{ij} = h(x_i - x_j), \quad h(x) = \sum_{r=-M}^M \frac{e^{irx}}{c(r^2)} \quad (8)$$

is a positive-definite symmetric matrix and so is easily handled numerically. Furthermore, it depends only upon the coordinates of the zone points where the data is to be fit, and so needs to be computed only once for a particular crystal, and inverted just once to solve the λ -equations. This enables the construction of interpolation coefficients if data for a particular zone needs to be interpolated at the same places several times during the course of a larger calculation,

$$f(x) = \sum_{j=1}^N h(x - x_j) \lambda_j = \sum_{i,j=1}^N h(x - x_j) (H^{-1})_{ji} y_i \quad (9)$$

As an example of the use of this method in an actual crystal calculation, Fig. 3 shows pseudopotential coefficients as computed

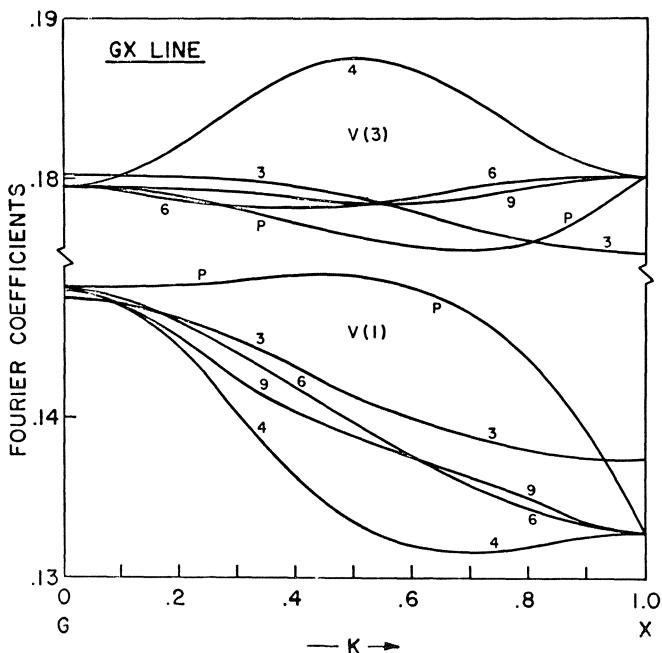


Fig. 3. Interpolation of Pseudopotential Coefficients

at the G, X, L, and W points and interpolated along the G-X axis by the various schemes described above. The curves marked 4 are the result of the straightforward fit of 4 symmetry-adapted "stars" to the four data points. The boundary conditions are correct, and the curves pass through the data, but perform a large excursion along the G-X line. A least-squares fit of three "stars" has a much smoother behavior, and the proper boundary conditions, but misses the data points. A polynomial fit P is relatively smooth, and reaches the data points, but with a non-zero slope at the zone boundary, violating the boundary conditions. The results of using 6 and 9 "stars" with the method of this paper finally show how smoothness is obtained without sacrifice of either the boundary conditions or the computed data. Increasing the number of stars beyond nine had virtually no effect on the shape of the curves. In the curves shown, the first (two) derivatives were equally weighted in the roughness functional. In other calculations, it has been found that the choice of roughness functional can be varied over quite wide limits, with relatively little effect on the resulting interpolation function, although it is well to try several choices to insure that the results characterize the data and physical problem, and not some unfortunate choice of roughness functional.

In summary, the straightforward method of fitting the first N symmetry-adapted functions to a set of data has been shown to yield poor results. Instead, a method of interpolation has been given which uses a larger than minimum set of functions and employs the extra degrees of freedom to reduce the roughness, as measured by an appropriately defined roughness functional. The resulting numerical equations involve solving an $N \times N$ linear system with a positive-definite symmetric matrix of coefficients, which often needs to be done only once for a given crystal problem. Dramatic improvement in the smoothness of the interpolant is obtained with very little additional computation, and the method is adaptable to a wide range of problems and dimensions.

EXACT SOLUTION OF THE TWO-BAND DENSITY OF STATES PROBLEM

N. W. Ashcroft

Laboratory of Atomic and Solid State Physics

Cornell University, Ithaca, New York 14850

Abstract

The band structure arising from the familiar two-band problem admits of an exact solution for the density of states function, which in turn may provide a useful check on its determination by numerical methods.

Many physical aspects of electrons moving in weak single particle potentials are well described by the use of a two band formulation of the problem in which the energy ϵ as a function of \vec{k} is given by the solution of

$$\begin{vmatrix} \frac{\hbar^2}{2m} \vec{k}^2 - \epsilon & V \\ V & \frac{\hbar^2}{2m} (\vec{k} - \vec{k})^2 - \epsilon \end{vmatrix} = 0$$

i.e.

$$\epsilon = \frac{\hbar^2}{2m} \left(\frac{\vec{k}^2 + (\vec{k} - \vec{k})^2}{2} \right) \pm \left\{ \left(\frac{\hbar^2}{2m} \right)^2 \left(\frac{\vec{k}^2 - (\vec{k} - \vec{k})^2}{2} \right)^2 + V^2 \right\}^{\frac{1}{2}}. \quad (1)$$

* Work supported by the Advanced Research Projects Agency through the Materials Science Center at Cornell University, MSC Report #1369.

We are dealing here with a zone plane situated at $\frac{1}{2} \tilde{k}$ and a weak (pseudo-) potential V . In (1), (+) denotes the upper band and (-) the lower. For the present purposes it is convenient to resolve k into its components parallel ($k_{||}$) and perpendicular (k_{\perp}) to \tilde{k} . Then

$$\begin{aligned}\epsilon_k &= \frac{\hbar^2}{2m} k_{\perp}^2 + g_{\pm}(k_{||}) \\ g_{\pm}(k_{||}) &= \frac{\hbar^2}{2m} (k_{||}^2 + \frac{k^2 - 2k_{||}K}{2}) \\ &\pm \left\{ (\frac{\hbar^2}{2m} \frac{(K^2 - 2k_{||}K)}{2})^2 + V^2 \right\}^{1/2}\end{aligned}\quad ..(2)$$

The density of states function (for both spins) is written:

$$\begin{aligned}N(\epsilon) &= \sum_{\tilde{k}} \delta(\epsilon - \epsilon_{\tilde{k}}) \\ &= \frac{V}{4\pi^3} \int dk_{||} 2\pi \int_0^{\infty} dk_{\perp} \delta(\frac{\hbar^2}{2m} k_{\perp}^2 - (\epsilon - g_{\pm}(k_{||}))) \\ &= \frac{V}{4\pi^2} \int dk_{||} (\frac{2m}{\hbar^2}) \int_0^{\infty} d(\frac{\hbar^2}{2m} k_{\perp}^2) \delta(\frac{\hbar^2}{2m} k_{\perp}^2 - (\epsilon - g_{\pm}(k_{||})))\end{aligned}$$

i.e.

$$N(\epsilon) = \frac{V}{4\pi^2} (\frac{2m}{\hbar^2}) \int dk_{||} \theta(\epsilon - g_{\pm}(k_{||}))$$

which is of the form

$$N(\epsilon) = \frac{V}{4\pi^2} (\frac{2m}{\hbar^2}) [k_{||}^{\max} - k_{||}^{\min}]$$

where $k_{||}^{\max}$ and $k_{||}^{\min}$ are the solutions (for each band) to

$$\epsilon = g_{\pm}(k_{||}).$$

Solutions have been given in the past by expanding $g(k_{||})$ in powers of V . We now show that an exact solution can be given.

Let Δ_k be the energy difference between the two bands.
i.e.

$$\Delta_k = 2 \left\{ \left(\frac{\hbar^2}{2m} \right)^2 \left(\frac{k^2 - 2k_{||} K}{2} \right)^2 + V^2 \right\}^{\frac{1}{2}}$$

The surface of constant energy difference ($\Delta_k = \Delta$ say) has equation

$$k_{||} = \frac{K}{2} \left\{ 1 - \frac{1}{4\epsilon_K} \left\{ \Delta^2 - (2V)^2 \right\}^{\frac{1}{2}} \right\} \quad (3)$$

where

$$\epsilon_K = \frac{\hbar^2}{2m} \left(\frac{K}{2} \right)^2$$

is the free electron energy at the zone face. If we now choose a value for ϵ corresponding values of $k_{||}$ and thus of Δ satisfy

$$\begin{vmatrix} H_{11} & V \\ V & H_{22} \end{vmatrix} = 0$$

where

$$H_{11} = \frac{\hbar^2}{2m} \left(\frac{K}{2} \right)^2 \left(1 - \frac{1}{4\epsilon_K} \left\{ \Delta^2 - (2V)^2 \right\}^{\frac{1}{2}} \right)^2 - \epsilon$$

and

$$H_{22} = \frac{\hbar^2}{2m} \left(\frac{K}{2} \right)^2 \left(1 + \frac{1}{4\epsilon_K} \left\{ \Delta^2 - (2V)^2 \right\}^{\frac{1}{2}} \right)^2 - \epsilon.$$

The solutions for Δ are easily shown to be

$$\Delta_1 = \pm 4\epsilon_K \left\{ 1 - \left\{ \frac{\epsilon_F}{\epsilon_K} + \frac{V}{2\epsilon_K} \right\}^2 \right\}^{\frac{1}{2}} \quad (4)$$

$$\Delta_2 = \pm 4\epsilon_K \left\{ 1 + \left\{ \frac{\epsilon_F}{\epsilon_K} + \frac{V}{2\epsilon_K} \right\}^2 \right\}^{\frac{1}{2}} \quad (5)$$

and now from (3) and (2) we find for $\epsilon < \epsilon_K - |V|$

$$N(\epsilon) = \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right) \cdot \frac{K}{2} \left[\left\{ \left(1 + \left\{ \frac{\epsilon}{\epsilon_K} + \left(\frac{V}{2\epsilon_K} \right)^2 \right\}^{\frac{1}{2}} \right)^2 - \left(\frac{V}{2\epsilon_K} \right)^2 \right\}^{\frac{1}{2}}$$

$$- \left\{ \left(1 - \left\{ \frac{\epsilon}{\epsilon_K} + \left(\frac{V}{2\epsilon_K} \right)^2 \right\}^{\frac{1}{2}} \right)^2 - \left(\frac{V}{2\epsilon_K} \right)^2 \right\}^{\frac{1}{2}} \right]$$

which may be simplified by writing $e = \frac{\epsilon}{\epsilon_K}$ and $\alpha = \left| \frac{V}{\epsilon_K} \right|$. Then

$$N(\epsilon) = N_0 \left(\frac{\epsilon}{\epsilon_K} \right)^{\frac{1}{2}} h(e)$$

where

$$h(e) = \left\{ \left(1 + \left\{ e + \frac{1}{4} \alpha^2 \right\}^{\frac{1}{2}} \right)^2 - \frac{1}{4} \alpha^2 \right\}^{\frac{1}{2}}$$

$$- \left\{ \left(1 - \left\{ e + \frac{1}{4} \alpha^2 \right\}^{\frac{1}{2}} \right)^2 - \frac{1}{4} \alpha^2 \right\}^{\frac{1}{2}},$$

$$e < 1 - \alpha \quad (6)$$

In the energy range

$$1 - \alpha < e < 1 + \alpha$$

the solution is

$$h(e) = \left\{ \left(1 + \left\{ e + \frac{1}{4} \alpha^2 \right\}^{\frac{1}{2}} \right)^2 - \frac{1}{4} \alpha^2 \right\}^{\frac{1}{2}} \quad (7)$$

and for

$$e > 1 + \alpha$$

the function $h(e)$ is given by

$$h(e) = \left\{ \left(1 + \left\{ e + \frac{1}{4} \alpha^2 \right\}^{\frac{1}{2}} \right)^2 - \frac{1}{4} \alpha^2 \right\}^{\frac{1}{2}}$$

$$- \left\{ \left(1 - \left\{ e + \frac{1}{4} \alpha^2 \right\}^{\frac{1}{2}} \right)^2 - \frac{1}{4} \alpha^2 \right\}^{\frac{1}{2}} \quad (8)$$

$$= \frac{4 \left\{ e + \frac{1}{4} \alpha^2 \right\}^{\frac{1}{2}}}{\left\{ \left(1 + \left\{ e + \frac{1}{4} \alpha^2 \right\}^{\frac{1}{2}} \right)^2 - \frac{1}{4} \alpha^2 \right\}^{\frac{1}{2}} + \left\{ \left(1 - \left\{ e + \frac{1}{4} \alpha^2 \right\}^{\frac{1}{2}} \right)^2 - \frac{1}{4} \alpha^2 \right\}^{\frac{1}{2}}}$$

$$\approx 2(e)^{\frac{1}{2}} + O(\alpha^2).$$

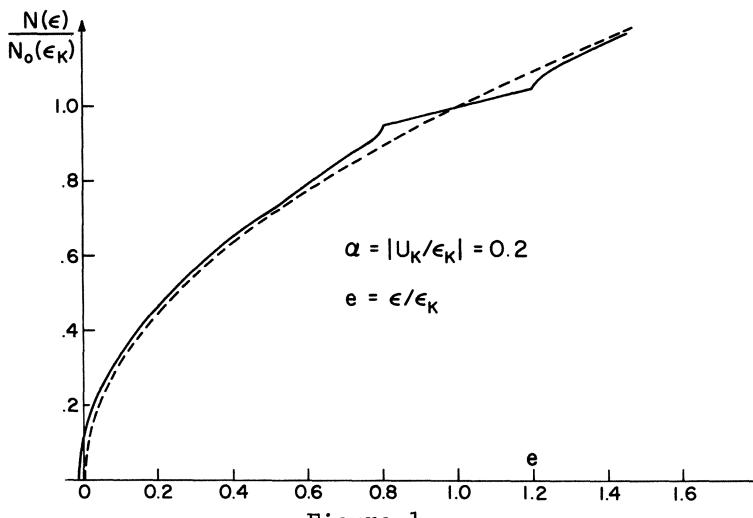


Figure 1

Equations (6), (7), and (8) give curves of the type shown in figure 1 (plotted for $\alpha = 0.2$). These may be particularly useful for checking numerical methods for estimating the density of states.

Finally we make two observations. First, as expected the derivative of the density of states is singular at $e = 1 \pm \alpha$ as can be readily verified from (6). Second, it is a straightforward matter to (i) construct from (6), (7) and (8) the integrated density of states

$$n(e) = \int_{-\frac{1}{4}\alpha^2}^e e' N(e') de'$$

and its deviation from an average of $3/5$ e per particle.¹

1. Results for $U(e)$ may also be found in A. R. Williams and D. Weaire, J. Phys. C; 3, 387 (1969).

THERMOELECTRIC TRANSPORT COEFFICIENTS OF CUBIC CRYSTALS VIA
K-SPACE INTEGRATION

J. M. Schoen

Bell Telephone Laboratories, Inc.

Allentown, Pennsylvania

In cubic crystals the electrical and thermal conductivities and the absolute thermoelectric power can all be expressed in terms of the integral

$$K_o(\epsilon) = \frac{1}{12\pi^3 k^2} \sum_q \int \nabla_k \epsilon \tau(\underline{k}, \epsilon) dS_q(\epsilon) \quad (1)$$

and its first energy derivative at the Fermi surface.¹ Here q is the index of the constant-energy surface sheet and $\tau(\underline{k}, \epsilon)$ is a generalized relaxation time function of the wavevector \underline{k} and the energy ϵ . The integrand of Eq. (1) is analytic everywhere; therefore it should be possible to calculate $K_o(\epsilon)$ with an empirical value of τ or a function $\tau(\underline{k}, \epsilon)$ derived from a model.

$K_o(\epsilon)$ can be evaluated at and near the Fermi energy by means of the algorithm used by L. F. Mattheiss to calculate the critical field anisotropy in superconducting Nb.² The details of the band structure of the crystal can be represented as discrete sets of wavevector radii to the various sheets of the constant-energy surfaces. Lattice harmonic expansions of these wavevector radii are determined by least squares fitting techniques for each sheet and for each energy. The wavevector radii are interpolated to a grid of points by means of the expansions. Each constant-energy-surface-sheet volume is divided into tetrahedra with edges formed

¹J. M. Ziman, "Principles of the Theory of Solids," Cambridge University Press, 1964, pp. 194-204.

²L. F. Mattheiss, Phys. Rev. B 1 373 (1970).

by three neighboring wavevector radii on the grid. Let V_i be the volume enclosed by the i -th tetrahedron and S_i be the constant-energy-surface area intersected by it. Also, let Δ denote the change in V_i with energy. Then

$$\nabla_k \epsilon \approx \Delta \epsilon S_i(\epsilon) / \Delta V_i \quad (2)$$

and using Eq. (2) the integral $K_o(\epsilon)$ can be expressed as

$$K_o(\epsilon) = \frac{1}{12\pi^3 k^2} \sum_q \sum_i \frac{\Delta \epsilon (S_i(\epsilon))^2}{\Delta V_i} \tau(k_i, \epsilon) \quad (3)$$

The above procedure is not applicable to a multiply-connected constant-energy sheet; however, the open sheet can be subdivided into several closed sheets by introducing arbitrary, artificial closing surfaces. A proper choice of the fictitious surfaces will result in a set of closed, constant-energy surfaces with unique wavevector radii. Care must be taken in the evaluation of $K_o(\epsilon)$ to ignore the contributions to the integral from the artificial closing surfaces.

5. THE BAND PROBLEM WITH BOUNDARY CONDITIONS

BANDS, BONDS, AND BOUNDARIES*

K. H. Johnson and F. C. Smith, Jr.

Massachusetts Institute of Technology

INTRODUCTION

Many of the most complex forms of matter may be considered to be built up of large, isolated clusters of atoms or large numbers of periodic and aperiodic clusters of several atoms or more. As examples we may cite:

1. Isolated polyatomic molecules
2. "Molecular" crystals with many atoms per unit cell.
3. Impurity and defect clusters in an otherwise perfect crystal.
4. Amorphous materials
5. Macromolecules important in polymer science and biology.

There is much current interest in developing quantitative theories for the electronic structures and related properties of such systems. However, the practical application of quantum theory to these problems depends, first of all, on one's understanding of the chemical bonding of component polyatomic clusters which are often arranged in quite arbitrary stereochemical configurations.

In this paper we wish to propose that the self-consistent-field scattered-wave (SCF-SW) approach to molecular-orbital theory,^{1,2} by virtue of its relative computational simplicity, its flexibility, and its relationship to the Korringa-Kohn-Rostoker (KKR)³ method

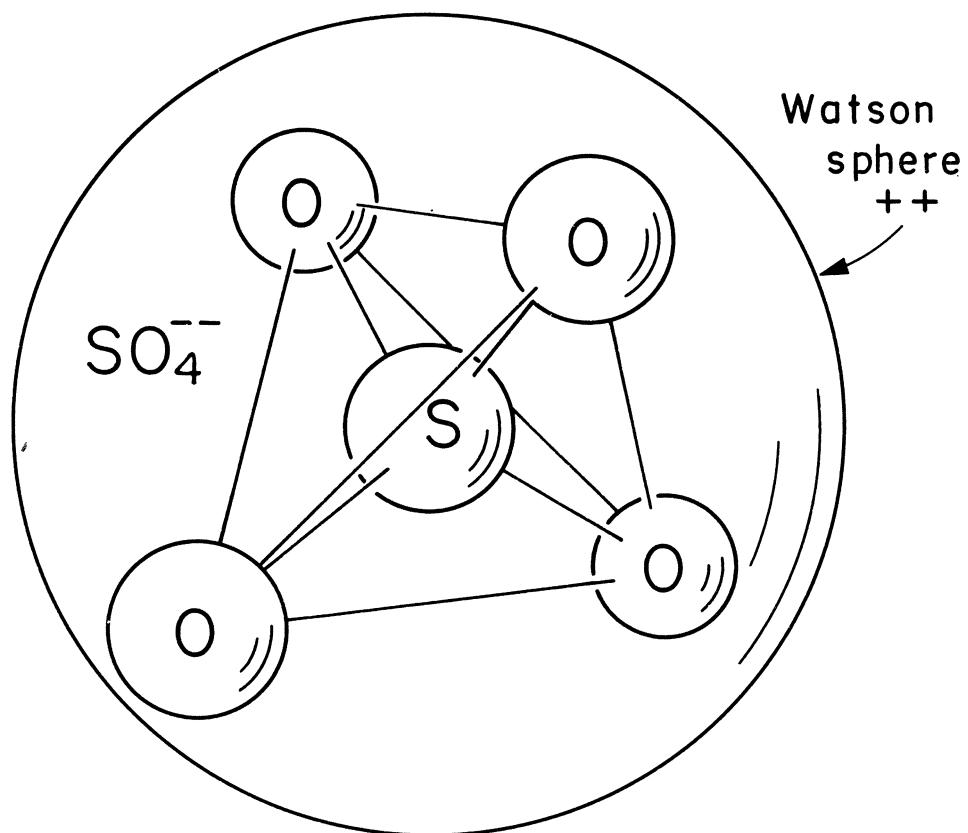


Fig. 1. Sulfate molecular ion

of band theory can be used as the basis for a cluster technique of calculating the electronic structures of complex molecules and solids.

POLYATOMIC MOLECULES

To demonstrate how the SCF-SW model is set up and executed for a polyatomic molecule, let us consider the problem of determining the molecular orbitals of the dinegative sulfate ion (SO_4^{2-}). The sulfate cluster in its equilibrium tetrahedral configuration is illustrated in Fig. (1). A calculation of the one-electron orbitals of a polyatomic molecule such as SO_4^{2-} is complicated, first of all, by its ionicity, by the presence of several "inner" electron shells, and, when conventional Hartree-Fock SCF-LCAO methods⁴ of quantum chemistry are used, the necessity of having to compute many multi-center integrals. For example, a recent ab initio SCF-LCAO calculation on SO_4^{2-} is reported to have required 7.5 hours of computational time on a large-scale computer, most of this time being required for multi-center integral evaluation.⁵ Furthermore, the sulfate complex is not stable in free space, i.e. the gaseous phase, but typically exists as the anion in an ionic crystal (e.g. K_2SO_4). Thus it is important to include, explicitly or implicitly, in a molecular-orbital calculation on such a system the stabilizing effects of the crystal environment. As we shall show, stable orbitals can be calculated for SO_4^{2-} by the SCF-SW technique with relatively little computational effort (a few minutes on an IBM 360 Mod 65 computer).

In the SCF-SW model we partition the sulfate cluster into three contiguous regions:

- I. Intraatomic: the region within nonoverlapping spheres centered on the constituent sulfur and oxygen atoms.
- II. Interatomic: the region between the inner atomic spheres and an outer sphere (the "Watson sphere") centered on the sulfur atom and surrounding the entire cluster.
- III. Outeratomic: the region outside the Watson sphere.

These artificial spherical boundaries are most generally assumed to be touching. The extension of the model to truly overlapping spheres can be carried out in fashion similar to that described by Williams⁶ for the KKR method. The actual sphere radii depend on the nature of the model Hartree-Fock potential chosen to initiate the SCF-SW calculation.

At an arbitrary point \vec{r} of the cluster we can represent the initial molecular potential as a superposition

$$V(\vec{r}) = V^S(|\vec{r}-\vec{R}_0|) + \sum_{j=1}^4 V^O(|\vec{r}-\vec{R}_j|) \quad (1)$$

of spherically symmetric Hartree-Fock-Slater free-atom and free-ion potentials centered at \vec{R}_j ($j=0, \dots, 4$; \vec{R}_0 =center of cluster), including the Slater $X\alpha^7$ statistical approximation

$$V_{X\alpha}(\vec{r}) = -6\alpha[(3/8\pi)\rho(\vec{r})]^{1/3} \quad (2)$$

to exchange correlation. The superposition (1) is then spherically averaged within each atomic sphere j and in the outeratomic region (III). In region (II) the starting potential is that due to a constant charge density obtained by averaging the true charge density over the interatomic volume.

To stabilize the orbitals of SO_4^{2-} we have added a contribution to the potential similar to that originally adopted by Watson⁸ for the calculation of free-ion Hartree-Fock wavefunctions. We distribute a positive charge uniformly around the outer spherical boundary separating the interatomic and outeratomic regions. This "Watson sphere" is therefore a crude approximation to the stabilizing effects of a crystal environment. Charges of +2 [as indicated in Fig. (1)] and of +1 have both been tested.

The partitioning of the space of the molecule into bounded regions of spherically averaged potential allows one to introduce a rapidly convergent, composite partial-wave representation of the molecular orbitals. Within each atomic sphere j of radius b_j (sulfur: $j=0$; oxygen: $j=1, 2, 3, 4$), we expand the orbital wavefunctions in the single-center form

$$\psi_I^j(\vec{r}) = \sum_L C_L^j R_\lambda^j(E; r) Y_L(\vec{r}) \quad (0 < r \leq b_j) \quad (3)$$

where $Y_L(\vec{r})$ are real spherical harmonics of degree $L=(\ell, m)$. The functions $R_\lambda^j(E; r)$ are solutions of the radial Schrödinger equation

$$[-\frac{1}{r^2} \frac{d}{dr} r^2 \frac{d}{dr} + \frac{\ell(\ell+1)}{r^2} + V^j(r) - E] R_\lambda^j(E; r) = 0 \quad (4)$$

for the spherical average $v_j^j(r)$ of the superposition (1) with respect to the j th atomic site. The solutions are generated by outward numerical integration for each trial energy parameter E and partial-wave component ℓ .

In the outeratomic region the orbitals are expanded with respect to the center of the cluster in the representation

$$\psi_{III}(\vec{r}) = \sum_L D_L^0 R_\ell^{out}(E; r) Y_L(\vec{r}) \quad (b_W \leq r < \infty) \quad (5)$$

where b_W is the radius of the Watson sphere. The functions $R_\ell^{out}(E; r)$ are solutions of a radial Schrödinger equation similar to (4) for the spherical average of the potential in the outeratomic region. For localized molecular orbitals, these radial functions must decay exponentially at large distances from the molecule. The solutions are generated by inward numerical integration of the radial equation for trial values of E and ℓ .

For the interatomic region we expand the orbitals in the multicenter partial-wave representation

$$\begin{aligned} \psi_{II}(\vec{r}) = & \sum_L B_L^0 j_\ell(\kappa r_0) Y_L(\vec{r}_0) \quad (b_0 \leq r_0 \leq b_W) \\ & + \sum_{j=0}^4 \sum_L A_L^j f_\ell(\kappa r_j) Y_L(\vec{r}_j) \quad (b_j \leq r_j) \\ & \quad (r_0 \leq b_W) \end{aligned} \quad (6)$$

in which $\vec{r}_j \equiv \vec{r} - \vec{R}_j$ ($j=0, 1, \dots, 4$)

$$\kappa \equiv (E - V_0)^{\frac{1}{2}} \quad (8)$$

$$f_\ell(\kappa r) = \begin{cases} h_\ell^{(1)}(\kappa r) & (E < V_0 < 0; \kappa \text{imag}) \\ n_\ell(\kappa r) & (V_0 < E < 0; \kappa \text{real}) \end{cases} \quad (9)$$

In the above expressions, V_0 is the average interatomic potential, j_ℓ is a spherical Bessel function, $h_\ell^{(1)}$ is a spherical Hankel function of the first kind, and n_ℓ is a spherical Neumann function.

The composite molecular-orbital wavefunctions (3), (5), and (6) and their respective first derivatives are required to be continuous across the adjacent spherical boundaries. This is accomplished via the scattered-wave formalism described in Refs. 1 and 2 and leads to the following relations among the multicenter and

single-center partial-wave coefficients,

$$A_L^j = -ikb_j^2 [j_\ell(\kappa b_j), R_\ell^j(E; b_j)] C_L^j \quad (10)$$

$$B_L^O = -ikb_W^2 [R_\ell^{out}(E; b_W), f_\ell(\kappa b_W)] D_L^O \quad (11)$$

$$\text{where } [j(x), R(x)] \equiv j(x) [dR(x)/dx] - R(x) [dj(x)/dx] \quad (12)$$

The secular equations which lead, in turn, to the molecular-orbital energies and independent partial-wave coefficients can be written in the linear, homogeneous form

$$\sum_{j'=0}^4 \sum_{L'} [T^{-1}(E)]_{LL'}^{jj'} A_L^{j'} - \sum_{L'} S_{LL'}^{jO}(E) B_L^O = 0 \quad (13)$$

$$\sum_{j'=0}^4 \sum_{L'} S_{LL'}^{Oj'}(E) A_L^{j'} - \sum_{L'} [g^{-1}(E)]_{LL'}^{OO} B_L^O = 0$$

in which

$$[T^{-1}(E)]_{LL'}^{jj'} \equiv \delta_{jj'} \delta_{LL'} [t_\ell^j(E)]^{-1} - (1 - \delta_{jj'}) G_{LL'}^{jj'}(E) \quad (14)$$

$$t_\ell^j(E) \equiv \frac{[j_\ell(\kappa b_j), R_\ell^j(E; b_j)]}{[f_\ell(\kappa b_j), R_\ell^j(E; b_j)]} \quad (15)$$

$$G_{LL'}^{jj'}(E) \equiv -4\pi i^{\ell-\ell'} \sum_{L''} i^{-\ell''} I_{L''}(L; L') \times f_{\ell''}(\kappa R_{jj'}) Y_{L''}(\vec{R}_{jj'}) \quad (16)$$

$$g_{LL'}^{OO}(E) \equiv \delta_{LL'} \frac{[f_\ell(\kappa b_W), R_\ell^{out}(E; b_W)]}{[j_\ell(\kappa b_W), R_\ell^{out}(E; b_W)]} \quad (17)$$

$$S_{LL'}^{jo}(E) \equiv -4\pi i^{\ell-\ell'} \sum_{L,L'} i^{-\ell''} I_{L,L'}(L;L') \\ \times j_{\ell''}(kR_{jo}) Y_{L'}(\vec{r}_{jo}) \quad (18)$$

The vectors $\vec{R}_{jj'} \equiv \vec{R}_j - \vec{R}_{j'}$ connect any two atoms of the cluster, and the vectors $\vec{R}_{jo} \equiv \vec{R}_o - \vec{R}_j$ connect each atom with the center of the cluster. The quantities

$$I_{L,L'}(L;L') \equiv \int Y_{L'}(\vec{r}) Y_L(\vec{r}) Y_{L'}(\vec{r}) d\Omega(\vec{r}) \quad (19)$$

are the well known Gaunt integrals.

Real spherical harmonics are used throughout, so that, under the conditions $|\ell-\ell'| \leq \ell'' \leq \ell+\ell'$ and $\ell''+\ell+\ell' =$ even integer for nonvanishing Gaunt integrals, the matrix elements turn out to be real and symmetric. It should also be noted from (18) that

$$S_{LL'}^{jo}(E) = -\delta_{LL'}(j=0) \quad (20)$$

For the purposes of computation, it is also useful to recall that the Bessel and Hankel functions of imaginary argument (which occur in the above formulae for the energy range $E < V_o < 0$) can be written in the real, modified form

$$i_\ell(x) = i^{-\ell} j_\ell(ix) \quad (21)$$

$$k_\ell^{(1)}(x) = -i^{-\ell} h_\ell^{(1)}(ix) \quad (22)$$

The matrix of the secular equations (13) can be factorized considerably by exploiting the equilibrium tetrahedral symmetry of the sulfate ion and using symmetrized combinations of spherical harmonics instead of the ordinary spherical harmonics. This leads to symmetrized secular arrays of only 3x3, 4x4, and 5x5 dimensions if we include allowed partial-wave components up to $\ell_{max}=3$ for the sulfur atom and outeratomic region and up to $\ell_{max}=1$ for equivalent oxygen atoms.

We also have the option of partitioning⁹ and contracting the original set of secular equations (13) prior to symmetrization. This reduces the secular matrices

even further. It also leads to a very interesting interpretation of the SCF-SW model which is particularly useful for extending the cluster formalism to complex "molecular" crystals and to the impurity problem (see below). The contracted secular equations can be written as

$$\sum_{j'=0}^4 \sum_L \left\{ \delta_{jj'} \delta_{LL} [t_\ell^j(E)]^{-1} - W_{LL}^{jj'}(E) \right\} A_L^{j'} = 0 \quad (23)$$

where $W_{LL}^{jj'}(E) \equiv (1 - \delta_{jj'}) G_{LL}^{jj'}(E)$

$$+ \sum_L \sum_{L'} S_{LL'}^{jo}(E) g_{L'}^{oo}(E) S_{L'L}^{oj}(E) \quad (24)$$

Expression (23) is just the partial-wave representation of the inverse of the "T-matrix"¹⁰ for a single electron multiply scattered among a system of nonoverlapping spherical potentials j ($j=0,1,--4$). The fact that this expression is set equal to zero is just the condition for the existence of bound single-particle states (the "T-matrix" itself has "poles" at these states). The "amplitude of scattering" at each potential for each partial-wave component ℓ of energy E is described by the individual "atomic t-matrix" $t_\ell^j(E)$ as defined in expression (15). The "propagation" of the partial waves between any two atoms j and j' is described by the matrix elements $W_{LL'}^{jj'}(E)$ as defined in equation (24). However, the latter is not equal solely to the partial-wave representation $G_{LL'}^{jj'}(E)$ of the "free-space" single-particle Green's function, as it would be for simple truncated "muffin-tin" type potentials. Through the partitioning and contraction of the original secular equations (13), the second term on the right side of (24) effectively "renormalizes" the free-space propagator $G_{LL'}^{jj'}(E)$ to the original boundary condition placed on the cluster orbitals, in the case of SO_4^{2-} the decay of the wavefunction beyond the Watson sphere.

The generation of a set of occupied molecular orbitals and energies for a model potential of the type described above is the starting point for a full SCF-SW calculation within the framework of the statistical $X\alpha^7$ exchange approximation. The initial set of orbitals

leads to a charge density which is used as the basis for generating a new potential. This potential is spherically averaged in the intraatomic and outeratomic regions of the molecule, and is volume averaged in the interatomic region. This result, in turn, serves as the model potential for the first iteration. A new set of orbitals and energies is computed, and the process is repeated until self-consistency in the potential is attained. Five to ten iterations have been sufficient in most applications of the SCF-SW method, thus far, to yield convergence of the molecular-orbital energies to ± 0.005 Hartree. Accuracy greater than this is unwarranted, in view of the uncertainties in the model potential for polyatomic molecules as complex as SO_4^{2-} . For the SCF-SW calculations on SO_4^{2-} reported below, a total computation time of only 5 minutes on an IBM 360 Mod 65 computer was required. This may be contrasted with the 7.5 hours of computing time (on a comparable machine) reported for the more elaborate SCF-LCAO applications to the sulfate ion.⁵

Our calculated results for the occupied molecular-orbital energies of the sulfate cluster are listed in Table I.¹¹ These calculations are based on a choice of $X\alpha$ exchange parameter $\alpha=1$ and Watson sphere of charge +2. The equilibrium sulfur-oxygen bond length in SO_4^{2-} is 1.44 Å. The sulfur and oxygen spheres were assumed to be touching at respective radii where the corresponding starting potentials were equal. The Watson sphere was then assumed to be tangent to the oxygen spheres. Included in the same table for comparison are the results of the SCF-LCAO studies⁵ as well as those of a semiempirical, Wolfsberg-Helmholz (WH) calculation.¹² It is obvious that the SCF-SW model leads to more stable molecular orbitals of SO_4^{2-} than do either of the LCAO applications. The positive occupied orbital energies resulting from the SCF-LCAO work are a peculiar feature of adopting limited atomic-orbital basis sets. It is doubtful that such energies have much physical significance, since one expects the orbitals to be stabilized by the crystal environment. Although the semiempirical WH-LCAO method leads to occupied orbitals which all have negative energy,¹² the ordering of these orbitals is not consistent with experiment. For example, the results of ESR measurements¹³ suggest that the highest occupied orbital should be of symmetry T_1 . The SCF-SW results are consistent with the expected ordering of levels. Moreover, the core-orbital energies calculated by the SCF-SW technique are in reasonable agreement with

Table I. Occupied molecular-orbital energies of SO_4^{2-} .

Molecular Orbital Symmetry	SCF-SW Method E (Hartrees)	SCF-LCAO ^a Method E (Hartrees)	WH-LCAO ^b Method E (Hartrees)
1T ₁	-0.490	+0.1772	-0.323
5T ₂	-0.510	+0.0968	-0.306
1E	-0.556	+0.0044	-0.352
4T ₂	-0.690	-0.1366	-0.426
5A ₁	-0.804	-0.2301	-0.294
3T ₂	-1.106	-0.7922	-0.991
4A ₁	-1.254	-1.0032	-1.138
2T ₂	-6.748	-6.2709	
3A ₁	-8.688	-8.5863	
1T ₂	-19.511	-19.7580	
2A ₁	-19.511	-19.7573	
1A ₁	-90.389	-90.6766	

^a See Ref. 5.^b See Ref. 12.Table II. Distribution of electronic charge in SO_4^{2-} for starting and SCF-SW potentials.

Region of Molecule	Initial Charge (electrons)	SCF Charge (electrons)
Sulfur sphere	-12.9	-12.4
Oxygen sphere	-6.0	-6.4
Interatomic region	-10.6	-11.3
Outeratomic region	-2.5	-0.8

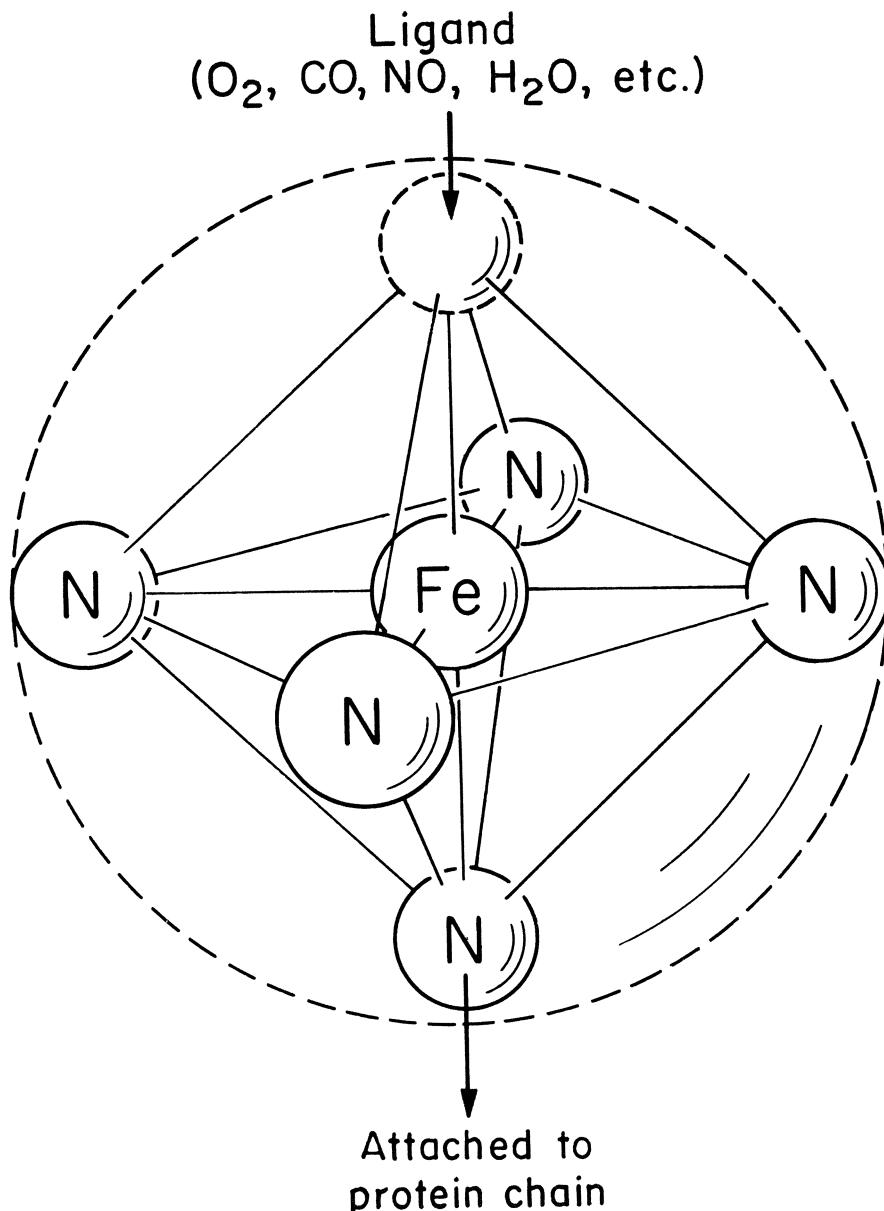


Fig. 2. Iron-porphyrin cluster which constitutes the biologically active center of the hemoproteins.

chemical shift data.¹⁴ Finally, our calculated results for the distribution of electronic charge in SO_4^{2-} are summarized in Table II. A comparison of the initial and SCF charge distributions implies a net transfer of electronic charge between sulfur and oxygen atoms, in general accord with the traditional chemical viewpoint.

Similar SCF-SW calculations are in progress on other polyatomic molecules, e.g. transition-metal complexes such as MnO_4^- , $\text{Fe}(\text{CN})_6^{4-}$, and the iron-porphyrin clusters [See Fig. (2)] which constitute the biologically active centers of the hemoproteins (e.g. hemoglobin, myoglobin, cytochrome, etc.). The results of this work suggest that the scattered-wave model is indeed a practical and reliable new approach to calculating the theoretical electronic structures of complex polyatomic molecules where more conventional methods of quantum chemistry are difficult and costly to implement.

COMPLEX CRYSTALS

Consider now a somewhat different problem, namely that of determining the electronic band structure of a complex crystal such as CuS_2 , a pyrite crystal having 12 atoms per unit cell of which the eight sulfur atoms are clustered into four distinct diatomic "molecules" [see Fig. (3a)]. Application of the conventional KKR³ method of band theory to this problem requires the computation of 66 sets of so called "off diagonal" structure constants $A_{LL'}^{jj'}(E; \vec{k})$ ($j \neq j'$) plus one set of simple-cubic "on-diagonal" structure constants $A_{LL}(E; \vec{k})$. Each of these quantities is a complicated lattice sum that must be evaluated, in principle, at every chosen point \vec{k} of the cubic Brillouin zone, for each trial energy E , and for all allowed partial-wave or lattice-harmonic components L . However, by initially focusing one's attention on the 12-atom unit cell and viewing it as a "molecular cluster" problem similar to that described above for SO_4^{2-} (but with different boundary conditions on the wavefunction), it is possible for one to simplify the band calculation considerably.

To be specific, we can immediately write down the secular equations for the pyrite band-structure problem in a form analogous to (13)

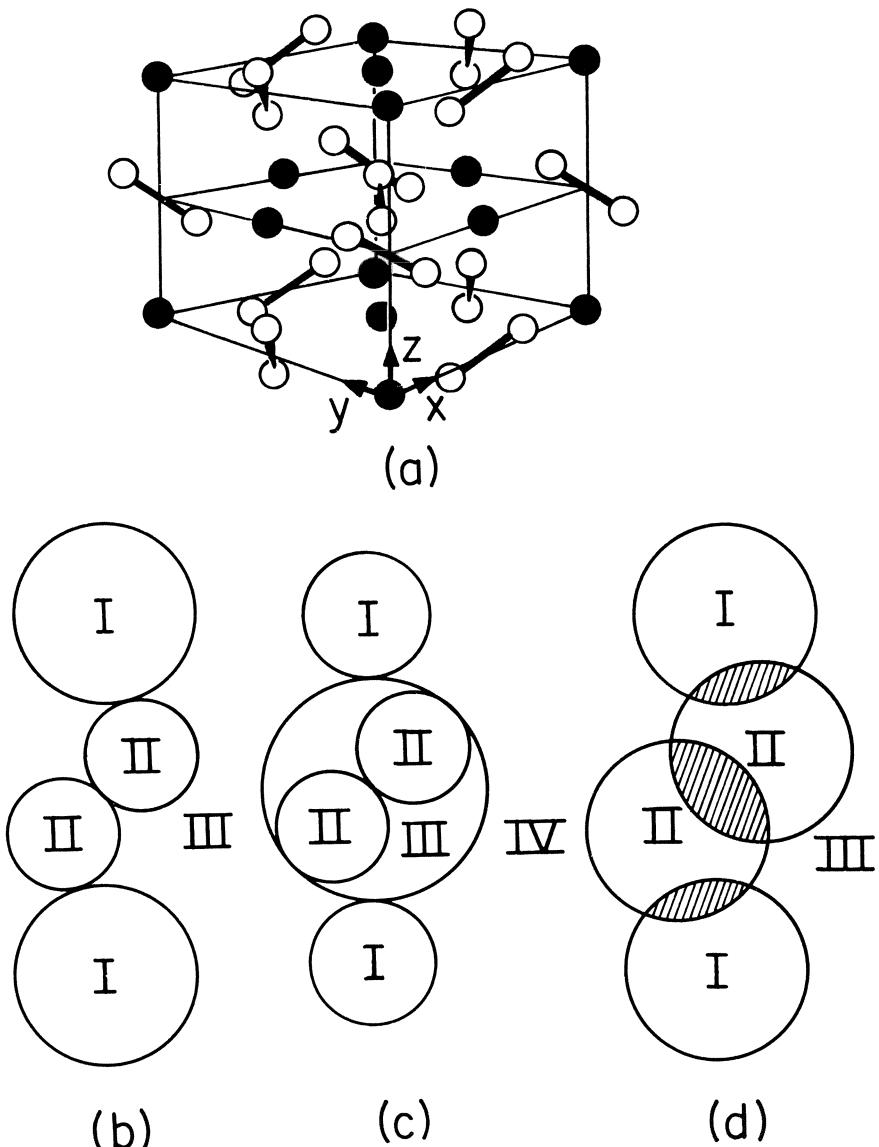


Fig. 3. (a) Unit cell of CuS_2 (pyrite) crystal.
 (b) "Muffin-tin" representation of potentials.
 (c) "Outer-sphere" modification to potentials.
 (d) Overlapping spherical potentials (see Ref. 6).

$$\sum_{j'=0}^{12} \sum_L [T^{-1}(E)]_{LL}^{jj'} A_L^{j'} - \sum_L S_{LL}^{jo}(E) B_L^o = 0 \quad (25)$$

$$\sum_{j'=0}^{12} \sum_L S_{LL}^{oj'}(E) A_L^{j'} - \sum_L [g^{-1}(E; \vec{k})]_{LL}^{oo} B_L^o = 0$$

All the quantities here are as originally defined in equations (14) through (18), except for the matrix elements $g_{LL}^{OO}(E; \vec{k})$, which we have now written as an explicit function of the Bloch wavevector \vec{k} as well as of the energy E . For the calculation of bound molecular orbitals, such as those of SO_4^{2-} , the boundary condition on the isolated polyatomic cluster is the exponential decay of the wavefunction beyond a sphere (the "Watson sphere") artificially surrounding the entire molecule. This boundary condition is described by the quantities $g_{LL}^{OO}(E)$ as originally defined in expression (17). In the present pyrite band-structure problem, we are dealing with a periodic 12-atom "molecule", so that the boundary condition on the wavefunction is just the Bloch condition. This condition is analogously described by the matrix elements $g_{LL}^{OO}(E; \vec{k})$ which can be shown to be identical, except for sign, to the simple-cubic "on-diagonal" KKR structure constants, i.e.

$$g_{LL}^{OO}(E; \vec{k}) = -A_{LL}(E; \vec{k}) \quad (26)$$

We can choose to solve the secular equations as written in the form (25), or we can partition⁹ and contract these equations to a form similar to expression (23), namely

$$\sum_j \sum_L \left\{ \delta_{jj'} \delta_{LL} [t_\ell^j(E)]^{-1} - w_{LL}^{jj'}(E; \vec{k}) \right\} A_L^{j'} = 0 \quad (27)$$

where $w_{LL}^{jj'}(E; \vec{k}) \equiv (1 - \delta_{jj'}) G_{LL}^{jj'}(E)$

$$+ \sum_L \sum_{L'} S_{LL'}^{jo}(E) g_{L-L'}^{OO}(E; \vec{k}) S_{L-L'}^{oj}(E)$$
(28)

Expression (28) may be compared with expression (24). Since the origin of the unit cell is arbitrary in the band-structure problem, we can choose it to be any of the atoms (i.e. $0 = j'$; $j' = 1, 2, \dots, 12$). Because of condition (20) on $S_{L,L'}^{Oj'}(E)$ for $0=j'$, expression (28) reduces to

$$W_{LL'}^{jj'}(E; \vec{k}) = (1 - \delta_{jj'}) G_{LL'}^{jj'}(E) - \sum_{L''} S_{LL''}^{jj'}(E) g_{L''L'}^{OO}(E; \vec{k}) \quad (29)$$

Finally, substitution of identity (26) into (29) and identification of the matrix elements $W_{LL'}^{jj'}(E; \vec{k})$ with the "off-diagonal" structure constants, i.e.

$$W_{LL'}^{jj'}(E; \vec{k}) = -A_{LL'}^{jj'}(E; \vec{k}) \quad (30)$$

lead to the exact formula

$$A_{LL'}^{jj'}(E; \vec{k}) = (\delta_{jj'} - 1) G_{LL'}^{jj'}(E) - \sum_{L''} S_{LL''}^{jj'}(E) A_{L''L'}^{jj'}(E; \vec{k}) \quad (31)$$

relating each "off-diagonal" KKR structure constant to the "on-diagonal" structure constants.

Expression (31) as applied to the pyrite crystal requires the computation of only one independent set of lattice sums $A_{LL'}^{jj'}(E; \vec{k})$, instead of the 67 sets required in the conventional formulation of the KKR³ method. However, the practical use of this formula depends on how fast the sum over $L'' = (\ell'', m'')$ converges. In our work on CuS₂ we have found this convergence to be sufficiently rapid ($4 \leq \ell_{\max}'' \leq 8$) to result in a considerable saving of computer time over the use of the standard KKR technique. The convergence generally improves with increasing number of atoms per unit cell. In the case of crystals with fewer than four atoms per unit cell, the conventional KKR formalism should be adopted. The present theoretical approach is also useful in establishing, within the limits of the model, the degree to which the band structure of the solid is determined by the short-range "molecular clustering."

Table III. Comparison of the energy bands of the CuS₂ crystal at $\vec{k}=0$ with the occupied orbital energies of the S₂ molecule and with the energy levels of the Cu atom.

Symmetry	CuS ₂ Bands E (Hartrees)	Symmetry E (Hartrees)	S ₂ Orbitals E (Hartrees)	Symmetry Cu Levels E (Hartrees)
Γ_2-	-0.303			
Γ_2+	-0.346	$2\pi g$	-0.258	4s -0.254
Γ_3+	-0.434	$5\sigma g$	-0.350	
Γ_4+	-0.474	$2\pi_u$	-0.370	$3d^{10}$ -0.372
Γ_2-	-0.513			
Γ_1-	-0.768	$4\sigma_u$	-0.629	
Γ_1+	-0.914	$4\sigma_g$	-0.763	

In Table III we have listed our calculated energies for a few of the bands of CuS₂ at the center ($\vec{k}=0$) of the cubic Brillouin zone. Because there are 16 sulfur 3s, 32 sulfur 3p, 40 copper 3d, and four copper 4s electrons per unit cell, space does not permit us to list all 46 bands (some of them degenerate) which are necessary to accommodate these electrons within the energy range shown. For this preliminary work we have adopted a "muffin-tin" crystal potential generated from a superposition of Hartree-Fock-Slater copper and sulfur atomic charge densities. However, since the conventional muffin-tin representation of the potential, illustrated schematically in Fig. (3b), is probably not very accurate for crystals as complex as CuS₂, we are also investigating alternative ways of partitioning the potential, such as those shown in Fig. (3c) and Fig. (3d). Included in Table III are the occupied 2π-4σ orbital energies of an isolated, "free-space" sulfur diatomic molecule, determined with the SCF-SW model for a composite potential of the type described in Ref. 2. The 3d and 4s energy levels of a free copper atom in the Hartree-Fock-Slater approximation are also tabulated. The lowest bands Γ_{1+} and Γ_{1-} correspond to the relatively localized S₂ 4σ and 4σ_u levels. Hybridization with the copper 3d and 4s electrons is at least partially responsible for the "chemical shifts" of the S₂ levels to lower energies in the crystal environment. An energy gap of approximately 0.25 au separates these bands from a higher and wider set of bands corresponding to the S₂ 2π_u, 5σ_g, 2π_g and copper 3d, 4s states. The overlap and hybridization of copper and sulfur bands in this energy range are of considerable importance in interpreting available experimental data (Fermi surface, optical, etc.) for CuS₂, which is a metal. While there are uncertainties in the initial choice of crystal potential, such as the effects of "non-muffin-tin" corrections [see Fig. (3d)] these results nevertheless suggest that the occupied sulfur and copper bands are determined largely by the localized charge distributions of the S₂ "molecules" and Cu ions in the crystal.

THE IMPURITY PROBLEM

As our final example, we shall consider the very difficult problem of determining the localized one-electron states of a substitutional impurity cluster in an otherwise perfect crystal. For simplicity we will assume the crystal has only one atom per unit cell. An exactly solvable model for this problem has been derived

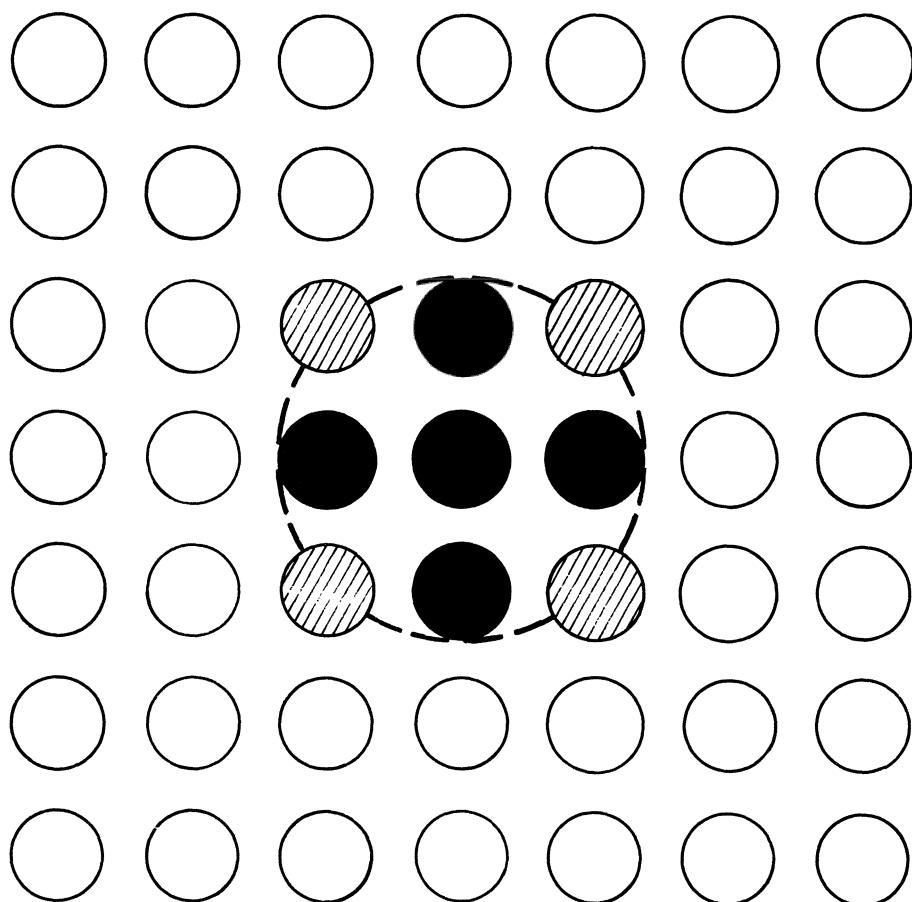


Fig. 4. Substitutional impurity cluster in an otherwise perfect crystal.

earlier by one of the authors (KHJ),¹⁵ and we refer the reader to this reference for details of the formalism. It is sufficient for our purposes here to point out that this model can be interpreted in terms of our scattered-wave approach to molecular clusters. We view the impurity complex (solute atoms plus neighboring "perturbed" atoms of the solvent crystal) as a "molecule" substitutionally introduced into the crystal matrix [see Fig. (4)]. By expanding both the cluster orbitals and the one-electron wavefunctions of the "unperturbed" atoms of the crystal on an equal basis in the partial-wave representation, we are led, via the partitioning and contraction of the original equations, to a finite set of secular equations for bound states of the impurity complex. These secular equations are formally identical to those given in the contracted form (23) for the "free-space" molecular cluster. However the quantities $t_{\ell}^{j'}(E)$ are now the "t-matrices" for the impurity atoms [illustrated in Fig. (4) by the dark shading], as well as for the "perturbed" nearest neighboring atoms of the crystal [illustrated in Fig. (4) by the light shading]. Furthermore, the molecular matrix elements (24) are replaced by more general ones of the form

$$W_{LL'}^{jj'}(E) = -\frac{\tau}{(2\pi)^3} [1 - t_{\ell}^{j'}(E)/t_{\ell}^{j'}(E)] \quad (32)$$

$$\times \int d^3k \exp(-ik \cdot \vec{R}_{jj'}) [\{[A(E; \vec{k})]^{-1} + t(E)\}^{-1}]_{LL'}$$

where the vectors $\vec{R}_{jj'} (j, j' = 1, 2, \dots, N)$ connect various pairs of atoms in the N-atom impurity cluster, $t_{\ell}^{j'}(E)$ are the "t-matrices" for the atoms of the perfect crystal, $A_{LL'}(E; \vec{k})$ are the KKR³ "structure constants" for the perfect crystal, τ is the volume of the unit cell, and the integral is over the Brillouin zone.

The matrix elements (32) contain all the band structure information for the perfect host crystal. This information, however, is not expressed directly in terms of wavefunctions for the perfect crystal, as in more conventional approaches to the impurity problem,¹⁶ but directly in terms of the t-matrices and structure constants used in a KKR calculation of the energy bands. Thus the impurity-cluster problem is reduced to a

"molecular-cluster" problem, analogous to that for the sulfate ion discussed earlier, but "renormalized" to include the band-structure effects of the crystal environment. While we have concentrated on the particular problem of a substitutional impurity complex, the present approach is quite flexible, and, with only slight modification, can be used to determine the localized electronic states associated with vacancies and interstitial impurities. The effects of lattice distortion are also within the scope of the model, and a similar approach has been proposed for the calculation of the localized states of surface and line defects.¹⁷

CONCLUSIONS

Extensive applications of the SCF-SW cluster model to the individual areas of interest in complex molecules and solids discussed above are in progress and will reported in greater detail in separate publications. The purpose of this paper has been to outline the possible advantages of treating several difficult problems of current interest within one self-consistent conceptual and computational framework. The present approach requires that we specify the geometrical and atomic configurations of each cluster. However, the efficiency with which we can now perform reliable numerical calculations on real polyatomic systems as a function of configuration gives us hope of eventually being able to implement some of the recently proposed cluster theories for the structures of disordered and amorphous materials.¹⁸

ACKNOWLEDGMENTS

We are grateful to Professor John C. Slater who originally suggested to us the possibility of adapting the statistical exchange approximation and methods of band theory for the treatment of polyatomic molecules.¹⁹ We wish to thank Dr. R. V. Kasowski for his cooperation in the band structure calculations on CuS₂ and Dr. J. W. D. Connolly for helpful discussions. Finally, we are grateful to the Air Force Office of Scientific Research, U. S. Air Force, for support of this research.

REFERENCES

- * Research supported by the Air Force Office of Scientific Research, U. S. Air Force (Contract No. F44620-69-C-0054).
- 1. K. H. Johnson, "Multiple-Scattering Model for Polyatomic Molecules," *J. Chem. Phys.* 45, 3085 (1966); K. H. Johnson, "Multiple-Scattering Model for Polyatomic Molecules II," *Inter. J. Quantum Chem.* 1s, 361 (1967). Also see Ref. 19. Simpler versions of this approach have been proposed independently by L. Egyes, "Solution of the Schrödinger Equation for a Particle Bound to More than One Spherical Potential," *Phys. Rev.* 111, 683 (1958) and by B. Segall, "On the Calculation of Molecular Orbitals," *Bull. Am. Phys. Soc.* 1, 302 (1956) and private communication.
- 2. F. C. Smith, Jr. and K. H. Johnson, "Scattering Model of Molecular Electronic Structure," *Phys. Rev. Letters* 22, 1168 (1969).
- 3. J. Korringa, "On the Calculation of the Energy of a Bloch Wave in a Metal," *Physica* 13, 392 (1947); W. Kohn and N. Rostoker, "Solution of the Schrödinger Equation in Periodic Lattices with an Application to Metallic Lithium," *Phys. Rev.* 94, 1111 (1954); B. Segall, "Calculation of the Band Structure of Complex Crystals," *Phys. Rev.* 105, 108 (1957).
- 4. J. C. Slater, Quantum Theory of Matter, Second Edition (McGraw-Hill Book Company, New York, 1968), p. 392.
- 5. I. H. Hillier and V. R. Saunders, "Ab Initio Calculations, Using a Small Gaussian Basis Set, of the Electronic Structure of the Sulphate Ion," *Inter. J. Quantum Chem.* 4, 203 (1970).
- 6. A. R. Williams, "Non Muffin-Tin Energy Bands for Silicon by the KKR Method," *Phys. Rev.* (in press).
- 7. J. C. Slater, T. M. Wilson, and J. H. Wood, "Comparison of Several Exchange Potentials for Electrons in the Cu⁺ Ion," *Phys. Rev.* 179, 28 (1969); J. C. Slater, J. B. Mann, T. M. Wilson, and J. H. Wood, "Nonintegral Occupation Numbers in Transition Atoms in Crystals," *Phys. Rev.* 184, 672 (1969); J. C. Slater, "Present Status of the X α Statistical

- Exchange," Semi-Annual Progress Report Number 71, Solid-State and Molecular Theory Group, M.I.T. (July 15, 1969).
8. R. E. Watson, "Analytic Hartree-Fock Solutions for O^{2-} ," Phys. Rev. 111, 1108 (1958).
 9. P. O. Löwdin, "Studies in Perturbation Theory, Part VI. Contraction of Secular Equations," J. Mol. Spect. 14, 112 (1964).
 10. A. Messiah, Quantum Mechanics (North-Holland Publishing Company, Amsterdam, 1962), p. 806.
 11. F. C. Smith, Jr. and K. H. Johnson, "SCF Molecular-Orbital Studies of the Sulfate Ion by the Scattered-Wave Model," Semi-Annual Progress Report Number 73, Solid-State and Molecular Theory Group, M.I.T. (July 15, 1970).
 12. D. M. Bishop, M. Randic, and J. R. Morton, "Electronic Structure of the Sulfate, Thiosulfate, and Related Ions. I. Calculation of Molecular Orbital Energy Levels," J. Chem. Phys. 45, 1880 (1966).
 13. J. R. Morton, D. M. Bishop, and M. Randic, "Electronic Structure of the Sulfate, Thiosulfate, and Related Ions. II. ESR Spectra of SO_4^- and $S_2O_3^-$," J. Chem. Phys. 45, 1885 (1966).
 14. R. Manne, "Molecular Orbitals and Inner-Electron-Shell Chemical Shifts for Sulfur and Chlorine Oxy-anions," J. Chem. Phys. 46, 4645 (1967).
 15. K. H. Johnson, "Scattering Model for the Bound Electronic States of an Impurity Complex in a Crystal," Inter. J. Quantum Chem. 2s, 233 (1968).
 16. G. F. Koster and J. C. Slater, "Wave Functions for Impurity Levels," Phys. Rev. 94, 1392 (1954); 95, 1167 (1954).
 17. K. H. Johnson and J. W. D. Connolly, "On Calculating the Localized Electronic States of Surface and Line Defects," Phys. Letters 28A, 291 (1968).
 18. W. H. Butler and W. Kohn, "Small Cluster Theory of Disordered Systems," NBS Conference on Electronic Density of States (to be published).

19. J. C. Slater, "Suggestions from Solid-State Theory Regarding Molecular Calculations," *J. Chem. Phys.* 43, S228 (1965).

ABSTRACT

The scattered-wave approach to molecular orbital theory, in conjunction with the Korringa-Kohn-Rostoker method of band theory, is used as the basis for a cluster method of calculating the electronic structures of complex molecules, complex solids, and "molecules in solids."

$\vec{K} \cdot \vec{\pi}$ INTERPOLATION AND THE CALCULATION OF VACANCY STATES IN PbTe⁺

G. W. Pratt, Jr. and E. K. Li; [†] F. J. Arlinghaus ^{*}

[†] Massachusetts Institute of Technology

^{*} General Motors Research Laboratories

I INTRODUCTION

The purpose of this paper is to describe the implementation of the Koster-Slater¹ theory of impurity states as applied to vacancies in PbTe. The nature of the vacancy levels and their physical consequences has already been discussed in the literature.² Parada³ has also published a more detailed account of his vacancy calculations. Also the $\vec{K} \cdot \vec{\pi}$ method using APW Bloch functions has appeared in press before. However, many of the key steps have never been fully described. The basic contributions to the $\vec{K} \cdot \vec{\pi}$ method using APW wave functions were made by Ferreira⁴ in his evaluation of the deformation potentials for PbTe. The techniques developed by Ferreira for finding matrix elements of the strain Hamiltonian were later used by him to calculate the momentum matrix elements required in the $\vec{K} \cdot \vec{\pi}$ scheme for PbTe⁵ and later for Bi.⁶ This paper is chiefly concerned with explaining Ferreira's approach to $\vec{K} \cdot \vec{\pi}$. In addition, however, we review the application of the $\vec{K} \cdot \vec{\pi}$ results to the evaluation of vacancy states.

II THE VACANCY PROBLEM

The strategy for the solution of the vacancy problem is outlined in Figure 1 and this section gives a brief description. The starting point is the APW band energies and Bloch functions at $\vec{k} = 0$. The Bloch functions are used to evaluate the $\vec{K} \cdot \vec{\pi}$ matrix elements. As explained by Parada³ the $\vec{\pi}$ operator can be replaced, by \vec{p} for this problem all other contributions being quite small. We treat the momentum matrix elements in detail in the next section. Having found the $\langle b_n | \vec{p} | b_m \rangle$'s the $\vec{K} \cdot \vec{\pi}$ secular equation

[†]Supported by the U.S. Army Research Office--Durham.

was set up and solved over a mesh of points in the Brillouin zone. Parada³ used 1000 points in 1/48th of the zone. The $\vec{k} \cdot \vec{\pi}$ secular equation must be of sufficient size so that the bands of primary interest (valence and conduction) are accurately reproduced by the $\vec{k} \cdot \vec{\pi}$ scheme. Parada³ used 32 bands at $\vec{k} = 0$ to generate 9 bands throughout the zone. As indicated by Figure 1, the band energies $E_n(\vec{k})$ are used to construct the Green's functions $G_n(\vec{R}_p - \vec{R}_q, E)$ for the unperturbed crystal. These are described in references 1, 2, and 3. The eigenvectors of the $\vec{k} \cdot \vec{\pi}$ secular equation are the Bloch functions away from $\vec{k} = 0$. These were used to set up the Wannier functions $a_m(\vec{r} - \vec{R}_q)$. Special care has to be used so that the Wannier functions are well localized and this is commented on in Section IV. Having the $a_m(\vec{r} - \vec{R}_q)$'s allows us to evaluate the matrix elements of the perturbing potential. The very localized nature of a vacancy allows a major simplification discussed in Section V. Finally, having the G and U matrices one solves the secular equation $\det |1 - GU| = 0$ as called for in the Koster-Slater scheme.

III THE MOMENTUM MATRIX ELEMENTS

In the $\vec{k} \cdot \vec{\pi}$ scheme³ the \vec{k} -dependent effective Hamiltonian $H(\vec{k})$ can be generated by subjecting the one electron Hamiltonian to a unitary transformation

$$\exp[-i(\vec{k} - \vec{k}_0) \cdot \vec{r}] \delta(\vec{r} - \vec{r}').$$

Resulting Hamiltonian $H(\vec{k})$ for Bloch solutions at \vec{k} has non-diagonal elements between Bloch states at \vec{k}_0 in the form

$$\vec{k} \cdot \langle b_{\vec{k}_0, n}(\alpha, i) | \vec{\pi} | b_{\vec{k}_0, m}(\beta, j) \rangle,$$

where $\vec{k} = \vec{k} - \vec{k}_0$ and where n and m are band indices. α and i (also β and j) label the irreducible representation and partner number of the point group of the wave vector $G(\vec{k}_0)$ to which a Bloch function belongs. We restrict this discussion to symorphic crystals. $\vec{\pi}$ transforms like a vector operator. Its most significant part is the linear momentum p . Because we intend to use p with augmented plane waves (APW) and because it originated from the term $(-\vec{V}^2)^{1/2}$, which is more correctly $(\vec{V} \cdot \vec{V})^{1/2}$, p is written as $[1/i \partial/\partial_z - 1/i \partial/\partial_i]$. The underlying arrows mean operating on wave functions to the right or to the left.

The Wigner Eckart theorem reduces the number of independent $\vec{\pi}$ elements. The particular form of this theorem suitable here is described by Koster et.al.¹ Cartesian components π_x , π_y , π_z of the operator transform as basis functions for a few particular irreducible representations (IR) of every point group. For example, the IR Γ_{15} of group O_h . Now suppose π_j transforms as IR γ , and the IR α is contained exactly once in the direct product of representations β and γ , then the theorem states that

$$I = \langle b_{\vec{k}}(\alpha, i) | \pi_j^\gamma | b_{\vec{k}}(\beta, l) \rangle / (U_{j,i}^{\beta \alpha})^*$$

is independent of partner indices k , j , and ℓ . The coupling coefficients $U_{j\ell i}^{\gamma\beta\alpha}$ are tabulated in ref. 7.

The symmetrized APW method provides Bloch sums as linear com-

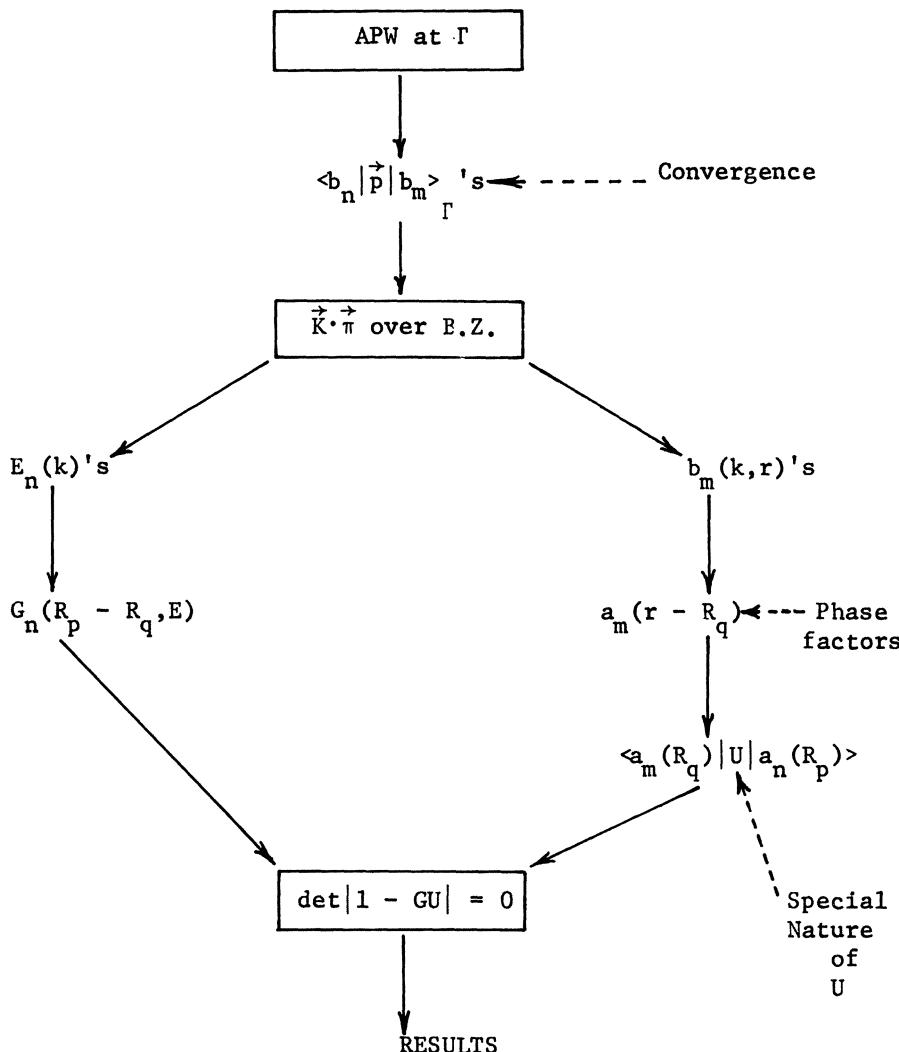


Figure 1. $\vec{K} \cdot \vec{\pi}$ interpolation and vacancy state calculation, outline of major steps. Boxed are names for some computational procedures. Unboxed are quantities constructed from the results of, or fed into, these procedures. Straight arrows indicated flow of operations. Dashed arrows point out a few problems that may require attention.

binations of symmetrized augmented plane waves (SAPW).⁹ Consequently, one might consider that the evaluation of momentum matrix elements between SAPW's would be the obvious thing to do. As we point out this is not the case. A momentum matrix element between SAPW's would have the form

$$\langle \vec{k}_i | A_{LM} | p_z | \vec{k}_j | B'_{NO} \rangle = \sum_R \sum_S A(S)_{LM}^* B(R)_{NO} \langle S\vec{k}_i | p_z | R\vec{k}_j \rangle \quad (1)$$

with $\vec{k}_i = \vec{k} + \vec{k}_i$. B stands for an irreducible representation of the point group $G_o(\vec{k})$. On the right side, its matrices $B(R)_{NO}$ make up the projection operator that forms from the raw APW $|\vec{k}\rangle$ the 0th linearly independent SAPW $|\vec{k}_j | B'_{NO} \rangle$ that transform as partner N. Instead of using SAPW's new functions will be used as explained below. We briefly mention here that time reversal symmetry can further reduce the number of independent matrix elements if we constructed SAPW's that transform as one member of a Kramers pair.³

Since the operator p_z is invariant under a subgroup G_{sub} of the group of the wavevector \vec{k} , the double sum over all elements R and S of $G_o(\vec{k})$ in equation (1) can be reduced to a sum over all $G_o(\vec{k})$ elements followed by a smaller sum of g/g_s elements over a selected set of elements $\{X_1, \dots, X_g/g_s\}$. g and g_s are the orders of groups $G_o(\vec{k})$ and G_{sub} respectively. However, to take advantage of the symmetry property of p_z , one must transform from the SAPW representation to new basis states.

$$\begin{aligned} |\vec{k}_i | A'_{\ell M} \rangle &= \sum_L U_{\ell L} |\vec{k}_i | A_{LM} \rangle \\ &= \sum_L \sum_S U_{\ell L} A(S)_{LM}^* |S\vec{k}_i \rangle \\ &= \sum_S \sum_p A'(S)_{\ell p} U_{pM}^* |S\vec{k}_i \rangle \end{aligned}$$

On this basis, all representation matrices $A'(Y)_{\ell p}$ for elements Y that belong to the subgroup are in reduced form. They have square block matrices coming from irreducible representations of G_{sub} along their diagonal and zero elements outside. L. Ferreira invented these basis vectors and called them transformed symmetrized APWs (TSAPW). He also emphasised that they are no longer SAPW's.⁶

One finds that a momentum matrix element between TSAPW's $\langle \vec{k}_i | A'_{\ell M} | p_z | \vec{k}_j | B'_{nO} \rangle$ vanishes unless

1. The partners ℓ of A' and n of B' belong to the same subgroup irreducible representation $\alpha'(\ell)$ contained in A' (also labelled as $\beta'(n)$ contained in B');

2. Partner indices p_ℓ and p_n , which the function originally indexed by ℓ and n presently takes on in the subgroup irreducible representation, must be identical.

When the conditions 1 and 2 are satisfied, we have

$$\langle \vec{k}_i A' | p_z | \vec{k}_j B' \rangle_{n0} = \sum_T \sum_X (X, T) \langle \vec{k}_i | p_z | T \vec{k}_j \rangle \quad (2)$$

where

$$(X, T) = (g_s / n_\alpha') \sum_t [\sum_L U_{tL} A(X) * \sum_{LMN} V_{tN} B(T)]$$

and n_α' is the dimension of the irreducible representation $\alpha'(\ell)$. If the sum on T covers all elements of $G(\vec{k})$, then the sum on X only involves a set of right coset representatives $\{X_1, X_2, \dots, X_g\}$. The term "right coset representatives" means the set is chosen so that for every element S of $G(\vec{k})$ there is exactly one X in the set and one subgroup element Y that makes $S = YX$. It can be proved that a selection exists¹¹ and there are $(g_s)^g / g_s$ possible selections. In the appendix we shall derive eq. (2).

Although equation (2) was discussed in the framework of the SAPW method, it can be adapted to any other case by substituting for TSAPW's the corresponding transformed symmetrized basis functions. Its result is completely general. If the operator whose matrix elements we are calculating is not a vector but has other transformation properties, a suitable expression can be derived in the same spirit. In particular, one can show that when the operator is invariant under $G_o(\vec{k})$ the formula reduces to the familiar single sum over all elements of the group.¹⁰

The particular form of APW wave functions and the particular crystal structure under investigation often allows further simplification. In the remainder of this section we shall describe an expression for a momentum matrix element between two raw APW's and substitute it into equation (2). Ferreira's computing algorithm for the PbTe case will then be displayed thru a rearrangement of terms.

The matrix elements of p_z between two raw APW's defined by the wave vectors \vec{k}_i and \vec{k}_j (and the energies E_i and E_j) can be expressed in the atomic system as⁶

$$\begin{aligned} 1/m \langle \vec{k}_i E_i | p_z | \vec{k}_j, E_j \rangle &= (k_{iz} + k_{jz}) \delta(\vec{k}_i, \vec{k}_j) \\ &\quad + (4\pi/\Omega) \sum_n R_n^2 \exp[i(\vec{k}_i - \vec{k}_j) \cdot \vec{r}_n] \\ &\quad \times \left\{ j_1(|\vec{k}_j - \vec{k}_i| R_n) / |\vec{k}_i - \vec{k}_j| \right\} (k_{iz} + k_{jz}) \quad (3) \\ &\quad + \sum_{\ell=0}^{\infty} F(\ell, j, i) j_{\ell+1}(k_j R_n) j_{\ell}(k_i R_n) I_n(i, j, \ell) \\ &\quad + \sum_{\ell=0}^{\infty} F(\ell, j, i) j_{\ell}(k_j R_n) j_{\ell+1}(k_i R_n) I_n(j, i, \ell) \end{aligned}$$

where n labels the atomic spheres with positions \vec{r}_n and Radii R_n in a unit cell. The j 's are spherical Bessel functions. The functions F and I are

$$\begin{aligned} F(\ell, i, j) &= (k_{iz}/k_i) P'_{\ell+1}(\cos \theta_{iz}) - (k_{jz}/k_j) P'_{\ell}(\cos \theta_{ij}) \\ I_n(i, j, \ell) &= (Q_{\ell, i}(R_n) Q_{\ell+1, j}(R_n))^{-1} \int_0^{R_n} [Q_{\ell, i}(r) Q'_{\ell+1, j}(r) \\ &\quad - Q_{\ell+1, j}(r) Q'_{\ell, i}(r) + (2(\ell+1)/r) Q_{\ell, i}(r) Q_{\ell+1, j}(r)] dr \end{aligned}$$

where P' are derivatives of Legendre polynomials, Q are r times the radial wave functions u_{nlj} for energies E_i and E_j , and $\cos\theta_{ij} = \hat{k}_i \cdot \hat{k}_j$. We have followed Ferreira⁶ very closely in the above.

A derivation of equation (3) is very similar to Slater's derivation of the matrix elements of $(H-E)$. To do so one may start from a raw APW

$$\psi_{k_i E_i} = \begin{cases} e^{i\hat{k}_i \cdot \hat{r}} & (\text{outside spheres}) \\ e^{i\hat{k}_i \cdot \hat{r}_n} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} (2l+1) i^l ((l-|m|)!/(l+|m|)!) \\ \times (j_l(k_i R_n)/u_{nlj}(R_n)) P_{l,m}(\cos\theta_i) P_{l,m}(\cos\theta_j) u_{nlj}(r') & (\text{in a sphere at } \hat{r} = \hat{r} - \hat{r}') \end{cases}$$

and differentiate it by $\partial/\partial z$. Then one may use relations

$$\sin^2 \theta d/d(\cos\theta) P_{lm}(\cos\theta) = -(\ell(\ell+1-m)/(2\ell+1)) P_{\ell+1,m} + ((\ell+1)(\ell+m)/(2\ell+1)) P_{\ell-1,m}$$

$$\cos\theta P_{lm}(\cos\theta) = ((\ell+1-m)/(2\ell+1)) P_{\ell+1,m} + ((\ell+m)/(2\ell+1)) P_{\ell-1,m}$$

that can be obtained from ladder operators and addition formulae for orbital angular momentum.¹² On integration over θ' and ϕ' one uses the orthogonality and normalization properties of P_{lm} . Finally the addition theorem and the recurrence relations for P_{lm} brings equation (3).¹³

We make the following observations before substituting equation (3) into equation (2). First of all functions $j_\ell, j_{\ell+1}$ and $I_n(i,j,l)$ are independent of the directions \hat{k}_i and \hat{k}_j . Secondly the argument of Legendre polynomials in the function F , and the quantity $|k_j - k_i|$ depend only on the direction cosine $v = \cos\theta_{ij}$. Moreover in PbTe, one atom at $\hat{r}_n = (1,1,1)\pi/a$. This makes $\exp[i\hat{k}_i \cdot \hat{k}_j] \cdot \hat{r}_n$ very simple functions of $(\hat{k}_i \cdot \hat{k}_j)$. These observations leads to the following algorithm.

Three kinds of sums over symmetry operations X and T that give constant $(X\hat{k}_i \cdot T\hat{k}_j)$ are performed. They are

$$FI(v) = \sum_{X,T} (X,T) (X\hat{k}_i)_z / \hat{k}_i$$

$$FJ(v) = \sum_{X,T} (X,T) (T\hat{k}_j)_z / \hat{k}_j$$

$$FIJ(v) = \sum_{X,T} (X,T) [(X\hat{k}_i)_z + (T\hat{k}_j)_z]$$

the primes indicate $(X\hat{k}_i \cdot T\hat{k}_j) = v$

Subsequently, the momentum matrix element is computed from the expression

$$\begin{aligned}
 & \langle \vec{k}_i | A'_{\ell M} | p_z | \vec{k}_j | B'_{n0} \rangle = \sum_v \{ (x_{\vec{k}_i}, T_{\vec{k}_j}) \\
 & - (4\pi/\Omega) \sum_n R_n^2 e^{i(T_{\vec{k}_j} - x_{\vec{k}_i}) \cdot \vec{r}_n} (j_1(|T_{\vec{k}_j} - x_{\vec{k}_i}| R_n) / |T_{\vec{k}_j} - x_{\vec{k}_i}|) \cdot F_{IJ}(v) \} \\
 & + (4\pi/\Omega) \sum_{\ell} \sum_n \{ j_{\ell+1}(k_j R_n) j_{\ell}(k_i R_n) I_n(ij\ell) \sum_v e^{i(T_{\vec{k}_j} - x_{\vec{k}_i}) \cdot \vec{r}_n} \\
 & \quad \times [P'_{\ell+1}(T_{\vec{k}_j} \cdot x_{\vec{k}_i}) F_{IJ}(v) - P'_{\ell}(T_{\vec{k}_j} \cdot x_{\vec{k}_i}) F_{IJ}(v)] \\
 & + j_{\ell}(k_j R_n) j_{\ell+1}(k_i R_n) I_n(ji\ell) e^{i(T_{\vec{k}_j} - x_{\vec{k}_i}) \cdot \vec{r}_n} \\
 & \quad \times [P'_{\ell+1}(T_{\vec{k}_j} \cdot x_{\vec{k}_i}) F_{IJ}(v) - P'_{\ell}(T_{\vec{k}_j} \cdot x_{\vec{k}_i}) F_{IJ}(v)] \}
 \end{aligned}$$

Ferreira's computations indicate that sums on converge rapidly.¹⁰

IV THE WANNIER FUNCTIONS

The Wannier function for the nth band

$$a_{n,i}(\vec{k} - \vec{R}_q) = (\Omega_k)^{-1/2} \int_{BZ} d\vec{k} e^{-i\vec{k} \cdot \vec{R}_q} b_{n,(\beta,i)}(\vec{k}, \vec{r}) e^{i\theta_n(\vec{k})}$$

Here $b_{\Gamma\beta}(\vec{k})$ is the Bloch function of the nth band which transforms like the i th partner of the irreducible representation $\Gamma_\beta(k)$ of the group of the wave vector \vec{k} , Ω_k is the volume of the first Brillouin zone and $\theta_n(\vec{k})$ is a phase factor which must be chosen carefully in order that the Wannier functions are localized satisfactorily. Ferreira and Parada¹⁴ have made a very interesting variational formulation of the problem of choosing the phase factors that lead to a minimum second moment for the charge density of the Wannier function. We will mention only their results and leave the reader to consult reference 14 for details. It is shown there that the Bloch functions to be used in the construction of minimum width Wannier functions is that set that is generated by a point by point (PP) $\vec{k} \cdot \vec{\pi}$ expansion starting at some initial point in \vec{k} -space, say \vec{k}_0 . Knowing the Bloch functions at \vec{k}_0 one finds them at $\vec{k}_0 + \vec{s}_k$ by $\vec{k} \cdot \vec{\pi}$. The phases of the Bloch functions at $\vec{k}_0 + \vec{s}_k$ are correctly related to those at \vec{k}_0 for minimum widths. One goes point by point through the zone making new $\vec{k} \cdot \vec{\pi}$ expansions about successive points to determine the appropriate phases. In other words, $\theta_n(\vec{k}) = 0$ for all \vec{k} , if point by point Bloch functions had been used.

Ferreira and Parade have derived a first order differential equation in \vec{k} whose solutions are the set of phases that give the point to point Bloch functions in terms of single point Bloch functions (SP). The latter are obtained directly from the $\vec{k} \cdot \vec{\pi}$ scheme as discussed in sections I and III, where one expands Bloch functions for all \vec{k} in terms of those at a single point \vec{k}_0 . They also have shown that in the special case of weak spin orbit coupling in a crystal with inversion symmetry the SP Bloch functions coincide with PP. This is what was done for the PbTe problem using $k = \Gamma$ as the single expansion point.

V MATRIX ELEMENTS FOR SHARPLY LOCALIZED DEFECT POTENTIALS

An essential simplification results from
 $U(\vec{r}) \exp[i(\vec{k} - \vec{k}') \cdot \vec{r}] = U(\vec{r}) + i(\vec{k} - \vec{k}') \cdot \vec{r} U(\vec{r})$
 being a very good approximation when $U(r)$ is sharply localized.
 The matrix elements of $U(r)$ between well localized Wannier functions

$$\begin{aligned} & \langle a_n'(\vec{r} - \vec{R}_q') | U(\vec{r}) | a_n(\vec{r} - \vec{R}_q) \rangle \\ &= (1/N_R) \sum_{\vec{k}} \sum_{\vec{k}'} \exp[i(\vec{k}' \cdot \vec{R}_q' - \vec{k} \cdot \vec{R}_q)] \\ & \quad \times \sum_{m,m} c_{n'm}(\vec{k}') c_{nm}(\vec{k}) \\ & \quad \times \int d^3r [b_m^{PP}(ko, \vec{r})]^* \exp[i(\vec{k} - \vec{k}') \cdot \vec{r}] U(\vec{r}) b_m^{PP}(\vec{k}o, \vec{r}) \end{aligned}$$

no longer requires a separate integral for every vector value of $(\vec{k} - \vec{k}')$. Instead only two types of integrals

$$\langle b_m^{PP} | U(\vec{r}) | b_m^{PP} \rangle$$

and

$$\langle b_m^{PP} | \vec{r}U(\vec{r}) | b_m^{PP} \rangle$$

are necessary. Furthermore if $U(\vec{r})$ possesses the full point group symmetry of the crystal, then these discussion in Section III again applies.

Finally we give the definition of the G matrix

$G_{n'n}(\vec{R}_p - \vec{R}_q, E) = \delta_{n'n} N_k^{-1} [E - E_n(\vec{k})]^{-1} \exp[i\vec{k} \cdot (\vec{R}_p - \vec{R}_q)]$
 where N_k is the number of allowed \vec{k} -vectors in the Brillouin zone.
 The defect states are solutions to $\det|I - GU| = 0$.

REFERENCES

1. G.F. Koster and J.C. Slater, Phys. Rev. 96, 1208 (1954).
2. N.J. Parada and G.W. Pratt, Jr., Phys. Rev. Lett. 22, 180 (1969).
3. N.J. Parada, Phys. Rev. in press. See also Ph. D. Thesis
Dept. of Elec. Eng. MIT (1968).
4. L.G. Ferreira, Phys. Rev. 137, A1601 (1965).
5. G.W. Pratt, Jr. and L.G. Ferreira, Proc. Int. Conf. on Phys.
of Semicond. Dunod, Paris (1964).
6. L.G. Ferreira, J. Phys. Chem. Sol. 28, 1891 (1967),
7. J.C. Slater, Phys. Rev. 51, 846 (1937).
8. G.F. Koster, J. Dimmock, R.G. Wheeler and H. Statz, "Properties
of the 32 Point Groups" (MIT Press 1963). See section on
coupling coefficients.
9. L.F. Mattheis, J.H. Wood, A.C. Switendick, on the APW method
in "Methods in Computational Physics", Vol 8 (AP 1968).
10. L.G. Ferreira, "Calculation of Electronic Properties of
Strained Lead Telluride", Ph.D. Thesis Dept. of Elec. Eng,
MIT (1964).

11. See for example, J.F. Cornwel "Group Theory and Electronic Energy Bands in Solids", (North Holland - J. Wiley 1969).
12. E.U. Condon and G. H. Shortley, "Theory of Atomic Spectra" (Cambridge University Press 1967).
13. E. Janke and F. Emde, "Tables of Functions", (Dover paperback)
14. L.G. Ferreira and N.J. Parada, Phys. Rev. in press.

APPENDIX

Let the unitary transforms U and V that transforms representation A to A' and B to B' respectively be chosen such that each subgroup irreducible representation contained in A' is either identical to or non-equivalent to subgroup irreducible representations contained in B' . This choice makes the usual orthogonality theorem for representation matrices applicable as we shall see later. Write the momentum matrix element between TSAPWs as

$$\langle \vec{k}_k | A'_{\ell M} | p_z | \vec{k}_j | B'_{n0} \rangle = \sum_{R,S} \sum_{p,q} A'(S)_{\ell p}^* U_{pM}^* B'(R)_{nq}^* V_{q,0} \langle S \vec{k}_i | p_z | R \vec{k}_j \rangle$$

Using subgroup G_s , partition the group of the wave vector $G_s(\vec{k})$ into right cosets with respect to coset representatives $x_1^o, x_2^o, \dots x_{g/g_s}^o$. That is to say, we select the set $\{x_1^o, x_2^o, \dots x_{g/g_s}^o\}$ as discussed in the text and have $S = YX$, so the element becomes

$$\sum_R \sum_Y \sum_{p,q} U_{pM}^* V_{q0}^* A'(YX)_{\ell p}^* B'(R)_{nq}^* \langle X \vec{k}_i | p_z | Y^{-1} R \vec{k}_j \rangle$$

where we also used $Y p_z Y^{-1} = p_z$. The rearrangement theorem then gives, $R = YT$, and

$$\begin{aligned} & \sum_T \sum_X \sum_{p,q} U_{pM}^* V_{q0}^* \sum_{r,s} A'(X)_{r,p}^* B'(T)_{s,q}^* \\ & \cdot [\sum_Y A'(Y)_{\ell r}^* B'(Y)_{ns}^*] \langle X \vec{k}_i | p_z | T \vec{k}_j \rangle \end{aligned}$$

The orthogonality theorem applied to reduced matrices $A'(Y)$ and $B'(Y)$ shows that the sum over all $Y \in G_{\text{sub}}$ vanishes unless

1. The four partners ℓ, r and A' and n, s of B' belong to the same irreducible representation of G_{sub} ;
2. ℓ and n are the same partners in that subgroup irreducible representation; similarly for r and s .

Equation (2) thus follows after we transform back to matrices $A(X)$ and $B(T)$.

A KKR METHOD FOR TWO-DIMENSIONAL LATTICES AND ITS APPLICATION TO
BAND CALCULATION

Kyozaburo Kambe

Fritz-Haber-Institut der Max-Planck-Gesellschaft

(1) Berlin 33, Faradayweg 4-6, Germany

ABSTRACT

A method of band calculation which is a kind of combination of KKR and plane-wave expansion is described. A Bloch function is expanded into plane waves only on planes between atomic layers. Matrix elements between these plane waves are calculated from the scattering property of the atomic layers by means of the KKR method modified to a two-dimensional form. Apart from the obvious possibilities of application to crystals with surfaces (e.g., LEED) or interfaces, there are several characteristics which appear to favor this method in proper band calculation.

1. Introduction

In the theory of LEED the KKR method has been modified to a two-dimensional form for the treatment of the scattering of electrons by atomic layers (ref. 1,2). It has been pointed out by McRae (ref. 3) and Kambe (ref. 2) that this method can be applied to the calculation of Bloch functions in crystals. Since the crystals are divided into layers, this method is closely related to the method of Marcus and Jepsen (ref. 4) which also calculates Bloch functions in connection with LEED.

It is obvious that this method is particularly suitable for problems such as LEED which have to do with crystals

with surfaces and also with interfaces such as stacking faults and twin boundaries. Aside from the possibilities of application this method appears to have several advantages which suggest that it may be proposed as an alternative to the usual methods of band calculations. The purpose of the present article is to discuss the present method from this point of view.

There are as yet only few examples of application of the present method to band calculation. In fact all these were done in connection with LEED. Jennings and McRae (ref.5) calculated the band structure of W along (100) line using Mattheiss' potential (ref.6) and obtained a good agreement with the results of Mattheiss' APW calculation. Jennings (ref.7) calculated the band structure of a model potential, the atom of which consisting of the nucleus and the constant negative charge filling up the muffin-tin sphere. His results agreed well with the usual KKR calculation of Williams (ref.8) using the same potential.

2. Theory

The crystal is divided into equivalent atomic layers parallel to some crystallographic plane (in the case of LEED parallel to the surface). A two-dimensional unit of the atomic layer may be one or more atoms according to the complexity of crystal structure.

We take as reference planes those which lie somewhere (usually just halfway) between the atomic layers. Thus the reference planes are an infinite set of equally-spaced planes. It is assumed that in the neighbourhood of the reference planes the crystal potential is constant. This assumption is not essential for the present method but convenient for understanding the following plane-wave expansion (ref.9). Since the potential is constant near the reference planes every solution of the Schrödinger equation in the lattice can be expanded into plane waves on the reference planes. For Bloch functions which are going to be calculated the energy E and the component of the propagation vector parallel to the layers K_t are assumed to be given and the perpendicular component k_z is to be calculated, the z-axis being taken normal to the layers. If E and K_t are given, the above plane-wave expansion can contain only a discrete set of plane waves. Thus

$$(1) \quad \Psi(\underline{r}) = \sum_p [\varphi_p^+ \exp\{i \underline{K}_p^+ \cdot \underline{r}\} + \varphi_p^- \exp\{i \underline{K}_p^- \cdot \underline{r}\}] ,$$

where

$$(2) \quad \underline{K}_p^+ = (K_{pt}, \Gamma_p), \quad \underline{K}_p^- = (K_{pt}, -\Gamma_p),$$

where

$$(3) \quad \underline{K}_{pt} = \underline{K}_t + 2\pi \underline{b}_{pt},$$

where \underline{b}_{pt} is a two-dimensional reciprocal-lattice vector, and

$$(4) \quad \Gamma_p = \sqrt{x^2 - |\underline{K}_{pt}|^2} \quad \text{if} \quad x^2 > |\underline{K}_{pt}|^2,$$

and

$$\Gamma_p = i\sqrt{|\underline{K}_{pt}|^2 - x^2} \quad \text{if} \quad x^2 < |\underline{K}_{pt}|^2,$$

where

$$(5) \quad x^2 = 2mE / \hbar^2.$$

φ_p^+ and φ_p^- are the amplitudes of plane waves which propagate to the $+z$ and $-z$ directions respectively. It is to be noted, however, that Γ_p is real only for a limited number of waves, all other waves being "evanescent" having imaginary Γ_p and decaying exponentially to the $+z$ and $-z$ directions respectively.

The infinite set of reference planes are enumerated successively, and the values of φ_p^+ and φ_p^- on the n -th reference plane are written as $\varphi_{p(n)}^+$ and $\varphi_{p(n)}^-$. If the values of $\varphi_{p(n-1)}^+$ and $\varphi_{p(n-1)}^-$ are known, the values of $\varphi_{p(n)}^+$ and $\varphi_{p(n)}^-$ can be derived from the scattering property of the atomic layer lying between the $(n-1)$ -th and n -th reference planes. This relation can be written in the vector-matrix form

$$(6) \quad \underline{\varphi}_{(n)} = \underline{Q} \underline{\varphi}_{(n-1)}$$

$\underline{\varphi}_{(n)}$ is a vector which has $\varphi_{p(n)}^+$ and $\varphi_{p(n)}^-$ as its elements:

$$(7) \quad \underline{\varphi}_{(n)} = (\dots, \varphi_{p(n)}^+, \dots, \dots, \varphi_{p(n)}^-, \dots)$$

\underline{Q} is a matrix which is called by McRae (ref.3) the transfer matrix. How the energy bands and the Bloch functions are derived from the eigenvalues and the eigenvectors of \underline{Q} is described in detail elsewhere (ref.2,3,9) so that it is omitted here for brevity.

The transfer matrix \underline{Q} is calculated from the scattering property of the atomic layer by means of a KKR method for a two-dimensional lattice (Kambe, ref.1,2). The results of ref.1,2 can be easily generalized to give the relations between the wave amplitudes $\varphi_{p(n-1)}^\pm$ and $\varphi_{p(n)}^\pm$. From ref.2, Eq.(2.12)

$$(8) \quad \begin{aligned} & \sqrt{4\pi} \sum_p [Y_{lm}^*(\theta_{kp}^+, \varphi_{kp}) \exp\{i \underline{k}_p^+ \cdot \underline{c}_i\} \varphi_{p(m-1)}^+ \\ & \quad + Y_{lm}^*(\theta_{kp}^-, \varphi_{kp}) \exp\{i \underline{k}_p^- \cdot \underline{c}_i\} \varphi_{p(m)}^-\] \\ & = x_{lm}^{(i)} \cos \hat{\eta}_l^{(i)} + \sum_{i' l'm'} \frac{A_{l'm'l'm'}^{(i'i')}}{x_{i'-e'}} (\sin \hat{\eta}_{l'}^{(i')}) x_{l'm'}^{(i')}, \end{aligned}$$

where $Y_{lm}(\theta, \varphi)$ are spherical harmonics, θ_{kp}^\pm and φ_{kp} the polar coordinates of the vectors \underline{k}_p^\pm , \underline{c}_i the position vectors of the i-th atom center, $x_{lm}^{(i)}$ the reduced expansion coefficients of the wave function on the i-th muffin-tin sphere, $\hat{\eta}_l^{(i)}$ the atomic phase shifts reduced in the range $(-\pi/2, \pi/2)$, $A_{l'm'l'm'}^{(i'i')}$ the structure constants. From ref.2, Eqs.(4.5) and (4.6)

$$(9) \quad \begin{aligned} \varphi_{p(m-1)}^- &= \varphi_{p(m)}^- + \frac{i 2\pi}{A \times \Gamma_p} \sum_{i' l'm'} \exp\{-i \underline{k}_p^- \cdot \underline{c}_i\} \sin \hat{\eta}_{l'}^{(i')} \sqrt{4\pi} Y_{lm}(\theta_{kp}^-, \varphi_{kp}) x_{lm}^{(i')}, \\ \varphi_{p(m)}^+ &= \varphi_{p(m-1)}^+ + \frac{i 2\pi}{A \times \Gamma_p} \sum_{i' l'm'} \exp\{-i \underline{k}_p^+ \cdot \underline{c}_i\} \sin \hat{\eta}_{l'}^{(i')} \sqrt{4\pi} Y_{lm}(\theta_{kp}^+, \varphi_{kp}) x_{lm}^{(i')}, \end{aligned}$$

where A is the area of the two-dimensional unit cell. From (8) $x_{lm}^{(i)}$ can be expressed as functions of $\varphi_{p(m-1)}^+$ and $\varphi_{p(m)}^-$. In practice it implies a numerical inversion of the matrix of coefficients of $x_{lm}^{(i)}$ on the right-hand side of (8). Substitution into (9) and rearrangement lead to the expression (6).

The structure constants are given by

$$(10) \quad A_{l'm'l'm'}^{(i'i')} = 4\pi i^{l-l'} \sum_L i^{-L} C_{L,m-m'; lm; l'm'} D_{L,m-m'}^{(i'i')},$$

where

$$(11) \quad C_{LM; lm; l'm'} = \int Y_{LM}(\theta, \varphi) Y_{lm}^*(\theta, \varphi) Y_{l'm'}(\theta, \varphi) d\Omega,$$

and

$$(12) \quad |l-l'| \leq L \leq l+l'.$$

In practice $C_{LM; lm; l'm'}$ are evaluated by Gaunt's formula (ref.10). For the case of only one atom per two-dimensional unit cell the structure constants $D_{LM}^{(i'i')}$ are given (dropping i, i'), if $L-M$ is odd, by

$$D_{LM} = 0 \quad , \text{ and if } L-M \text{ is even,}$$

by Ewald's summation method by

$$(13) \quad D_{LM} = D_{LM}^{(1)} + D_{LM}^{(2)} + \delta_{LD} D_{LM}^{(3)} ,$$

$$(14) \quad D_{LM}^{(1)} = -\frac{1}{Ax} \frac{i^{|M|+1}}{\sum} [(2L+1)(L+|M|)! (L-|M|)!]^{\frac{1}{2}} \\ \times \sum_p \exp\{-iM\varphi_p\} \sum_{m=0}^{(L-|M|)/2} \frac{(K_{pt}/x)^{L-2m} (\Gamma_p/x)^{2m-1}}{m! [\frac{1}{2}(L-|M|-2m)]! [\frac{1}{2}(L+|M|-2m)]!} \\ \times \Gamma((1-2m)/2, e^{-\pi i} \propto \Gamma_p^2/x^2) ,$$

where $K_{pt} = |K_{pt}|$, $\Gamma(a, x)$ is an incomplete gamma function, and

$$(15) \quad D_{LM}^{(2)} = -\frac{x}{4\pi} \frac{(-1)^L (-1)^{(L-|M|)/2}}{2^L (\frac{L-|M|}{2})! (\frac{L+|M|}{2})!} [(2L+1)(L-|M|)! (L+|M|)!]^{\frac{1}{2}} \\ \times \sum_m^{|a_{mt}|=0} \exp\{-i(K_{ta_{mt}} + M\varphi_{a_{mt}})\} \left(\frac{x a_{mt}}{2}\right)^L \int_0^x u^{-\frac{3}{2}-L} \exp\{u - \frac{(x a_{mt})^2}{4u}\} du ,$$

where \underline{a}_{mt} are the real two-dimensional lattice vectors, $a_{mt} = |\underline{a}_{mt}|$, $\varphi_{a_{mt}}$ the polar angle of \underline{a}_{mt} , and

$$(16) \quad D_{00}^{(3)} = -\frac{x}{2\pi} \left[2 \int_0^{\sqrt{x}} e^{t^2} dt - \frac{e^x}{\sqrt{x}} \right] .$$

In practice the incomplete gamma functions are derived by a recurrence formula from $\Gamma(1/2, x)$, which is expressed by Chebyshev series. The integral in (15) is also derived by a recurrence formula from a complex error function, which is evaluated by a standard technique (ref.11).

$D_{00}^{(3)}$ is expressed by a Chebyshev series.

The convergence of the series in (14) and (15) depends on the choice of the separation constant α . Estimation of the terms shows that for large real and reciprocal lattice vectors the terms decrease exponentially. It can be shown that

$$(17) \quad \alpha = \frac{x^2 A}{4\pi}$$

is the proper choice for obtaining the same exponential decrease in (14) and (15). It has been found in practice that this choice works satisfactorily.

3. Discussion

In the present method the wave function is represented inside the atomic layers by expansion in spherical harmonics by use of the KKR method. The plane-wave expansion is used only on the reference planes. Since these planes lie usually well outside the ion cores, it is expected from the same reasoning as in the theory of pseudopotentials that the plane-wave expansion is relatively rapidly convergent. Jennings and McRae (ref.5) have found that in usual conditions all the evanescent waves (with imaginary Γ_p) can be neglected. This may be seen easily from the fact that the evanescent waves decay usually to a negligible amount before it arrives from the level of atom centers to the level of the reference plane.

It should be noted that the number of plane waves is much smaller than in the usual three-dimensional plane-wave expansion. Jennings and McRae (ref.5) found, for example, that for the energy of 1 Ry above the zero level between the muffin-tin atoms a 18x18 matrix is satisfactory for a good accuracy of results. It means that only 9 waves in the +z and 9 in the -z directions are included.

The whole calculation is divided into two main steps. One is the calculation of the matrix elements of \underline{Q} using the partial-wave expansion; the other is the diagonalization of the matrix \underline{Q} between the plane-wave amplitudes. Thus, the present method is something intermediate between the partial-wave expansion and the plane-wave expansion. Jennings and McRae (ref.5) found that the calculation time is about half and half for the two steps. It appears that a reasonable compromise is reached between the reduction of the number of matrix elements and the corresponding increase of the labor of evaluation of matrix elements.

A large part of the calculation time of the matrix elements is spent for the evaluation of structure constants. This is, however, much less time-consuming than in the usual three-dimensional KKR method, because we have to do with lattice summations in two dimensions. The application of Ewald's summation method has revealed to be very effective. It has been found that 9 real and 9 reciprocal lattice points are usually completely satisfactory for giving reasonable accuracy.

Finally, a very important advantage of the present method should be mentioned. It has been found that some physical insight in the band structure is obtained by

inspection of the scattering property of the single atomic layers which appears in an intermediate stage of the calculation. Jenning and McRae (ref.5) found for W that two lowest gaps of the fully symmetric bands along (100) line appears as though they may be attributed to the two peaks which appear in the reflection coefficient of a (100) atomic layer in normal direction. These peaks are interpreted as resonances occurring in the layer (ref.12). Thus, at least this particular case is concerned, a band gap can be interpreted as caused by a resonance scattering by a single atomic layer, and not by a Bragg reflection by the set of atomic layers.

The author wishes to thank Dr.E.G.McRae for supplying materials on which this article is wholly based .

References

- (1) Kambe, K., Z.Naturforschg. 22a, 322 (1967)
- (2) Kambe, K., Z.Naturforschg. 23a, 1280 (1968)
- (3) McRae, E.G., Surface Sci. 11, 479 (1968)
- (4) Marcus, P.M. and D.W.Jepsen, Phys.Rev.Letters 20, 925 (1968)
- (5) Jennings, P.J. and E.G.McRae, Surface Sci. in press
- (6) Mattheiss, L.F., Phys.Rev. 139, A 1893 (1965)
- (7) Jennings, P.J., private communication
- (8) Williams, A., private communication
- (9) Jepsen, D.W. and P.M.Marcus, this volume
- (10) Gaunt, J.A., Phil.Trans.Roy.Soc.London A 228, 151 (1929)
- (11) Faddeeva, V.N. and N.M.Terent'ev, Tables of Values
of the Function
 $w(z) = \exp\{-z^2\} [1 + (2i/\sqrt{\pi}) \int_0^z \exp\{t^2\} dt]$
for Complex Argument, Pergamon Press, N.Y.1961
- (12) Kambe, K., Surface Sci. 20, 213 (1970)

THE PROPAGATION MATRIX METHOD FOR THE BAND PROBLEM WITH A PLANE
BOUNDARY

D. W. Jepsen and P. M. Marcus

IBM Thomas J. Watson Research Center

Yorktown Heights, New York

ABSTRACT

Solution of electronic problems involving plane surfaces on crystals requires solution of the band problem for real energy E in complex k space, and superposition of the generalized Bloch functions at the surface. A compact and general formulation of the problem of finding these Bloch functions and matching them across a plane makes use of a numerical matrix, the propagation matrix P , obtained from the Schrodinger equation. The eigenvectors of P are just the desired Bloch functions, and the eigenvalues give all k_{\perp} values at given E , k_{\parallel} (component parallel to the surface). Thus once P is found, the band problem is reduced to an ordinary eigenvalue problem; the bands can be followed along any line in k space parallel to k_{\perp} by varying k_{\parallel} ; the potential may be complex (to describe inelastic scattering). A procedure for generating P by integration of a matrix equation has the advantage that a general anisotropic potential can be used, but the disadvantage of a Fourier expansion parallel to the surface plane which does not hold well near the nucleus; hence it applies best for potentials that are weak or have a small number of Fourier coefficients. By generating P for a single layer by a two-dimensional version of KKR, this limitation is avoided for muffin-tin potentials.

I. INTRODUCTION

There is a class of band problems with boundary conditions that, in contrast to the impurity problem, is easily soluble, although more difficult than solution of the band problem alone.

This class of problems introduces boundary conditions on parallel plane surfaces and includes realistic physical models of such problems as low energy electron diffraction (LEED), surface states and tunneling. We describe here a general attack on this class of problems derived from studies of LEED but applicable more generally. We discuss the definition, generation and properties of the propagation matrix P , which contains the solution of the band problem, and can be used for a compact formulation of the surface matching problem. We also comment on the differences and advantages of this approach as a solution of the ordinary band problem. The solution of such surface-dependent problems is a multistep procedure. One must determine the band structure, obtain both the periodic and "exponential" or evanescent solutions of the Schrodinger equation of the solid, and carry out a matching with vacuum waves at the surface (Heine 1963, 1964). We have found that a method for determining the band structure and wave functions somewhat different from the usual techniques is uniquely useful and suitable to this problem. This new technique obtains the quantities needed in the final matching in a useful form almost directly, whereas other methods involve a larger amount of intermediate calculation. In particular it is much better suited to the calculation of evanescent waves than any other procedure we know of. Furthermore, examination of what is involved in this new procedure indicates that with suitable elaborations it may be more efficient than the standard KKR, and APW methods in ordinary band structure calculation because it obtains all the bands at a given energy in one matrix diagonalization. Hence, it seems certain that nothing is lost in power and efficiency by working this way on the surface problem.

The essence of the method is to formulate a suitable representation of an arbitrary wave function of energy E and of its normal derivative on a plane through the crystal which, typically, passes between the atoms and is parallel to the surface of interest. This representation should approximate the wave function and its derivative well using a suitable linear combination of basis functions with a finite number of coefficients. Next we determine the operator which translates a wave function from this plane to another equivalent plane one lattice spacing further from the surface. This operator P gives us the wave function and its normal derivative on this new plane from the values on the old plane, as these new values are uniquely determined by the Schrodinger equation once the energy is specified. In fact, P gives a linear transformation of the coefficients in the chosen representation and can hence be represented by a numerical matrix. We must solve the Schrodinger equation to obtain this matrix; several procedures will be discussed later. We shall denote both the operator and the matrix representing it to a finite approximation by P . An eigenvector of this matrix with eigenvalue λ gives us the value and normal (z) derivative of a wave function which is merely

multiplied by λ when translated one lattice spacing L_z into the crystal normal to the surface. If $|\lambda| = 1$ so that $\lambda = e^{ik_z L_z}$ with real k_z , then we have the translation property defining a Bloch wave. If $|\lambda| \neq 1$, the wave is an evanescent wave - necessary in matching boundary conditions at a surface but not a bulk solution of the band problem. The Bloch translational properties in the x and y directions are easily built into the representation as we shall see. Normally the wave functions used as a basis for this matrix are restricted to those with fixed values of the Bloch parameters k_x and k_y . Thus after fixing E , k_x and k_y , we obtain an unsymmetrical matrix, each of whose eigenvalues (this is an ordinary eigenvalue problem) provides a real or complex k_z . By varying the energy with k_x and k_y fixed, we trace out the complete band structure along a line through the Brillouin zone parallel to the z axis, including all of the real lines connecting the band edges in complex k_z space. We can also obtain constant energy surfaces by fixing E and varying k_x or k_y .

To further illustrate the method, we assume that we have a semi-infinite crystal bounded at a plane like the one we have chosen by either vacuum or a somewhat distorted or otherwise modified surface structure. To the extent that the representation we have chosen is complete, we obtain the values and normal derivatives of all of the Bloch and evanescent waves needed for matching. We need not calculate any properties of these functions off this plane. If there is a separate surface layer, a translation operator P_s for the surface layer can be calculated by the same methods and used to achieve the matching across this layer from bulk to vacuum.

II. GENERAL METHOD

These procedures will now be illustrated in more detail by means of the calculation of the scattering properties of a {100} face of an orthorhombic crystal. This calculation solves the low energy electron diffraction problem. The procedure will also be related to the calculation of surface states and to the calculation of transmission probabilities in tunneling between crystal media.

Because of the two-dimensional periodicity of the whole system of crystal plus vacuum in planes parallel to the surface (the x , y plane), the wave function which gives the solution to this problem has the property

$$\psi(\underline{r} + \underline{a}) = e^{i\vec{k}_0 \cdot \underline{a}} \psi(\underline{r}), \quad (1)$$

where a is any lattice vector of the crystal in the x , y plane

and k_o is the component in this plane of the \mathbf{k} vector of an incoming wave. Equation (1) follows from a two-dimensional form of Bloch's theorem. Hence $e^{ik_o \cdot \mathbf{r}} \psi(\mathbf{r})$ is periodic in x and y and can be expanded in a Fourier series. We shall use this Fourier series and the corresponding Fourier series for the z derivative (into the crystal) as our representation for the treatment. It can actually be written for any plane and we shall keep this generality by making the coefficients, $\psi_{\mathbf{k}}$, functions of z

$$\Psi(\underline{r}) = \sum_{\mathbf{K}} \psi_{\mathbf{K}}(z) e^{i(\underline{k}_o + \underline{K}) \cdot \underline{P}} . \quad (2)$$

Here the sum is over the two-dimensional reciprocal lattice vectors \mathbf{K} of the periodicity of the surface net, (the three-dimensional reciprocal lattice vectors are not needed in this method), and $\underline{P} = (x, y)$.

If the crystal is treated as made up of planes of scattering centers with free space between, as in the Darwin and Ewald theories of diffraction (James 1950), this representation is very good between the planes of scatterers. In this case the coefficients $\psi_{\mathbf{k}}(z)$ have the form

$$\psi_{\mathbf{K}}(z) = \alpha_{\mathbf{K}}^+ e^{i k_z(K) z} + \alpha_{\mathbf{K}}^- e^{-i k_z(K) z} \quad (3)$$

where

$$k_z(K) = [E - (K_x + k_{ox})^2 - (K_y + k_{oy})^2]^{1/2} . \quad (4)$$

For large K , k_z becomes an imaginary quantity of large magnitude and the corresponding term in the wave function increases strongly with either z or $-z$ and has many oscillations in the (x, y) plane. Since the true wave function does not have this sort of behavior between the atom layers, the coefficients of such terms are small. If the interaction with the atoms is weak, so that the wave function remains smooth near the centers of the atoms which do the scattering, this expansion can be used at all z . In this case the P matrix can be obtained simply by direct integration of the Schrodinger equation using the representation of the wave function we have chosen. Using Bohr radii and Rydbergs as units, the Schrodinger equation is

$$-\nabla^2 \Psi + V \Psi = E \Psi . \quad (5)$$

In diffraction and scattering problems, it is convenient to represent inelastic scattering by an imaginary contribution to the potential V (Slater 1937). Substituting in ψ as given by Eq. (1) and solving for the coefficients we obtain

$$\frac{d^2}{dz^2} \Psi_k(z) = -[k_z(k)]^2 \Psi_k(z) + \sum_{k'} V_{k-k'} \Psi_{k'}(z) \quad (6)$$

where

$$V_k(z) = \frac{1}{L_x L_y} \iint V(\underline{r}) e^{-i(k_x x + k_y y)} \quad (7)$$

This can be written as the first order matrix differential equation

$$\frac{d}{dz} \underline{\Psi} = A \underline{\Psi} \quad (8)$$

with

$$\underline{\Psi} = \begin{pmatrix} \Psi_{k_1} \\ \Psi_{k_2} \\ \vdots \\ \Psi'_{k_1} \\ \Psi'_{k_2} \\ \vdots \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 \\ -k_z^2 + V & 0 \end{pmatrix}, \quad (9)$$

where the vector Ψ contains as components the coefficients ψ_k followed by the normal derivatives $\partial \psi_k / \partial z \equiv \psi'_k$. The matrix A is made up of blocks corresponding to the ψ_k and ψ'_k blocks of C ; 0 is a zero matrix, 1 is a unit matrix, V is a matrix with components $V_{kk'} = V(z)_{k-k'}$, and $-k_z^2$ is a diagonal matrix corresponding to the first term on the right in Eq. (6). The importance of writing the equations in this form is that knowledge of the vector Ψ for any value of z determines it completely for any other z , simply by the integration of Eq. (8). It is also simple to truncate this infinite set of equations to a consistent finite set of dimension $2N$. We can obtain the general solution of these equations for arbitrary initial conditions by replacing the $2N$ dimensional vector Ψ by a $2Nx2N$ dimensional matrix P and using the initial condition that P be the unit matrix. Thus P is defined by

$$\frac{d}{dz} \underline{\underline{P}}(z, z_0) = \underline{\underline{A}} \underline{\underline{P}}(z, z_0) , \quad (10)$$

$$\underline{\underline{P}}(z_0, z_0) = \underline{\underline{I}} . \quad (11)$$

Then it can be easily seen that for arbitrary $\Psi(z_0)$

$$\underline{\Psi}(z) = \underline{\underline{P}}(z, z_0) \underline{\Psi}(z_0) \quad (12)$$

satisfies Eq. (8) and becomes equal to the given $\Psi(z_0)$ when z is equal to z_0 . The matrix P translates the vector Ψ . Furthermore, one finds that

$$\underline{\underline{P}}(z_2, z_1) \underline{\underline{P}}(z_1, z_0) = \underline{\underline{P}}(z_2, z_0) \quad (13)$$

where the ordinary matrix product is assumed. A variety of other interesting properties of P can be demonstrated, but they will not be used here. We shall only mention that the equations obtained by truncation continue to conserve electrons in the sense that the total flux of electrons in all of the components is guaranteed to be the same for every z . This has important consequences.

The finite set of equations obtained by truncation can be integrated numerically in matrix form quite easily to obtain $P(z_0, z)$ for any z_0 and z .

Up until this point the treatment may be used for a potential periodic in x and y but with an arbitrary variation in the z direction. Thus the procedure can be used to match wave functions across surface layers on solids provided that they have some sort of periodicity in the plane of the surface.

III. BLOCH WAVES

We shall now consider waves deeper in the crystal where the potential is also assumed to be periodic in the z direction. In this case the most interesting P matrix starts at some z_0 and translates Ψ a distance equal to one period of the potential in the z direction. A right eigenvector Ψ_λ of this P matrix is by definition merely multiplied by an eigenvalue λ when the matrix P is applied to it. Since applying P translates Ψ_λ by one period, we find that translating Ψ_λ by one period merely multiplies by λ . If $|\lambda| = 1$ this is the Bloch property, otherwise we are studying an evanescent wave. Using the coefficients in Ψ_λ in Eq. (2) gives us the corresponding function of x and y in the z_0 plane which is seen to have the usual Bloch properties in the x and y directions. This function can be extended to other z planes by using Eq. (12) and Eq. (2), and it is found that all of the Bloch-like properties are preserved in the full extension to all of three-dimensional space.

Figure 1 shows the extended band structure obtained by plotting the k_z values obtained from the λ 's by $\lambda = e^{ik_z L_z}$ as a function of energy. This was obtained in the course of a LEED calculation for a model of the (110) face of Al using an adaption of the potential which will be given in Eq. (22) with $Z = 3$. Here k_y was taken to be zero but k_x was given by $k_x = \sqrt{E} \sin 6^\circ$, since LEED experiments are not usually done at normal incidence. Hence the curve traversed through k space is near but not on the $\Gamma - X$ line. The dashed lines inside the rectangle give the real parts of complex k_z values along the "real line," The dashed lines outside the rectangle show the magnitudes of pure imaginary k_z 's to show how the bands are linked together. Similar diagrams have been given by Pendry and Forstmann (Pendry and Forstmann 1970). The calculation was done by numerical integration keeping 16 Fourier components.

The right eigenvectors and the left eigenvectors can be grouped together to give the pair of matrices which diagonalize P .

$$\underline{\underline{B}}^{-1} \underline{\underline{P}} \underline{\underline{B}} = \underline{\underline{\lambda}} \quad (14)$$

The matrix B is made up of the right eigenvectors of P and λ is a diagonal matrix made up of the eigenvalues which we have already discussed. It is convenient to choose the arrangement of these eigenvectors so those with eigenvalues with absolute value less than 1 and those corresponding to Bloch waves carrying flux into the crystal lie in the top half of the λ matrix and those with modulus greater than 1 or corresponding to Bloch waves moving back toward

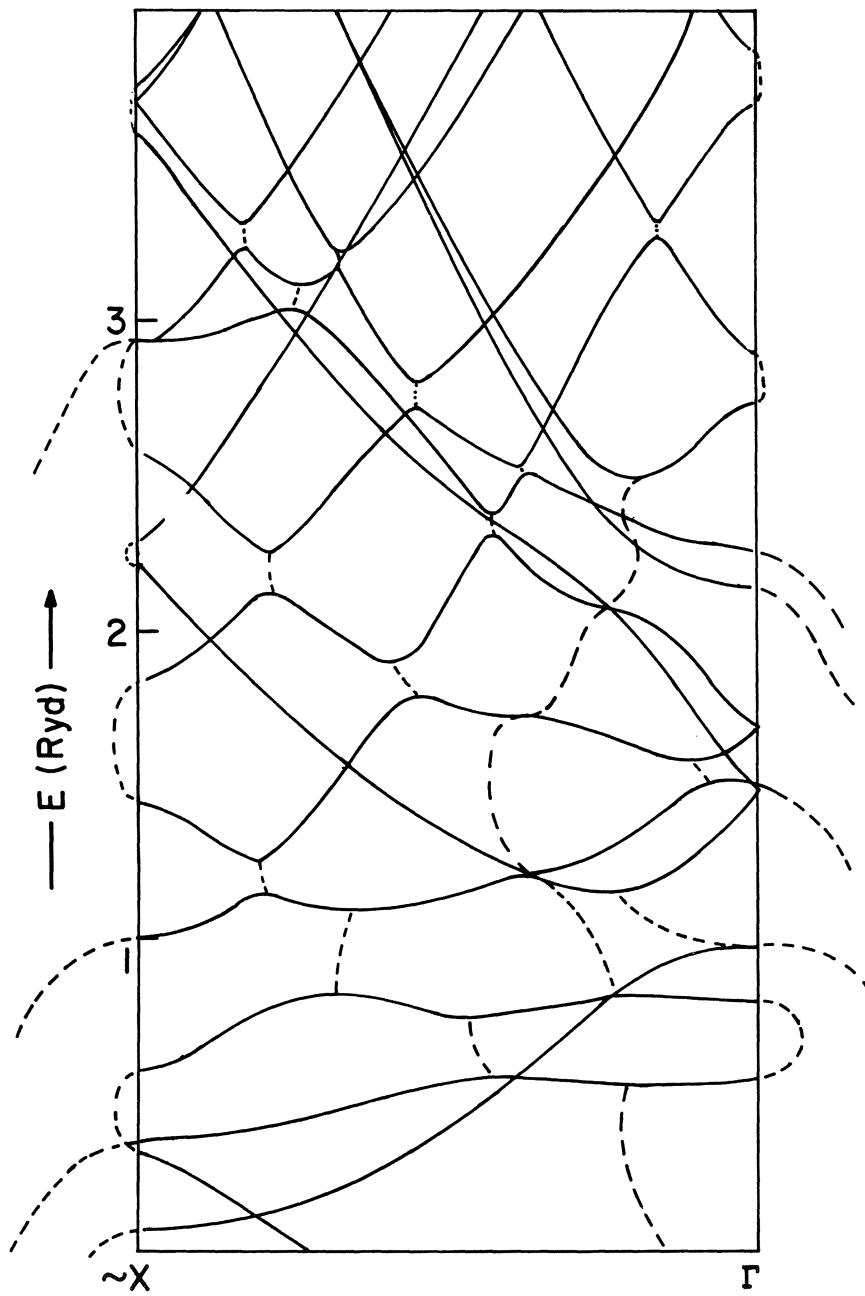


Figure 1. Calculated Extended Band Structure. See text.

the surface lie in the lower half. Note that a Ψ eigenvector contains the information required to calculate the sign of the z component of flux. The equation

$$\underline{\Phi} = \underline{B}^{-1} \underline{\Psi} \quad (15)$$

in which Ψ is an arbitrary vector of initial values, resolves this vector into a set of components on the Bloch-like waves found by the diagonalization procedure. Each of these generalized Bloch waves then goes into a multiple of itself as we translate by lattice vectors into the crystal. For the scattering problem we are considering, it is clear that no Bloch waves coming back from deep in the crystal, and no evanescent waves which increase as we go into the crystal, can be involved in the matching. This gives the condition

$$\underline{B}^{-1} \underline{\Psi} = \begin{pmatrix} X \\ 0 \end{pmatrix} \quad (16)$$

where the upper half X , of the Φ vector on the right-hand side is unknown, but the lower half must be zero if Ψ is to represent the solution for scattering by a semi-infinite crystal. We shall show that this is all that is needed to express the intensities of the beams diffracted back from the surface in terms of the intensity of the incoming electron beam.

IV. MATCHING

In the preceding work we have specified the values of the solution of the Schrodinger equation on a plane by the components of the vector Ψ . The equations

$$\begin{aligned} \Psi_k(z) &= \alpha_k^+ e^{ik_z(K)z} + \alpha_k^- e^{-ik_z(K)z} \\ \Psi'_k(z) &= ik_z(K) [\alpha_k^+ e^{ik_z(K)z} - \alpha_k^- e^{-ik_z(K)z}] \end{aligned} \quad (17)$$

$$\alpha_k^+ = \frac{1}{2} [\Psi_k(z) + (ik_z) \Psi'_k(z)] e^{-ik_z z}$$

$$\underline{\alpha}_k^- = \frac{1}{2} [\Psi_k(z) - (\frac{1}{\kappa_z}) \Psi'_k(z)] e^{+ik_z z}$$

tell us that if Ψ represents the wave function inside a crystal with a surface at z , then each pair (ψ_k, ψ'_k) of components of Ψ uniquely matches to two plane waves in vacuum at this surface with amplitudes α_k^+ and α_k^- moving toward positive and toward negative z , respectively. This means that instead of specifying the values of the various ψ_k and ψ'_k on a plane we could specify the amplitudes α_k^+ and α_k^- which would match these values if there were an interface with vacuum at this z . For large K the k_z 's which occur will be imaginary, but the same procedure goes through with the coefficients of the terms with positively and negatively imaginary k_z 's grouped into the α^+ and α^- vectors, respectively. The transformation to this new representation can be expressed by a simple matrix F

$$\underline{\Psi} = \underline{F} \underline{\alpha},$$

$$\underline{F} = \begin{pmatrix} 1 & 1 \\ i\chi & -i\chi \end{pmatrix} \begin{pmatrix} \underline{\epsilon}_z & 0 \\ 0 & \underline{\epsilon}_z^{-1} \end{pmatrix}, \quad \underline{\alpha} = \begin{pmatrix} \alpha^+(k_1) \\ \alpha^+(k_2) \\ \vdots \\ \alpha^-(k_1) \\ \alpha^-(k_2) \\ \vdots \end{pmatrix},$$

$$\underline{\chi} = \begin{pmatrix} \kappa_z(k_1) & & 0 \\ & \kappa_z(k_2) & \\ 0 & & \ddots & \\ & & & \kappa_z(k_n) \end{pmatrix}, \quad \underline{\epsilon}_z = e^{i\chi z},$$

$$\underline{F}^{-1} = \frac{1}{2} \begin{pmatrix} \underline{\epsilon}_z^{-1} & 0 \\ 0 & \underline{\epsilon}_z \end{pmatrix} \begin{pmatrix} 1 & -i\chi^{-1} \\ 1 & +i\chi^{-1} \end{pmatrix}.$$

(18)

The combination of matrices $B^{-1}F$ projects an arbitrary combination of plane waves outside the crystal into the components it produces on the generalized Bloch waves inside the crystal. If

this matrix is split into four blocks:

$$\underline{B}^T \underline{F} \begin{pmatrix} \alpha^+ \\ \alpha^- \end{pmatrix} \equiv \begin{pmatrix} J & L \\ M & N \end{pmatrix} \begin{pmatrix} \alpha^+ \\ \alpha^- \end{pmatrix} = \begin{pmatrix} X \\ 0 \end{pmatrix} \quad (19)$$

where the last equality expresses Eq. (16), this implies

$$\underline{N} \underline{\alpha}^- = -\underline{M} \underline{\alpha}^+ \quad (20)$$

which expresses the reflected waves α^- in terms of the incident wave given by α^+ .

The solid has a surface state for the given values of E , k_x and k_y when the equations we have developed have a solution even with no incoming wave. This gives the homogeneous equation

$$\underline{N} \underline{\alpha}^- = 0 \quad (21)$$

which must be satisfied for a surface state to exist. Surface states may be determined by varying E , k_x or k_y until the determinant of N is zero.

The calculation of the transmission coefficient in tunneling involves a procedure very similar to what we have described except the plane waves in vacuum are replaced by the Bloch waves of a new P matrix for the material on the initial side of the tunneling barrier.

V. CALCULATION OF THE P MATRIX BY SCATTERING THEORY

This procedure for carrying out the matching is independent of how the P matrix used is obtained. Obtaining the P matrix by direct numerical integration, as described in the general theory, is much more powerful than the usual techniques of diffraction theory because many beams can be handled. Unfortunately, however, it is

not accurate for strong atomic potentials because the expansion for the wave function we have given converges slowly in the neighborhood of an atomic core. This loss of accuracy is not surprising from what we know about plane wave expansions of wave functions, but these difficulties do appear to be limited to the (x,y) planes near the cores. This indicates that numerical integration can be used for layers of potential away from the cores, but some other procedure must be used for a thin layer containing each core. This conclusion is related to the discussion after Eq. (4), but can be considered more directly. Let us consider a simple potential with most of the physical properties of real potentials suitable for this type of calculation:

$$V(\underline{r}) = - \sum_j \frac{2Z}{|\underline{r} - \underline{R}_j|} e^{-\lambda |\underline{r} - \underline{R}_j|}. \quad (22)$$

The coefficients occurring in Eq. (7) for this potential are

$$V_K(z) = - \frac{4\pi Z}{L_x L_y} \frac{e^{-\beta(z-z_N)} + e^{-\beta(L_z - z + z_N)}}{\beta(1 - e^{-\beta z})}$$

where

$$\begin{aligned} \beta(k) &= [\lambda^2 + k_x^2 + k_y^2]^{1/2}, \\ 0 \leq (z - z_N) &\leq L_z, \end{aligned} \quad (23)$$

$L_x L_y$ is the area of the unit mesh in the x,y plane, and z_N is the z coordinate of the nucleus. The function is determined outside the given range by periodicity. We note that these coefficients go down exponentially with K when z is away from z_N but only as $1/K$ when $z = z_N$.

The problem of treating a strong potential can be solved through the use of pseudopotentials, but in the integration method only local pseudopotentials can be used because we do not know the wave function for values of z ahead of the point at which the integration is being done. To handle a general potential directly more powerful methods for obtaining the P matrix from the Schrodinger equation are needed, at least near the cores.

The P matrix can be transformed by means of F to a matrix Q, (the transfer matrix used by McRae 1968)

$$\underline{\underline{Q}} = \underline{\underline{F}}^{-1} \underline{\underline{P}} \underline{\underline{F}}. \quad (24)$$

Here it is convenient to choose the phase of the plane waves used for expansion to be zero on the plane of matching and the exponential waves equal to one. Then the ϵ matrix factors which occur in F and F^{-1} are unit matrices and Q is similar to P . It is easily seen that this transfer matrix connects the plane waves and exponential waves on one side of a slab with those on the other side by the relation

$$\begin{pmatrix} \alpha_{(+)}^+ \\ \alpha_{(+)}^- \end{pmatrix} = Q \begin{pmatrix} \alpha_{(-)}^+ \\ \alpha_{(-)}^- \end{pmatrix} \quad (25)$$

where the lower sign on the α symbols indicate that the plane waves on the right-hand side of the equation occur on the negative z side of the slab, while those on the left-hand side occur on the positive z side. Usually a physicist describes this type of situation by saying that the slab is scattering waves which are coming in from each side and producing reflected and transmitted waves corresponding to each incident wave. We note that $\alpha_{(-)}^+$ and $\alpha_{(-)}^-$ are vectors of incident waves while $\alpha_{(+)}^+$ and $\alpha_{(+)}^-$ contain outgoing waves. Thus we are led to consider the matrix S described by the equation

$$\begin{pmatrix} \alpha_{(+)}^+ \\ \alpha_{(-)}^- \end{pmatrix} = S \begin{pmatrix} \alpha_{(-)}^+ \\ \alpha_{(+)}^- \end{pmatrix} \quad (26)$$

which gives the outgoing waves in terms of the incoming waves. This is analogous to the usual S matrix used in scattering problems, but it is not unitary because attenuating waves and certain k_z factors in the non-attenuating waves make the conservation of flux more complicated in this case. Simple linear algebra enables us to express the Q matrix and the S matrix in terms of each other.

$$\begin{aligned} S &\equiv \begin{pmatrix} s_1 & s_2 \\ s_3 & s_4 \end{pmatrix}, \quad Q = \begin{pmatrix} s_1 - s_2 s_4^{-1} s_3 & s_2 s_4^{-1} \\ -s_4^{-1} s_3 & s_4^{-1} \end{pmatrix}, \\ Q &\equiv \begin{pmatrix} g_1 & g_2 \\ g_3 & g_4 \end{pmatrix}, \quad S = \begin{pmatrix} g_1 - g_2 g_4^{-1} g_3 & g_2 g_4^{-1} \\ -g_4^{-1} g_3 & g_4^{-1} \end{pmatrix}. \end{aligned} \quad (27)$$

From this we see that the solution of the most general scattering problem is contained in Q and that Q can be calculated from scattering theory, although in general exponentially attenuating waves must be included in the procedure. These attenuating waves are clearly required for matching when two slabs are placed near each other. Decaying exponentials from the two slabs which overlap can carry flux.

Thus the calculation of the P matrix can be considered as a problem in scattering theory to which a variety of scattering techniques might be applied. One procedure which seems promising is to calculate the scattering of atomic layers by pseudopotential methods. Pseudopotentials have been used for scattering problems with atoms and molecules, (e.g., Kestner, Jortner, Cohen, and Rice 1965).

VI. THE LAYER-KKR METHOD

A more powerful method for calculating the scattering of layers has been developed by Kambe (Kambe 1967a, 1968) and applied to the calculation of the Q matrix by McRae and Jennings (McRae and Jennings 1969, and to be published). This procedure is based on the KKR method in band theory but with the incident waves occurring in the equations as inhomogeneous terms. A simple derivation of the equations can be given based on the ideas of Korringa. This development is simpler than those in the literature and makes use of a relation between Bessel functions which appears to be new (Korringa 1947).

We wish to calculate the coefficients α^{scat} of the plane waves scattered out of each side of a plane containing a periodic array of scatterers, given a vector α^{in} of plane waves incident on the layer. The vectors $\alpha^{\text{out}} = \alpha^{\text{scat}} + \varepsilon_{2d}^{\text{II}} \alpha^{\text{in}}$ and α^{in} are the α vectors which occur in the left- and right-hand sides of Eq. (26) respectively. Here

$$\sum_{=2d}^{\text{II}} = \begin{pmatrix} \varepsilon_{2d} & 0 \\ 0 & \varepsilon_{2d} \end{pmatrix}$$

where ε is defined in Eq. (18) and $2d$ is the thickness of the layer.

The wave function around a typical scatterer in a plane of non-overlapping scatterers has the form

$$\Psi(\underline{r}) = \sum_L A_L J_L(\underline{r}) + \sum_L B_L Y_L(\underline{r}) \quad (28)$$

where

$$\mathcal{L} = (\ell, m)$$

$$J_L(z) \equiv i^\ell Y_\ell^m(z) j_\ell(kz)$$

and

$$H_L(z) \equiv i^{\ell+1} Y_\ell^m(z) h_\ell(kz). \quad (29)$$

Here the Y_ℓ^m 's are taken to be real spherical harmonics and j_ℓ and h_ℓ are standing wave and outgoing wave spherical Bessel functions respectively (Morse and Feshbach 1953, II, Chap. 11). The imaginary factors are introduced here to simplify later formulas. A small imaginary part to k will improve convergence in (28) and later formulas. The scattering properties of a spherical scatterer are completely characterized by the relation

$$B_L = \frac{e^{2i\delta_\ell}}{2i} - \frac{1}{A_L}, \quad (30)$$

containing the phase shifts δ_ℓ , which tells how the outgoing waves H_L from the center are determined from the incident waves J_L which would be present in this part of space even if the scattering center were not there. The phase shifts of a spherical potential of finite extent can be determined by integration of the radial Schrodinger equation. By combining the scattering of spherical scatterers of finite extent, using the scattering theory to be developed, one obtains the scattering and band properties of a crystal with a muffin tin potential.

The coefficients A_L which describe the waves incident on this center are made up of the contributions F_L due to waves incident on the layer from outside plus the contributions P_L from waves scattered from other centers in the layer which, for simplicity, we shall assume to be identical to the one we are considering,

$$A_L = F_L + P_L. \quad (31)$$

In general the waves incident on the layer will have the form

$$\sum_k \alpha_k^{in} e^{i\vec{k}(k) \cdot \underline{z}} = \sum_k \alpha_k^{in} e^{i\vec{k}(k) \cdot \underline{z}_0} \cdot e^{i\vec{k}(k) \cdot (\underline{z} - \underline{z}_0)} \quad (32)$$

where \underline{r}_0 is the displacement of the center from the plane on which α^{in} is defined. Expanding the last exponential using the standard formula in the appendix, Eq. (A5), shows that F_L is given by

$$F_L = 4\pi \sum_k \alpha_k^{in} Y_L(\vec{k}(k)) e^{i\vec{k}(k) \cdot \underline{z}_0} \quad (33)$$

where the vector $k(K)$ has the form

$$\vec{k}_{\pm}(K) = (\vec{k}_{ox} + K_x, \vec{k}_{oy} + K_y, \pm [E - (\vec{k}_{ox} + K_x)^2 - (\vec{k}_{oy} + K_y)^2]^{1/2}) \quad (34)$$

with the + and - chosen for waves moving in the positive and negative z directions, respectively. If Y is defined as the usual polynomial in the elements of k , multiplied by $E^{-\vec{k}/2}$, these formulas can also be used when k_z is not real.

Using Eq. (1), the wave function expressed about another center in the layer displaced from this one by R_n must be given by

$$\Psi(\underline{z}) = e^{i\vec{k}_0 \cdot \underline{R}_n} \sum_L A_L J_L(\underline{z} - \underline{R}_n) + e^{i\vec{k}_0 \cdot \underline{R}_n} \sum_L B_L \mathcal{H}_L(\underline{z} - \underline{R}_n). \quad (35)$$

Using the important relation

$$\begin{aligned} \mathcal{H}_L(\underline{z} - \underline{R}) &= 4\pi \sum_{L_1 L_2} C_{LL_1 L_2} \mathcal{H}_{L_1}(-\underline{R}) J_{L_2}(\underline{z}), \quad |\underline{R}| > |\underline{z}|, \\ C_{LL_1 L_2} &\equiv \int Y_L(\underline{k}) Y_{L_1}(\underline{k}) Y_{L_2}(\underline{k}) d^2 \Omega_{\underline{k}} \end{aligned} \quad (36)$$

derived in the appendix, we see that the contributions to the P_L 's at the original site from the waves scattered from the other atoms in the layer are given by

$$\mathcal{P}_{L_2} = 4\pi \sum_{R_n \neq 0} e^{i k_o \cdot R_n} \sum_L \mathcal{B}_L \sum_{L_1} C_{L L_1 L_2} \mathcal{D}_{L_1}^o (-R_n), \quad (37)$$

Substituting (37) into (31), and this into (30) we obtain

$$\mathcal{B}_L = \frac{e^{i k_o \delta_L} - 1}{2i} \left[F_L - 4\pi \sum_{L_1, L_2} \mathcal{B}_{L_2} C_{L L_1 L_2} \mathcal{D}_{L_1}^o \right]$$

with

$$\mathcal{D}_{L_1}^o = - \sum_{R_n \neq 0} e^{i k_o \cdot R_n} \mathcal{D}_{L_1}^o (-R_n). \quad (38)$$

Equation (38) can be transformed to

$$\cot \delta_L \mathcal{B}_L = F_L - \frac{4\pi}{k} \sum_{L_1, L_2} \mathcal{B}_{L_2} C_{L L_1 L_2} \mathcal{D}_{L_1}, \quad (39)$$

with

$$\mathcal{D}_{L_1}^o = k \left[\mathcal{D}_{L_1}^o - i(4\pi)^{-\frac{1}{2}} S_{L_1} \right] \quad (40)$$

by adding and subtracting iB_L in the bracket, transferring the subtracted term to the left-hand side, and using

$$C_{o L_1 L_2} = (4\pi)^{-\frac{1}{2}} S_{L_1 L_2} \quad (41)$$

on the added term. This can be solved for the B_L 's in terms of F_L 's to give the outgoing waves from one center in terms of the incoming waves.

Using Eq. (A6) of the appendix the D_L^o 's and the D_L 's can be evaluated by using the generating function relation

$$\begin{aligned} G(\underline{z}) &\equiv \sum_{R_n} \frac{e^{ik|\underline{z}-\underline{R}_n|+i\kappa_0 R_n}}{4\pi|\underline{z}-\underline{R}_n|} \\ &= -k \sum_{R_n \neq 0} \sum_L e^{i\kappa_0 R_n} \mathcal{H}_L(-R_n) J_L(\underline{z}) - \frac{e^{ik|\underline{z}|}}{4\pi|\underline{z}|} \end{aligned} \quad (42)$$

$$G(\underline{z}) = k \sum_L D_L^o J_L(\underline{z}) - \frac{\cos kr}{4\pi r} - i \frac{\sin kr}{4\pi r}$$

$$= \sum_L D_L J_L(\underline{z}) - \frac{\cos kr}{4\pi r} \quad (43)$$

Since $G(r)$ is the KKR Green's function for this planar problem, we see that the D_L 's are the two-dimensional analogues of the structure factor constants defined in KKR theory by this equation (Ham and Segall 1961, Eq. 2.18, Kambe 1967a, Eq. 28). Kambe has shown how they can be calculated by the usual Ewald procedures (Kambe 1967a, 1968, in this Proceedings).

VII. THE KKR EXPRESSION FOR SCATTERED WAVES

The Green's function $G(r)$ can also be expressed in the representation used in Eq. (2). The result is (Kambe 1967b, Eq. (34))

$$G(\underline{z}'-\underline{z}) = \sum_K \frac{e^{i\kappa_0(K)|\underline{z}'-\underline{z}|+i(\kappa_0+\underline{k}) \cdot (\underline{p}'-\underline{p})}}{2iL_x L_y \kappa_0(K)} \quad (44)$$

This can be used to find the waves scattered out from a layer con-

taining the plane of scatterers. Considering each spherical scatterer to be at the center of a cell between layer faces at $z=d$ and $-d$, take r' on the cell face at $z=d$ and r within the inscribed sphere; then $z < z'$ and $|r| < |r' - R_n|$. The absolute value sign can be removed and (44) becomes (replacing r by $-r$)

$$G(\underline{k} + \underline{k}) = \sum_K g_K e^{i(\underline{k} + \underline{k}_0) \cdot \underline{r}'} e^{i\underline{k}_+ \cdot \underline{r}} \\ g_K = \frac{e^{i\underline{k}_3(K)d}}{2iL_x L_y k_3(K)} \quad (45)$$

where \underline{k}_+ is defined in (34). Then using (A5) to expand $e^{i\underline{k}_+ \cdot \underline{r}}$

$$G(\underline{k}' + \underline{k}) = \sum_K g_K \sum_L 4\pi Y_L(\underline{k}_+) j_L(\underline{k}) e^{i(\underline{k} + \underline{k}_0) \cdot \underline{r}'} \quad (46)$$

Now $G(r' + r)$ can also be expanded in spherical waves in the form

$$G(\underline{k}' + \underline{k}) = -k \sum_L j_L(\underline{k}) \sum_{R_n} e^{i\underline{k}_0 \cdot \underline{R}_n} \mathcal{Y}_L(\underline{k}' - \underline{R}_n) \quad (47)$$

which is derived like (42) under the conditions on r and r' above. Consider the total scattered wave at r' in the forms

$$\Psi^{\text{scat}}(\underline{r}') \equiv \sum_L B_L \Psi_L^{\text{scat}}(\underline{r}') \equiv \sum_L B_L \sum_K \Psi_{LK}^{\text{scat+}} e^{i(\underline{k} + \underline{k}_0) \cdot \underline{r}'} \\ \equiv \sum_K \alpha_K^{\text{scat+}} e^{i\underline{k}_+ \cdot \underline{r}'} \quad (48)$$

Since $\Psi^{\text{scatt}}(r')$ is made up of the $B_L H_L(r' - \underline{R}_n)$ terms in (35)

summed over all R_n including $R_n = 0$, i.e., the sum of all the scattered waves by scatterers in the layer,

$$\psi_L^{\text{scat}}(\underline{r}') = \sum_{R_n} e^{i \frac{k_0}{\lambda} R_n} \mathcal{H}_L(\underline{r}' - \underline{R}_n). \quad (49)$$

Then, comparing the coefficients of $J_L(r)$ in (46) and (47), from (48) and (49)

$$\psi_{LK}^{\text{scat}+} = \frac{2\pi i}{k k_3(K) L_x L_y} e^{i \frac{k_3(K)d}{\lambda}} Y_L(k_+). \quad (50)$$

Similarly on the $z' = -d$ surface we have

$$\psi_{LK}^{\text{scat}-} = \frac{2\pi i}{k k_3(K) L_x L_y} e^{i \frac{k_3(K)d}{\lambda}} Y_L(k_-). \quad (51)$$

These functions form a basis for the waves scattered out of the layer. The Fourier coefficients of the complete scattered wave field is given by

$$\alpha_K^{\text{scat}} = \frac{2\pi i}{k k_3(K) L_x L_y} \sum_L B_L Y_L(k). \quad (52)$$

Solving Eq. (39) for B_L and substituting this into (52) with F_L defined by (33), we obtain a complete expression for the scattering by a layer:

$$\begin{aligned} \alpha^{\text{scat}} &= -\frac{8\pi^2 i}{L_x L_y} K^{\frac{1}{2}} \sum_{L=1}^{\infty} Y_L [K^{-1} - \frac{S}{L}]^{-1} \sum_{m=1}^{\infty} C_m^L \alpha_m \\ S &= \sum_{L=2}^{\infty} -\frac{8\pi^2 i}{L_x L_y} K^{\frac{1}{2}} \sum_{L=1}^{\infty} Y_L [K^{-1} - \frac{S}{L}]^{-1} Y_L \sum_{m=1}^{\infty} C_m^L \end{aligned} \quad (53)$$

where

$$\mathcal{G}' = \left\| \mathcal{G}_{L_1 L_2} \right\|, \quad \mathcal{G}_{L_1 L_2} = 4\pi \sum_{L_3} C_{L_1 L_2 L_3} \mathcal{D}_{L_3},$$

$$\mathcal{K}^{-1} = \left\| -k \cot \delta_L S_{L_1 L_2} \right\|,$$

$$\underline{\underline{Y}} = \left\| Y_{k L} \right\|, \quad Y_{k L} = Y_L(k),$$

$\underline{\underline{Y}}$ is $\underline{\underline{Y}}$ transpose

$$\underline{\underline{X}}^{II-1} = \begin{pmatrix} \underline{\underline{X}}^{-1} & 0 \\ 0 & \underline{\underline{X}}^{-1} \end{pmatrix}, \quad \underline{\underline{\epsilon}}_d^{\text{II}} = \begin{pmatrix} e^{i\chi d} & 0 \\ 0 & e^{i\chi d} \end{pmatrix},$$

and χ is defined in Eq. (18). We have used the fact that r_0 is $(0, 0, \pm d)$ with respect to origins in the two surfaces of the layer. The notation here has been chosen to make contact with some treatments of multiple scattering and KKR methods in band theory in the literature. The matrix K^{-1} is the inverse of the K matrix discussed by Ziman (Ziman 1965) and others. The matrix G' is the two-dimensional analogue of the G' matrix of Williams (Williams 1970) and called B in Ham and Segall. These formulas also have close analogues in the work of others (e.g., Beeby 1968, Eq. 15). Various extensions and improvements of the KKR procedure discussed by these and other authors should also be applicable to the procedures we have discussed.

A typical column of the Q matrix gives the waves in and out of the positive z side of a layer with only one wave of unit amplitude going in or out of the negative z side and all other waves on this side constrained to be zero. This characterization follows from Eq. (25). Such a result can be obtained by a special calculation of the waves scattered by the layer in which extra incident waves moving in the negative z direction are introduced to cancel the scattered waves on the negative z side that are forbidden by the constraint. The choice of these other incident waves clearly depends on the multiple scattering inside the layer. Equation (52) with k taken as k_- gives the amplitude of a typical wave to be canceled. The F_L terms required to do this are

$$F_L^C = -\frac{8\pi^2 i}{k k_3(k) L_x L_y} \sum_{L_2} B_{L_2} Y(L_2) Y(L_2)$$

Summing this expression over all of the waves which we wish to constrain to be zero gives an expression to be put into Eq. (39) as part of F; the assumed wave of unit amplitude makes up the rest of F. This F^c contribution can be combined conveniently with the terms involving the D_L 's in this expression.

Solving this equation for the B_L 's, we can again get the waves scattered out of the layer from Eq. (52). On the low z side of the layer we know that the scattered waves will all be canceled by waves incident from the other side and only the unit amplitude wave (which may be incident from either side) will be seen. On the high z side there will be the scattered waves corresponding to $k+$ in Eq. (52) and also the incident waves, including those that are minus the $k-$ waves scattered to the other side. Thus we obtain a formula for the Q matrix:

$$\underline{Q} = \begin{pmatrix} e^{2ikz} & 0 \\ 0 & e^{-2ikz} \end{pmatrix} - \frac{8\pi^2 i}{L_x L_y} \begin{pmatrix} k' e^{ikz} & 0 \\ 0 & -k' e^{-ikz} \end{pmatrix} Y [K' - k' - \frac{Q}{k}]^{-1} Y \begin{pmatrix} e^{ikz} & 0 \\ 0 & e^{-ikz} \end{pmatrix}$$

$$\underline{\mathcal{L}} = \|\underline{\mathcal{L}}_{L_1 L_2}\|, \quad \underline{\mathcal{L}}_{L_1 L_2} = \frac{8\pi^2 i}{L_x L_y} \sum_{L_3} C_{L_1 L_2 L_3} \sum_{k-} \frac{1}{k_g(k-)} Y_{L_3}(k-)$$
(54)

This result can also be obtained directly by substituting Eq. (53) into the expression for Q in terms of S in Eq. (27) and performing algebraic manipulations.

The sum over $k-$ in $\underline{\mathcal{L}}$ should be carried out over those scattered waves which exponentially decrease to non-negligible values at the surface of the layer.

VIII. CONCLUSIONS

The general procedure described here constitutes a new approach to the band problem which focusses on its relation to the scattering properties of the single-atom-thick layers from which it can be built up. It utilizes an unusual expansion in plane waves on the interstitial plane between layers (all the plane waves have the same energy and reduced component of k parallel to the plane), and leads to a new type of secular equation for the band problem. In addition the construction from layers is a particularly natural and useful procedure in problems involving plane interfaces.

The formulation emphasizes a matrix approach which has

substantial conceptual and computational advantages. We have given explicit formulas for the transfer matrix Q and propagation matrix P , whose eigenvalues give the band structure, in terms of the scattering matrix of a typical layer of the crystal. By using the Korringa approach to KKR theory, applied to the two-dimensional case, together with appropriate expansion formulas in Gaunt coefficients, formulas have been obtained for P and Q in terms of phase shifts and structure constants, similar to, but more explicit, than previous work by McRae and Jennings and Kambe.

The calculation of band structures by this layer-KKR procedure has several advantages over the usual KKR band calculation. The sums in the structure constants are over a plane of atoms only and more than one third of them vanish by the reflection symmetry in the plane. A procedure developed for a mesh of parallelograms will handle any monatomic plane and from such planes the P or Q matrices of even the most complex crystal structures can be obtained by multiplying together the P or Q matrices of the planes that make up a repeating layer. Moreover this process does not increase the order of the matrices which must be handled. The order of the matrices used depends on the number of plane waves needed to describe the wave function between the layers, and is empirically found to be small. The matrix to be diagonalized is presently $2N \times 2N$ complex unsymmetric, but it can probably be simplified. The number $2N$ of plane wave components necessary is probably smaller than the $(l_{\max}+1)^2$ which is needed in the usual three-dimensional KKR method when more than two or three phase shifts are needed. An additional economy is that layer matrices could be saved from one calculation to another in which the same plane of atoms occurs; in particular results for compounds could be built up from layers calculated for pure materials.

Finally we note that inelastic scattering may also be introduced into this layer-KKR procedure as a uniform negative imaginary part in the potential, which then affects the formalism in two ways: the phase shifts are complex, and the propagation between muffin tins involves a complex energy parameter. It is also clear that procedures developed to extend the usual KKR method beyond spherical muffin tin potentials can also be used here.

Appendix

We shall assume that real spherical harmonics have been chosen so that the spherical harmonic addition theorem

$$(2\ell+1) P_\ell(\cos \theta_{ab}) = 4\pi \sum_m Y_\ell(a) Y_\ell(b) \quad (\text{A1})$$

and the orthonormality relation

$$\int Y_\ell(a) Y_\ell(b) d\Omega_a = S_{\ell\ell} \quad (\text{A2})$$

are satisfied. Then in our notation, the well-known formulas (Morse and Feshbach 1953, II, p. 1574)

$$\mathcal{E}^{ik\vec{r} \cdot \vec{r}} = \sum_\ell (2\ell+1) i^\ell P_\ell(\cos \theta) j_\ell(kr) \quad (\text{A3})$$

and

$$\frac{\mathcal{E}^{ik|\underline{a}-\underline{b}|}}{ik|\underline{a}-\underline{b}|} = \sum_\ell (2\ell+1) P_\ell(\cos \theta_{ab}) h_\ell(k|a|) j_\ell(k|b|) \quad (\text{A4})$$

$|a| > |b|$

become

$$\mathcal{E}^{ik\cdot r} = 4\pi \sum_\ell Y_\ell(\underline{k}) J_\ell(\underline{r}) \quad (\text{A5})$$

and

$$\frac{\mathcal{E}^{ik|\underline{a}-\underline{b}|}}{ik|\underline{a}-\underline{b}|} = \sum_\ell (-1)^\ell Y_\ell(\underline{b}) J_\ell(\underline{a}) \quad |b| > |a|$$

or

$$\frac{e^{ik|\underline{a} + \underline{b}|}}{4\pi k|\underline{a} + \underline{b}|} = \sum_L C_L(\underline{b}) J_L(\underline{a}) \quad |\underline{b}| > |\underline{a}| \quad (A6)$$

by using the relation

$$Y_\ell^m(-\underline{b}) = (-1)^\ell Y_\ell^m(\underline{b}) \quad (A7)$$

which should also hold for the spherical harmonics, and changing \underline{b} to $-\underline{b}$. Equations (A5) and (A2) give the integral representation

$$J_{L_2}(\underline{a} + \underline{b}) = \frac{1}{4\pi} \int e^{i\underline{k} \cdot (\underline{a} + \underline{b})} d\Omega_{\underline{k}} = \frac{1}{4\pi} \int e^{i\underline{k} \cdot \underline{a}} e^{i\underline{k} \cdot \underline{b}} d\Omega_{\underline{k}} \quad (A8)$$

Substituting (A5) for each exponential in the far right side, we obtain

$$J_{L_2}(\underline{a} + \underline{b}) = 4\pi \sum_{L_1, L_2} C_{L_1, L_2} J_{L_1}(\underline{a}) J_{L_2}(\underline{b}) \quad (A9)$$

with

$$C_{L_1, L_2} = \int Y_{L_1}(\underline{k}) Y_{L_2}(\underline{k}) Y_{L_2}(\underline{k}) d\Omega_{\underline{k}} \quad (A10)$$

Equation (A6) gives

$$\frac{e^{i\omega|\underline{a} + \underline{b} + \underline{c}|}}{4\pi k|\underline{a} + \underline{b} + \underline{c}|} = \sum_{L_2} H_{L_2}(\underline{c}) J_{L_2}(\underline{a} + \underline{b}) \quad |\underline{a} + \underline{b}| < |\underline{c}| \quad (\text{A11})$$

and also that the same quantity is equal to

$$= \sum_L H_L(\underline{b} + \underline{c}) J_L(\underline{a}) \quad |\underline{b} + \underline{c}| > |\underline{a}| \quad (\text{A12})$$

Substituting (A9) into (A11) and equating to (A12), we have

$$\sum_L H_L(\underline{b} + \underline{c}) J_L(\underline{a}) = \quad (\text{A13})$$

$$4\pi \sum_{L_2} \sum_{L_1} \sum_L H_{L_2}(\underline{c}) C_{LL_1L_2} J_{L_1}(\underline{b}) J_L(\underline{a})$$

If $|\underline{c}| > |\underline{b}|$ is satisfied, a small sphere can be found such that for \underline{a} extending from the origin to any point on this sphere $|\underline{b} - \underline{c}| > |\underline{a}|$ and $|\underline{b} - \underline{a}| < |\underline{c}|$ are both satisfied. This enables to equate coefficients of $J_L(\underline{a})$ on each side of (A13) to obtain

$$H_L(\underline{b} + \underline{c}) = 4\pi \sum_{L_1 L_2} H_{L_2}(\underline{c}) C_{LL_1L_2} J_{L_1}(\underline{b}) \quad |\underline{c}| > |\underline{b}| \quad (36)$$

Acknowledgments

The authors are much indebted to F. P. Jona for his participation in the development of propagation matrix theory and to A. R. Williams for his assistance in formulation of KKR theory for a layer.

References

- Beeby, J. L., 1968, "The Diffraction of Low Energy Electrons by Crystals," *J. Phys. C.*, Proc. Phys. Soc. 1, 82-87.
- Ham, F. S., and Segall, B., 1961, "Energy Bands in Periodic Lattices - Green's Function Method," *Phys. Rev.* 124, 1786-96.
- Heine, V., 1963, "On the General Theory of Surface States and Scattering of Electrons in Solids," *Proc. Phys. Soc.* 81, 300-10.
- 1964, "Some Theory About Surface States," *Surf. Sci.* 2, 1-7.
- James, R. W., 1950, The Optical Principles of the Diffraction of X-Rays (G. Bell and Sons, Ltd., London), Chap. II.
- Kambe, K., 1967a, "Theory of Low-Energy Electron Diffraction I. Application of the Cellular Method to Monatomic Layers," *Z. Naturforsch.* 22a, 322-30.
- 1967b, "Theory of Electron Diffraction by Crystals I. Green's Function and Integral Equation," *ibid.*, 22a 422-31.
- 1968, "Theory of Low-Energy Electron Diffraction II. Cellular Method for Complex Monolayers and Multilayers," *ibid.*, 23a, 1280-94.
- This Proceedings, previous paper.
- Kestner, N. R., Jortner, J., Cohen, M. H., and Rice, S. A., 1965, "Low-Energy Elastic Scattering of Electrons and Positrons from Helium Atoms," *Phys. Rev.* 140, A56-66.
- Korringa, J., 1947, "On the Calculation of the Energy of a Bloch Wave in a Metal," *Physica* 13, 392-400.
- McRae, E. G., 1968, "Electron Diffraction at Crystal Surfaces I. Generalization of Darwin's Dynamical Theory," *Surf. Sci.* 11, 479-91.
- McRae, E. G., and Jennings, P. J., 1969, "Surface-State Resonances in Low-Energy Diffraction," *ibid.*, 345-48.
- 1970, "Electron Diffraction at Crystal Surfaces IV. Computation of LEED Intensities for Muffin-Tin Models with Application to Tungsten (001)," to be published in *Surf. Sci.*.
- Morse, P. M., and Feshbach, H., 1953, Methods of Theoretical Physics (McGraw-Hill Book Company, Inc., New York).
- Pendry, J. B., and Forstmann, F., 1970, "Complex Band Structure in the Presence of Bound States and Resonances," *J. Phys. C.*, Proc. Phys. Soc. 3, 59-69.
- Slater, J. C., 1937, "Damped Electron Waves in Crystals," *Phys. Rev.* 51, 840-46.
- Williams, A. R., 1970, "Non-Muffin-Tin Energy Bands for Silicon

by the Korringa-Kohn-Rostoker Method," ibid., B1, 3417-26.
Ziman, J. M., 1965, "The T Matrix, The K Matrix, d Bands, and
 ℓ -Dependent Pseudo-Potentials in the Theory of Metals,"
Proc. Phys. Soc. 86, 337-53.

6. SELF-CONSISTENCY, EXCHANGE, AND CORRELATION

THE SELF-CONSISTENT FIELD METHOD FOR CRYSTALS

J. C. Slater

University of Florida, Gainesville, Florida 32601

THE $X\alpha$ EXCHANGE

In a self-consistent-field calculation,¹ we start with assumed spin-orbitals u_i , which can correspond either to spin up or spin down. We let the occupation number of the i th spin-orbital be n_i , which would be unity for occupied spin-orbitals, zero for empty ones, in a Hartree-Fock scheme, but which in certain cases can be considered to be continuously variable between zero and unity. The charge density of spin up and spin down can be defined as

$$\rho = \rho^\uparrow + \rho^\downarrow = \sum(i\uparrow) n_i u_i^* u_i + \sum(i\downarrow) n_i u_i^* u_i \quad (1)$$

where $\sum(i\uparrow)$ indicates that we are summing only over the spin-orbitals with spin up, and similarly for $\sum(i\downarrow)$.

We assume a Hamiltonian of the form $\sum(i) f_i + \sum(\text{pairs } i,j) g_{ij}$, where f_i is the kinetic energy of the i th electron and its potential energy in the field of the nuclei, g_{ij} is the Coulomb potential energy of interaction of the i th and j th electrons, and where we have omitted the potential energy of interaction between the nuclei, a constant. We wish to find the expectation value of this Hamiltonian. We shall write this in the form

$$\begin{aligned} \langle Ex\alpha \rangle &= \sum(i) n_i f_i + \frac{1}{2} \int \rho(1) \rho(2) g_{12} dv_1 dv_2 \\ &\quad + \frac{1}{2} \int [\rho^\uparrow(1) U_{X\alpha}^\uparrow(1) + \rho^\downarrow(1) U_{X\alpha}^\downarrow(1)] dv_1 \end{aligned} \quad (2)$$

The first term is the kinetic energy plus the potential energy in the field of the nuclei. The summation in this term is over both

spins. The second term is the Coulomb energy of interaction between electrons, where again we include both spins. The third term is the exchange energy, for which we assume a statistical approximation, which we denote as the $X\alpha$ exchange. In this term we assume

$$U_{X\alpha}^{\uparrow}(1) = -\frac{3}{2}\alpha \left[6(3\rho^{\uparrow}/4\pi)^{1/3} \right] \quad (3)$$

with a similar formula for spin down. The Rydberg is used as unit of energy. If we set $\alpha = 2/3$, the exchange energy is the one suggested by the theory of the free-electron gas. The best choice of α is discussed in detail in Ref. 1; it is in the neighborhood of 0.7.

Once we have set up this expression for total energy, we can vary the spin-orbitals u_i , keeping the occupation numbers constant, to minimize the total energy, using undetermined multipliers to maintain the orthonormal properties of the u_i 's. When this is done, we obtain one-electron equations for the u_i 's. These equations are

$$\left[-\nabla_1^2 + V_C(1) + V_{X\alpha}^{\uparrow}(1) \right] u_i^{\uparrow}(1) = E_{iX\alpha}^{\uparrow} u_i^{\uparrow}(1) \quad (4)$$

with a similar equation for spin down, where $V_C(1)$ is the Coulomb potential acting on electron 1 from the total charge of all nuclei and electrons, computed classically, and

$$V_{X\alpha}^{\uparrow}(1) = \frac{2}{3} U_{X\alpha}^{\uparrow}(1) = -\alpha \left[6(3\rho^{\uparrow}/4\pi)^{1/3} \right] \quad (5)$$

The factor $3/2$ occurring in Eq. (3) has disappeared in the process of minimizing the energy, as pointed out by Gaspar,² Kohn, and Sham.³ In carrying out the self-consistent-field calculation, we then use the initial values of u_i to set up Eq. (4), through the appearance of u_i in V_C and $V_{X\alpha}$, and demand that the final values of u_i obtained by solving Eq. (4) should agree with the initial values. This is the Hartree iteration method, ordinarily carried out at present by the digital computer.

The iteration is to be carried through at constant occupation numbers, and these numbers would ordinarily be chosen to be those for the ground state of the system, assuming that we are interested in ground-state wave functions. We can next ask, what are the occupation numbers corresponding to the ground state? In some cases the answer is obvious, but in others, such as crystals containing 3d transition elements, it is not at all straightforward. The criterion which has ordinarily been used for determining the answer is to assume that in the ground state all orbitals with eigenvalues $E_{iX\alpha}$ below the Fermi energy will have an occupation number of unity, all those above will have an

occupation number of zero. We can now show that this is rigorously correct, if we use for the total energy $\langle EX\alpha \rangle$ the expression of Eq. (2). For it can be proved, as described in Ref. 1, that

$$E_{iX\alpha} = \frac{\partial \langle EX\alpha \rangle}{\partial n_i} \quad (6)$$

Thus if an infinitesimal amount of charge Δn is shifted from the i th to the j th orbital, the change of total energy is

$$\Delta \langle EX\alpha \rangle = (-E_{iX\alpha} + E_{jX\alpha}) \Delta n \quad (7)$$

which will be negative if the j th state is lower than the i th. We shall thus have achieved the lowest possible energy if all levels with $E_{iX\alpha}$ below a Fermi energy are filled, all those above are empty. We can have fractional occupation numbers in the ground state only if the energy levels precisely at the Fermi energy are degenerate, so that we can achieve the condition of electrical neutrality by having each of these levels partially filled, all with identical fractional occupation numbers. Some of the consequences of this situation have been examined by the author, Mann, Wilson, and Wood.⁴

It is to be observed that in the process of iteration to get the self-consistent solutions for the ground state, it will ordinarily be necessary to reevaluate the Fermi energy, and reassigned the occupation numbers of the spin-orbitals so that those below the Fermi energy will be occupied, those above will be empty, at each stage of the iteration. Since the heights of the various energy levels change greatly with occupation numbers, this process is a very important one, particularly in such cases as systems containing 3d transition atoms, as was emphasized in Ref. 4. In some cases of insulating or semiconducting crystals, where the occupation numbers are obvious from the start, this readjustment of occupation numbers will not be required.

It is of great importance to note that the eigenvalues for the $X\alpha$ method, given by Eq. (6), are not identical with those of the Hartree-Fock method, which are

$$E_{iHF} = \langle EHF \rangle_{n_i=1} - \langle EHF \rangle_{n_i=0} \quad (8)$$

or the difference between the total energy of the atom and of the ion in which the i th electron has been removed, where we use the same spin-orbitals for the ion as for the atom. In the next section we shall take up some of the consequences of this difference between the $X\alpha$ and the Hartree-Fock eigenvalues.

THE X α AND HARTREE-FOCK EIGENVALUES

The difference between the X α eigenvalue of Eq. (6), and the Hartree-Fock eigenvalue of Eq. (8), is most easily discussed if we regard the total energy $\langle E \rangle$ of either the X α or the HF method as a function of n_i , which we can express as a power series. The use of non-integral n_i 's in the Hartree-Fock method demands a special technique, which is discussed in Ref. 4, where it is defined as the Hyper-Hartree-Fock method. We have, up to quadratic terms,

$$\langle E \rangle = \langle E \rangle_0 + \frac{\partial \langle E \rangle}{\partial n_i} \Bigg|_0 (n_i - n_{i0}) + \frac{1}{2} \frac{\partial^2 \langle E \rangle}{\partial n_i^2} \Bigg|_0 (n_i - n_{i0})^2 + \dots \quad (9)$$

The author, and collaborators,^{1,4} have investigated the possibility of expanding the total energy of a system in a general power series expansion, of which Eq. (9) gives a special case. We find that though at least third power terms are required to get good accuracy, the quadratic terms are enough for a qualitative understanding of the phenomena. We then find that if $n_i = 0$, $n_{i0} = 1$, the case considered in Eq. (8), we have

$$\begin{aligned} \langle E \rangle_{n_i=0} &= \langle E \rangle_{n_i=1} - \frac{\partial \langle E \rangle}{\partial n_i} \Bigg|_0 + \frac{1}{2} \frac{\partial^2 \langle E \rangle}{\partial n_i^2} \Bigg|_0 \\ \langle E \rangle_{n_i=1} - \langle E \rangle_{n_i=0} &= \frac{\partial \langle E \rangle}{\partial n_i} \Bigg|_0 - \frac{1}{2} \frac{\partial^2 \langle E \rangle}{\partial n_i^2} \Bigg|_0 \end{aligned} \quad (10)$$

That is, the Hartree-Fock type of eigenvalue (the finite difference) equals the X α type of eigenvalue (the first derivative), minus half the second derivative $\partial^2 \langle E \rangle / \partial n_i^2 \Big|_0$. The difference between the eigenvalues of the two cases comes almost entirely from this second derivative, not from the small differences between $\langle E X\alpha \rangle$ and $\langle E HF \rangle$.

To get the meaning of this second derivative, let us differentiate Eq. (9) with respect to n_i . We have

$$\frac{\partial \langle E \rangle}{\partial n_i} = \frac{\partial \langle E \rangle}{\partial n_i} \Bigg|_0 + \frac{\partial^2 \langle E \rangle}{\partial n_i^2} \Bigg|_0 (n_i - n_{i0}) + \dots \quad (11)$$

If we use X α orbitals and the energy $\langle E X\alpha \rangle$, the quantity of Eq. (11) is the X α eigenvalue, as a function of the occupation

number n_i . If n_i is decreased by unity, the eigenvalue decreases by $\partial^2\langle E \rangle / \partial n_i^2 |_0$. Of course, each eigenvalue of an atomic system will be lowered as we go from the neutral system to a positive ion, and we see from this that the second derivative measures the lowering of eigenvalue per unit change of occupation number. This is a positive quantity, so that we see that the Hartree-Fock type of eigenvalue, as given in Eq. (10), is lower than the $X\alpha$ eigenvalue. This is what is observed in calculations by the two methods.

There are two physical effects which are contributing to the second derivative $\partial^2\langle E \rangle / \partial n_i^2$. First, in the Coulomb energy of Eq. (2), we include the Coulomb interaction of an orbital with itself. As n_i decreases by one unit, the electronic charge in the i th orbital decreases, so that the Coulomb repulsive energy decreases, leading to a decrease of the one-electron energy in Eq. (11), or to a positive quadratic term in the total energy of Eq. (9) whenever n_i differs from n_{i0} . If we take just this part of the second derivative into account, and use $\langle EHF \rangle$, the eigenvalue of Eq. (10) agrees with the Hartree-Fock eigenvalue.

Secondly, however, as n_i decreases by one unit, the other orbitals will relax, and this will be bound to have a second-order effect of decreasing the total energy. This appears in Eq. (9) as a negative contribution to the second derivative, partly cancelling the term arising from the change of occupation number, so that the second derivative is not as large as we should think if we included only the first term. It is for this reason that the Hartree-Fock eigenvalues are numerically too large to agree with experiment. If we compute the excitation energies by making separate calculations of the total energy of the atom and ion, and subtracting, we take account of the relaxation, and get a considerably better estimate of the excitation energy than by the Hartree-Fock eigenvalue. Koopmans' theorem, which does not take account of the relaxation, deals with the Hartree-Fock eigenvalue as we have been discussing it.

It is clear from this discussion that the $X\alpha$ and Hartree-Fock eigenvalues, which are known to differ significantly, really are referring to different things, and we must understand the difference if we wish to use the theory correctly. In solving for the ground state of a system, it is necessary to find the correct occupation numbers, and in the case of a metal this demands very careful attention to the eigenvalues. A shift from $X\alpha$ to Hartree-Fock eigenvalues, in an energy-band problem, can shift one energy band, such as the 3d band, with respect to another, such as the 4s and hence can change completely our assignment of occupation numbers in the ground state of the crystal. It was through the study of such cases that we were led to the need of using the $X\alpha$ method.⁴ The evidence is very strong that an

exchange like the $X\alpha$ exchange, with α in the neighborhood of 0.7, is required to bring the bands into the correct relation to each other to lead to the experimental occupation numbers. And for study of electronic specific heat, electrical conductivity, and such problems, relating to the Fermi surface, the $X\alpha$ method, with the eigenvalues determined by Eq. (6) leads precisely to the Fermi statistics, so that this is the method which should be used for such problems.

However, for optical absorption, we wish the energy difference between two localized atomic states whose occupation numbers differ by integers. Hence we need something like the Hartree-Fock eigenvalues, but corrected for relaxation as we have just discussed. We conclude from this that those who compare energy-band theory with experiment have been making an unwarranted assumption when they have postulated that a single set of energy bands can explain both the Fermi surface phenomena, and optical properties of crystals. We go on in the next section to a discussion of this problem.

OPTICAL EXCITATIONS OF CRYSTALS

We must distinguish between two different sorts of excitation which theoretically could occur in a crystal. First, we could have a single electron shifting from one Bloch state to another. Secondly, we could have an electron shifting from an atomic-type orbital localized on a single atom of the crystal, to an excited orbital on the same atom or in its neighborhood. The first type of excitation is what is concerned in thermal excitation and electrical conductivity, but the second is what is ordinarily found in optical excitation. The conventional theory of the exciton, going back to about 1930, sets up excited states in which an electron is removed from an inner level of a single atom, and is located in an excited state surrounding that atom, not unlike the excited state of an isolated atom. The continuous absorption at the band gap, in this type of theory, comes at the series limit of the discrete set of exciton levels.

Frenkel⁵ in 1931 treated the exciton from the point of view of localized excitation. Slater and Shockley⁶ in 1936 showed that if one starts out with Bloch waves, and treats the configuration interaction between the many states corresponding to the shift of one electron from an occupied to an empty Bloch state, there will ordinarily be at least one discrete excited state which will be split off from the continuum of excited states, corresponding to a lower excitation energy than the Bloch excitations, but carrying the larger part of the oscillator strength. This discrete state can be identified with the exciton state of Frenkel. Wannier⁷ in 1937 developed a more elegant method of handling the

mathematical description of the exciton, in terms of a hole in the valence band, and an electron in the conduction band, attracting hydrogen-like orbits around their center of mass. Wannier's picture led to a spectrum of discrete excited levels, terminated by a continuum which would represent the ordinary Bloch excitation. Koster and Slater,⁸ in the 1950's, handled the discrete states by a more powerful mathematical method than Wannier's, based on the use of Green's function, and showed that as the attraction between the hole and electron grew stronger, or the energy band narrower, there could be more and more discrete levels separated from the continuum. In the case of a wide enough energy band, it is possible to have no discrete level at all, whereas with a narrow enough band there will always be discrete levels.

These discrete levels have a definite energy for a definite value of K , the total vector sum of the wave vectors of all electrons in the crystal, which is a good quantum number for a perfect crystal. The dependence of the energy on K leads to a calculation of the mobility of the exciton. It was shown by Schultz⁹ and Nettel¹⁰, among others, that in case of a slowly moving exciton, there can be displacements of the atoms surrounding the excited atom, which have the effect of diminishing greatly the mobility, and the dependence of the exciton energy on K . It seems likely that this reduction of mobility is great enough so that in the case of x-ray excitation, the excitation is practically permanently tied to the atom from which the original electron was excited.

This series of papers, and many others, have made it reasonable to think of optical excitation of a crystal as arising from a localized excitation similar to that in an isolated atom. We must expect, then, that the excitation energy will be very similar to that in an atom, resembling the Hartree-Fock excitation of Eq. (8). We have seen from Eqs. (10) and (11) that this differs from the difference of $X\alpha$ eigenvalues by the term in $\frac{\partial^2 \langle E \rangle}{\partial n_i} \Big|_0$, measuring the change in excitation energy of an atom when an electron is removed from its i th shell. If however we are dealing with excitation from one of the continuous Bloch waves to another, then in a crystal containing N atoms, only $1/N$ of an electron will be shifted in each atom, and the second derivative, measuring the change in excitation energy when one electron is shifted, will be only $1/N$ times as large as for an isolated atom. In other words, it will be negligible, and the correct excitation energy to use in this case is the difference of the $E_{iX\alpha}$'s of the initial and final states. Thus the calculation of excitation energy of the discrete and of the extended excitations will be definitely different from each other, in a way which is not considered in the existing theories of the exciton.

THE TRANSITION STATE

There is an interesting and simple way in which we can consider the localized excitation. Let us first note from Eqs. (10) and (11) that if we compute the $X\alpha$ type of eigenvalue of an isolated atom, $\partial\langle E \rangle / \partial n_i$, not for the initial state, but for a state with $n_i - n_{i0} = -1/2$, which means a state in which half the electron has been removed from the atom, the $X\alpha$ eigenvalue $E_{iX\alpha}$ will then equal the Hartree-Fock eigenvalue of the i th state, from Eq. (10), computed for the ground state. It is shown in the second paper of Ref. 1 that we can prove an extension of this result. We consider the excitation of an atom from the i th to the j th states, and make a self-consistent calculation by the $X\alpha$ method for what we may call a transition state, in which the electron in question is half in the i th state, half in the j th state of the problem. Then we can show, as in Ref. 1, that the difference between the $X\alpha$ eigenvalues of the i th and j th states in this transition state gives a very accurate measure of the excitation energy of the atom from the i th to the j th levels. This calculation can be shown to take account very accurately of the relaxation of the orbitals in going from the one state to the other.

This concept of the transition state can be extended to the crystalline excitation. In such a case, if we are interested in optical excitation, the general line of discussion suggested in Refs. 5--10 indicates that the excitation will be localized. The transition state is then one in which a definite atom has an electron half in the i th, half in the j th orbital. This atom will act like a perturbed atom. By the techniques suggested in Ref. 8, one can discuss the energy levels of such an atom in a self-consistent manner. One will find, as we have indicated earlier, that there will be discrete energy levels below the continuum, and it is the transition state in which the i th discrete level is half empty, and the j th half occupied, which we solve self-consistently. Through this mechanism, we are led to self-consistent localized excitations, and to a technique for solving for the excitation energy by the $X\alpha$ method as applied to the transition state.

A study of excitations in crystals using this method has not yet been carried out. However, various techniques which are available at present would allow one to make the calculations. Thus, Parada and Pratt¹¹ have applied the $k \cdot p$ method to APW energy bands, and have carried through a detailed study of impurity levels in the PbTe crystal, using the methods of Ref. 8. These studies of impurity levels are mathematically analogous to the problems which would be met in the exciton problem. Johnson and Connolly¹² have studied methods of handling similar problems by the KKR method. It seems entirely practical to carry through

such treatments for the exciton problem, and it is clear from more detailed discussion that the results will give optical excitation energies which are closer to those of the isolated atom problem than to those which would come from the direct use of the $X\alpha$ eigenvalues for determining optical spectra.

There is one very interesting situation which will prove to have a bearing on the problem. For the x-ray excitation, we definitely have a discrete exciton level split off from the continuum. On the other hand, we have mentioned that for sufficiently broad bands, it can come about that no discrete level will be split off from the continuum. In such a case, the $X\alpha$ eigenvalues would directly determine the optical spectrum. This situation would very likely occur in many cases of optical excitation, but not necessarily in all. It should be a very interesting study, in the future, to investigate the intermediate situation between these cases, in which for some bands we might be able to use the $X\alpha$ energy levels for determining optical spectra, while in other cases there could be appreciable departures from this simple picture.

We point out finally that, as the writer showed¹³ in 1937, there is a close analogy between exciton theory and the theory of spin waves and magnetic excitations. There is enough progress already in the study of ferromagnetic and antiferromagnetic crystals by energy-band methods¹⁴ so that one can be quite confident that we are at last on the track of a quantitatively successful theory of these difficult problems. The extensive literature of superexchange, Heisenberg exchange integrals, and so on, as well as of crystal field theory, has suffered because it was not possible to make accurate calculation of the many parameters which entered the theory. Now with the type of development which we are foreseeing, we can hope that we are on the threshold of the time when such quantities can be derived from the type of theory which we have been discussing. Some discussion of such problems is given in Ref. 4, and in Wilson's work¹⁴ on anti-ferromagnetic crystals.

REFERENCES

¹ For more detailed discussions of the points brought up in the present note, see J. C. Slater and J. H. Wood, Statistical Exchange and the Total Energy of a Crystal, Preprint No. 191, Quantum Theory Project, University of Florida, Gainesville, Florida 32601, to be published in International Journal of Quantum Chemistry, Symposium No. 4, January, 1971. See also J. C. Slater and J. H. Wood, Optical Transitions and the Statistical Exchange in Crystals, Phys. Rev. (to be published), and J. C. Slater, Transition Probab-

bilities and Fractional Occupation Numbers in Atoms, to be published in volume honoring E. U. Condon.

- 2 R. Gaspar, "Über eine Approximation des Hartree-Focksschen Potentials durch eine universelle Potentialfunktion, *Acta Phys. Hung.* 3, 263 (1954).
- 3 W. Kohn and L. J. Sham, Self-Consistent Equations Including Exchange and Correlation Effects, *Phys. Rev.* 140, A1133 (1965).
- 4 J. C. Slater, J. B. Mann, T. M. Wilson, and J. H. Wood, Non-Integral Occupation Numbers in Transition Atoms, *Phys. Rev.* 184, 672 (1969).
- 5 J. Frenkel, On the Transformation of Light into Heat in Solids, *Phys. Rev.* 37, 17 (1931).
- 6 J. C. Slater and W. Shockley, Optical Absorption in the Alkali Halides, *Phys. Rev.* 50, 705 (1936). See also J. C. Slater, Excited Energy Levels of Insulating Crystals, *Trans. Faraday Soc.* 34, 828 (1938).
- 7 G. H. Wannier, Structure of Electronic Excitation Levels in Insulating Crystals, *Phys. Rev.* 52, 191 (1937).
- 8 G. F. Koster and J. C. Slater, Wave Functions for Impurity Levels, *Phys. Rev.* 94, 1932; 95, 1167 (1954). G. F. Koster, Theory of Scattering in Solids, *Phys. Rev.* 95, 1436 (1954). G. F. Koster and J. C. Slater, Simplified Impurity Calculation, *Phys. Rev.* 96, 1208 (1954). See also J. C. Slater, Electrons in Perturbed Periodic Lattices, *Phys. Rev.* 76, 1592 (1949).
- 9 T. D. Schultz, Slow Electrons in Polar Crystals: Self-Energy, Mass, and Mobility, *Phys. Rev.* 116, 526 (1959).
- 10 S. J. Nettel, Many-Particle Theory of Impurity States in Polar Crystals, *Phys. Rev.* 128, 2573 (1962).
- 11 N. J. Parada and G. W. Pratt, Jr., New Model for Vacancy States in PbTe, *Phys. Rev. Letters* 22, 180 (1969).
- 12 K. H. Johnson, Scattering Model for the Bound Electronic States of an Impurity Complex in a Crystal, *Int. J. Quant. Chem.* 2S, 233 (1968). K. H. Johnson and J. W. D. Connolly, On Calculating the Localized Electronic States of Surface and Line Defects, *Phys. Letters* 28a, 291 (1968).

- 13 J. C. Slater, The Theory of Ferromagnetism: Lowest Energy Levels, Phys. Rev. 52, 198 (1937).
- 14 S. J. Cho, Spin-Polarized Electronic Energy-Band Structure in EuS, Phys. Rev. 157, 632 (1967).
- S. J. Cho, Antiferromagnetic Exchange Mechanism in Europium Chalcogenides, Phys. Letters 29A, 129 (1969).
- S. J. Cho, Spin-Polarized Energy Bands in Eu Chalcogenides by the APW Method, Phys. Rev. B (in press).
- J. W. D. Connolly, A Self-Consistent Calculation of the Energy Bands in Ferromagnetic Nickel, Int. J. Quant. Chem. 1S, 615 (1967).
- J. W. D. Connolly, Energy Bands in Ferromagnetic Nickel, Phys. Rev. 159, 415 (1967).
- J. W. D. Connolly, The Energy Band Structure of Magnetic Transition Metals, Int. J. Quant. Chem. 2S, 257 (1968).
- P. D. DeCicco and A. Kitz, Calculation of Neutron and X-Ray Scattering Amplitudes for bcc Iron, Phys. Rev. 162, 486 (1967).
- S. Wakoh and J. Yamashita, Band Structure of Ferromagnetic Iron by a Self-Consistent Procedure, J. Phys. Soc. Japan 21, 1712 (1966).
- T. M. Wilson, Spin-Polarized Energy Bands in the Antiferromagnetic MnO, Int. J. Quantum Chem. 2S, 269 (1968).
- T. M. Wilson, Spin-Polarized Energy-Band Structure of Antiferromagnetic MnO, J. Appl. Phys. 40, 1588 (1969).
- T. M. Wilson, A Study of the Electronic Structure of the First-Row Transition-Metal Compounds, Int. J. Quantum Chem. 3S, 757 (1970).
- T. M. Wilson, J. H. Wood, and J. C. Slater, Studies of the Statistical Exchange Approximation in the First Transition Row Atoms and Ions: The Mn^{+2} Ion, Phys. Rev. (submitted for publication).
- J. Yamashita, S. Wakoh, and S. Asano, Band Structure of Transition Metals Calculated by the Green's Function Method, "Quantum Theory of Atoms, Molecules, and the Solid State," P. O. Ljwdin, ed., Academic Press, New York, 1966, p. 497.
- J. Yamashita, S. Asano, and S. Wakoh, Band Theory of Metallic Antiferromagnetism, J. Appl. Phys. 39, No. 2, Pt. 2, 1274 (1968).
- J. Yamashita, H. Namba, and S. Asano, Stability of Non-Magnetic State in 3d Transition Metals, Prog. Theoret. Phys. 39, 1091 (1968).

APPROXIMATIONS OF THE EXCHANGE AND CORRELATION POTENTIALS*

L. J. Sham

University of California, San Diego

La Jolla, California 92037

ABSTRACT

A review is made of the approximate exchange and correlation potentials. A calculation of the gradient correction to the exchange potential is presented and the systematic expansion of the potentials in powers of the gradients of the density is critically examined. A related question concerning the validity of the local exchange raised by Overhauser is investigated here.

1. INTRODUCTION

There has been a number of numerical investigations of the various local exchange potentials. I shall not review the results here. Instead, I would like to review briefly the formal framework of the exchange and correlation potentials and to discuss the theoretical basis for the various approximations which have been made.

At the outset, I want to make a distinction between the exchange and correlation potential used for two different purposes: (1) for calculating the density distribution and hence, the total energy, and (2) for calculating the one-particle energies in a system with a large number of electrons. The potentials for these two purposes are related but not identical. This fact has important ramifications as we shall see.

2. SELF-CONSISTENT POTENTIAL FOR THE DENSITY DISTRIBUTION AND THE TOTAL ENERGY

Based on the work of Hohenberg and Kohn¹, Kohn and I² wrote the ground state energy in the form

$$E = \int dr v(r)n(r) + \frac{1}{2} \int dr \int dr' \frac{e^2}{|r-r'|} n(r)n(r') + T_I[n] + E_{xc}[n]. \quad (1)$$

where $T_I[n]$ is the kinetic energy of the independent Fermions which have the same density distribution $n(r)$. E_{xc} is, by definition, the exchange and correlation energy. By varying the density, we arrive at a set of one-particle Schrödinger equations with the self-consistent potential $\varphi(r) + V_{xc}(r)$; from which the density, and hence, the total energy can be determined. φ is the Hartree potential and

$$V_{xc}(r) = \delta E_{xc} / \delta n(r), \quad (2)$$

is the exchange and correlation potential. This is an exact result, showing how a many-body problem can be reduced to an essentially one-body problem. At this stage, we should be careful not to interpret the energy eigenvalues of the self-consistent potential as the one-particle energies.

For an approximate form of the exchange and correlation potential which is "computationally accessible," we follow the spirit of the Thomas-Fermi approximation, that is, we approximate the exchange and correlation effects of the electrons in a small neighborhood by those of the uniform system with the corresponding local density. This can be made systematic as a gradient expansion,³

$$E_{xc} = \int dr n \epsilon_{xc}(n) + \frac{1}{2} \int dr g_{xc}^{(2)}(n) (\nabla n)^2 + \dots \quad (3)$$

Keeping only the first term, we have

$$V_{xc}(r) = \mu_{xc}(n(r)), \quad (4)$$

where μ_{xc} denotes the exchange and correlation part of the chemical potential of the uniform system evaluated at the local density. The exchange part is just two thirds Slater exchange, sometimes referred to as the KSG exchange⁴, (or, more properly, the Dirac exchange⁵) and has been widely investigated in numerical studies. Since in solids, the correlation is just as important as the exchange and since, in this approximation, it is no more difficult to include, I hope to see more computations include this correlation approximation.

In the gradient expansion (3), if we keep the first two terms, we have the first gradient correction, the exchange part of which have been investigated by Herman, Van Dyke and Ortenburger.⁶ They have found a remarkable result. Now, the coefficient of the gradient term must have the following density dependence (see the Appendix):

$$g_x^{(2)}(n) = -\gamma_2 n^{-4/3} \quad (5)$$

where γ_2 is a constant. The constant, which Herman et.al use, is actually given by

$$\beta = -\gamma_2 (\pi/81)^{1/3}/e^2 \quad (6)$$

They obtain β in two ways: (1) in which they adjust β to minimize the total energy (the "optimized statistical" value) and (2) in which they adjust β to make the total energy equal to the Hartree-Fock energy (the "alternate statistical" value). For various atoms, β varies remarkably little from atom to atom. The values for β are 5×10^{-3} and 3.14×10^{-3} respectively.

However, it is possible to calculate γ_2 from first principles. (This has independently been done by I. B. Ortenburger.) We relegate the details to the Appendix. The coefficient γ_2 is related to the second order coefficient of the static density response function expanded in powers of the wave-vector q . In Column I of the Table, we give the calculated values and note that the value for β is about a factor of three to five too small to enable one to explain the results of Herman et. al in terms of rapid convergence of the gradient series.

TABLE I. COEFFICIENT OF THE GRADIENT TERM

	I	II	III
$-\frac{k_F}{F} \tilde{\chi}_1^{(2)} / \tilde{\chi}_1^{(0)}$	$5/72$	$1/36$	0
$-\gamma_2/e^2$	3.33×10^{-3}	4.77×10^{-3}	5.72×10^{-3}
β	1.13×10^{-3}	1.61×10^{-3}	1.94×10^{-3}

The Hartree-Fock static density response to first order in e^2 has been computed for a range of values of the wave-vector by Geldart and Taylor⁷. In column II, the coefficient gives a reasonable fit of the density response for wave-vectors ranging

from zero to about $k_F/2$. The value of β is increased only by 50%. In column III, we show that even if the part of density response of first order in e^2 were independent of q , the value of β is still below 2×10^{-3} .

A similar conclusion for the gradient correction to correlation in atomic systems was reached earlier by Ma and Brueckner⁸.

Kohn and I have proposed² to account for the effect of density variation by including the correction term,

$$E_{xc}[n] = \int dr n \epsilon_{xc}(n) - \frac{1}{4} \int dr \int dr' K_{xc}(r-r'; n(\frac{1}{2}r+\frac{1}{2}r')) \{n(r)-n(r')\}^2 \quad (7)$$

where K is minus the reciprocal of the irreducible density response function of the homogeneous system. (See the Appendix). This correction term includes a partial series of the gradient expansion and is also correct in the limit of small but possibly rapidly varying density deviation from the average. We can only proceed with numerical evaluation if we know K_{xc} , i.e. the density response function beyond the random phase approximation. There have been various approximations for it, and I have just learned that J. Woo has a numerical calculation of the density response⁹. It is now feasible and interesting to have this gradient correction tested numerically.

3. SELF-CONSISTENT POTENTIAL FOR THE ONE-PARTICLE ENERGIES

To include all the many-body effects, the one-electron (quasi-particle) energy is determined in principle from the one-particle Schrödinger equation with the self-energy $\Sigma(r, r'; E)$ as the effective potential. The complications of this potential are that it is (1) nonlocal (like the Fock exchange term), (2) energy dependent, and (3) non-Hermitian. The energies are determined¹⁰ by obtaining the eigenvalues $W_k(E)$ for each E and then solving $E = W_k(E)$.

We have shown¹⁰ that if we write

$$\Sigma(r, r'; E) = \varphi(r) \delta(r-r') + M(r, r'; E-\varphi(\frac{1}{2}r+\frac{1}{2}r')) , \quad (8)$$

then the exchange and correlation part M is short-ranged in two senses, (i) being confined to the region $|r-r'| \lesssim \text{Max}(1/k_F, 1/k_{FT})$ where k_F is the Fermi wave-vector and k_{FT} is the Thomas Fermi screening wave vector, (ii) having a short-ranged density dependence. Thus, we can approximate M by M_h of the homogeneous electron system with the local density. In particular, at the Fermi level, the potential reduces to $\mu_{xc}(n)$ which is the same as in

Eq. (4). Hence, the exchange and correlation potential discussed in the last section can be used to calculate the density and the total energy as well as the Fermi surface.

For energies in the neighborhood of the Fermi level, the energy dependence of M makes the procedure a little more complicated¹⁰. According to B. I. Lundqvist's calculation¹¹ for the homogeneous electron gas, the energy and momentum dependence of the self-energy is weak. Based on this, Hedin and S. Lundqvist¹² suggested a local potential which is 5/6 Slater's exchange. From the above considerations, it is clear that, if B. I. Lundqvist's result holds, a more logical approximation would be $\mu_{xc}(n)$.

It is of some interest to consider just the exchange term. Then, the mass operator M reduces to the usual exchange operator. The replacement by M_h of the homogeneous system at the local density gives the effective potential

$$M_h(k) = -\frac{2e^2 k}{\pi} F(k/k_F) = \epsilon_x(k) \quad (9)$$

where k and k_F are the local wave-vectors and

$$F(x) = \frac{1}{2} + \frac{1-x^2}{4x} \ln \left| \frac{1+x}{1-x} \right| \quad (10)$$

This momentum dependent potential was written down by us¹⁰ and used in calculating atomic energy levels by Liberman¹³. In Eq.(9) if we put $k = k_F$, we have the Dirac or KSG exchange and if we average over $k \neq k_F$, we have Slater's exchange.

4. OVERHAUSER'S OBJECTION

Recently, Overhauser¹⁴ raised a question on the validity of the local density approximation for the exchange operator. Consider the inhomogeneous electron system with the density

$$n(r) = n_0 [1 + 2\lambda \cos q \cdot r] \quad (11)$$

where the fractional deviation from the constant density, λ , is small and the spatial variation is slow, i.e. q small. Entirely within the Hartree-Fock approximation and to the lowest order in λ as well as e^2 (that is, to the lowest order in the electron-electron interaction), the exchange operator is

$$M(k + q, k) = \lambda n_0 \tilde{\Lambda}(k + q, k) / \tilde{\chi}(q) , \quad (12)$$

where $\tilde{\Lambda}$ and $\tilde{\chi}$ are given in the Appendix. Overhauser evaluated the exchange operator numerically and found that for $k = k_F$, as $q \rightarrow 0$,

$$M(k_F + q, k_F) \sim -\ln q \quad (13)$$

It is easy to show analytically that, for small q ,

$$\tilde{\Lambda}(k+q, k) = -\frac{\epsilon_x(k+q) - \epsilon_x(k)}{\epsilon_o(k+q) - \epsilon_o(k)} + \frac{2e^2 k_F}{\pi k^2} F(\frac{k}{k_F}) \quad (14)$$

where $\epsilon_x(k) = k^2/2$, and $\epsilon_x(k)$ is the exchange energy given by Eq. (9). The first term on the right contains the $\ln q$ divergence if we first put $k = k_F$. If we let $q \rightarrow 0$ first, then the first term becomes the exchange part of the electron velocity, which again diverges as $k \rightarrow k_F$ as is well known.

The relation between the divergence of the exchange operator and the divergence of the electron velocity at the Fermi level in the Hartree-Fock approximation can be exhibited more directly. By the Ward identities, it can be shown that¹⁰ as $q \rightarrow 0$,

$$\tilde{\Lambda}(k+q, k) \rightarrow -\partial M_h(k)/\partial \mu \quad (15)$$

and

$$\tilde{x}(q) \rightarrow \partial n/\partial \mu \quad (16)$$

where the chemical potential in the Hartree-Fock approximation is

$$\mu = \frac{1}{2}k_F^2 - e^2 k_F / \pi . \quad (17)$$

Hence, the exchange operator, by Eq. (12), is directly proportional to the electron velocity as $q \rightarrow 0$. Incidentally, by this method, it is not necessary to expand in powers of e^2 .

The method applies just as well if we include correlation. Kohn and L¹⁰ have considered just the problem Overhauser raised. In this case, the electron velocity is well behaved at the Fermi level and we have

$$M(k + q, k) \rightarrow \lambda n_o \partial M_h(k)/\partial n , \quad (18)$$

in complete agreement with the local density approximation.

The so-called Liberman exchange, Eq. (9), is most seriously affected by Overhauser's objection because Overhauser's result shows that it is at best an asymptotic approximation, with a rather divergent correction. If correlation is included, we saw above that the local density approximation is perfectly valid. Slater's exchange survives by the following argument. The average over k of $M_h(k)$ can alternatively be regarded as replacing the exchange density with a long oscillatory tail by a square hole¹⁵. Since the effect of correlation is to screen this tail, Slater's approximation has the effect of accounting for some correlation.

The exchange potential for the calculation of density and total energy, V_x given by Eq. (2), is totally unaffected by Overhauser's objection. It is incorrect to expect that $M(k+q, k)$ should reduce to V_x . Indeed for the system with the weak and slowly varying density, given in Eq. (11), it is easily seen from the results in the Appendix that the correction to the KSG exchange is finite. This instance demonstrates the importance of recognizing the distinction between the two exchange and correlation potentials.

APPENDIX. GRADIENT CORRECTION TO THE EXCHANGE ENERGY

Suppose that $v(r)$ is the weak and slowly varying potential which produces the density given by Eq.(11) and that $\varphi(r)$ is the total electrostatic (Hartree) potential. Then, the exchange operator and the density deviation from the average are of the form

$$M(k+q, k) = \tilde{\Lambda}(k+q, k) \varphi(q) , \quad (A1)$$

and

$$n(q) = \tilde{\chi}(q) \varphi(q) , \quad (A2)$$

where $\tilde{\Lambda}$ is called the irreducible vertex part and $\tilde{\chi}$ the irreducible density response.

In the Hartree-Fock approximation, it is straight forward to apply perturbation theory to find that $\tilde{\Lambda}$ is given by the integral equation,

$$\tilde{\Lambda}(k+q, k) = 1 - \int \frac{d^3 k}{(2\pi)^3} \theta(k'+q, k') u(k-k') \tilde{\Lambda}(k'+q, k') , \quad (A3)$$

$$\text{where, } u(q) = 4\pi e^2/q^2 , \quad (A4)$$

$$\text{and, } \theta(k'+q, k') = (f_{k'+q} - f_{k'})/[\epsilon(k'+q) - \epsilon(k')] . \quad (A5)$$

f_k is the occupation number of state k and $\epsilon(k)$ is the electron energy of the homogeneous system,

$$\epsilon(k) = \epsilon_o(k) + \epsilon_x(k) . \quad (A6)$$

The density response is given by

$$\tilde{\chi}(q) = 2 \int \frac{d^3 k}{(2\pi)^3} \theta(k+q, k) \tilde{\Lambda}(k+q, k) . \quad (A7)$$

The coefficient $g_x^{(2)}(q)$ of the gradient correction to the exchange term is given by the q^2 coefficient of $K_x(q)$ in powers of q , and

$$K_x(q) = -1/\tilde{\chi}(q) + 1/\tilde{\chi}_0(q). \quad (A8)$$

The suffix zero denotes no interaction between electrons, i.e. $e^2=0$. If we expand,

$$\tilde{\chi}(q) = \tilde{\chi}^{(0)} + \tilde{\chi}^{(2)} q^2 + \dots, \quad (A9)$$

then, $g_x^{(2)} = \tilde{\chi}^{(2)}/\{\tilde{\chi}^{(0)}\}^2 - \tilde{\chi}_0^{(2)}/\{\tilde{\chi}_0^{(0)}\}^2.$ (A10)

Now we go further in asserting that $g_x^{(2)}$ is determined if we know the coefficients of $\tilde{\chi}$ to the first order in e^2 . We know that the density dependence of $g_x^{(2)}$ must be $n^{-4/3}$ because, if we scale the whole system by multiplying the linear dimension with a factor s , the total exchange energy scales by a factor s . From Eqs.(A3--7), we see that $\tilde{\chi}(q)$ is k_F times a function of q/k_F and e^2/k_F only. Thus, the known density dependence of the coefficients of $\tilde{\chi}$ as deduced from relations such as Eq.(A10) means that the coefficients can be determined if we know them to the first order in e^2/k_F . Hence,

$$g_x^{(2)} = [\tilde{\chi}_0^{(2)}/\{\tilde{\chi}_0^{(0)}\}^2] \cdot [\tilde{\chi}_1^{(2)}/\tilde{\chi}_0^{(2)} - 2\tilde{\chi}_1^{(0)}/\tilde{\chi}_0^{(0)}], \quad (A11)$$

where the suffices denote the orders of e^2 and the superscripts denote the orders of q .

This simplifies the problem enormously as we can expand every quantity in the integral equation (A3) in powers of e^2 and solve by iteration. Thus, $\tilde{\chi}_0(q)$ is just the well known random phase approximation result:

$$\tilde{\chi}_0(q) = -\frac{k}{\pi} F(q/2k_F).$$

The density response of the first order in e^2 is

$$\begin{aligned} \tilde{\chi}_1(q) = & -2 \int \frac{d^3 k}{(2\pi)^3} \theta_0(k+q, k) \left[\int \frac{d^3 k'}{(2\pi)^3} \theta_0(k'+q, k') u(k-k') + \right. \\ & \left. + \{\varepsilon_x(k+q) - \varepsilon_x(k)\}/\{\varepsilon_0(k+q) - \varepsilon_0(k)\} \right] \end{aligned} \quad (A12)$$

Where $\theta_0(k+q, k)$ is given by Eq.(A5) with $\varepsilon_0(k)$ in place of $\varepsilon(k)$.

This is the expression which has been evaluated by Geldart and Taylor.⁷ However, we need the coefficients of the power series in q^2 . We have seen in Sec.4 that either term in the square bracket of Eq.(A12) contains a logarithmic singularity as $q \rightarrow 0$ and $k \rightarrow k_F$ and that the two singularities cancel out. Since the final result is a regular function of q for $q \ll k_F$, we can avoid unpleasantness of the singularities in intermediate steps by using

the Yukawa interaction

$$u(q) = 4\pi e^2 / (q^2 + \lambda^2) , \quad (A13)$$

and let λ tend to zero in the end. Incidentally, if we let λ be the Thomas-Fermi screening wave-vector, we have the model of statically screened exchange.

With the Yukawa interaction, the exchange energy of the electron becomes,

$$\begin{aligned} \epsilon_x(k) &= - \int \frac{d^3 k'}{(2\pi)^3} f_{k'} u(k-k') \\ &= - \frac{e^2 k_F}{\pi} \left[1 - \left(\frac{\lambda}{k_F} \right) \left\{ \tan^{-1} \left(\frac{k+k_F}{\lambda} \right) - \tan^{-1} \left(\frac{k-k_F}{\lambda} \right) \right\} \right. \\ &\quad \left. + \frac{(k_F^2 + \lambda^2 - k^2)}{4kk_F} \ln \left\{ \frac{(k+k_F)^2 + \lambda^2}{(k-k_F)^2 + \lambda^2} \right\} \right] . \end{aligned} \quad (A14)$$

It reduces to the expression (9) if we let $\lambda=0$. A related quantity which is important for later use is

$$P(k) \equiv \left(\frac{\partial}{\partial \mu_0} + \frac{\partial}{\partial \epsilon_0(k)} \right) \{ k \epsilon_x(k) \} . \quad (A15)$$

$$= - \frac{2e^2 k_F}{\pi k} + \frac{e^2 \lambda}{\pi k} \left\{ \tan^{-1} \left(\frac{k+k_F}{\lambda} \right) - \tan^{-1} \left(\frac{k-k_F}{\lambda} \right) \right\} . \quad (A16)$$

The expression is written out to exhibit the fact that, if we let $\lambda \rightarrow 0$ last, then,

$$P(k) \rightarrow - \frac{2e^2}{\pi} \left(\frac{k_F}{k} \right) , \quad (A17)$$

$$\text{but, } \frac{\partial P}{\partial k} \Big|_{k_F} \rightarrow e^2 / \pi k_F . \quad (A18)$$

A straight forward way to obtain the series in q^2 from (A12) is to regard the occupation number f_k as a function of energy, $f(\epsilon_0(k))$, and expand

$$\theta_0(k+q, k) = f'(\epsilon_0(k)) + \frac{1}{2} \{ \epsilon_0(k+q) - \epsilon_0(k) \} f''(\epsilon_0(k)) + \dots \quad (A19)$$

Then, the q independent term is

$$\tilde{x}_1^{(o)} = - \epsilon_x(k_F) / \pi^2 k_F^2 + 2 \frac{\partial}{\partial \mu_0} \int \frac{d^3 k}{(2\pi)^3} \epsilon_x(k) \frac{\partial f_k}{\partial \mu_0} \quad (A20)$$

$$= - e^2 / \pi^3 . \quad (A21)$$

having let $\lambda \rightarrow 0$ in the last step. We could, of course, have obtained this from the compressibility sum rule, Eq.(16).

The coefficient of the q^2 term is

$$\begin{aligned}\tilde{\chi}_1^{(2)} = & -\epsilon_x(k_F) / 12\pi^2 k_F^3 + \int \frac{d^3k}{(2\pi)^3} [\epsilon_x(k) \left\{ \frac{1}{3} f'''(\epsilon_o(k)) \right. \\ & \left. + \frac{1}{9} \epsilon_o(k) f^{IV}(\epsilon_o(k)) \right\} + \partial \epsilon_x(k) / \partial \mu_o \left\{ -\frac{1}{2} f''(\epsilon_o(k)) \right. \\ & \left. - \frac{1}{9} \epsilon_o(k) f'''(\epsilon_o(k)) \right\}]\end{aligned}\quad (A22)$$

The integrals on the right are regrouped to give a sum of three integrals:

$$\begin{aligned}I_1 &= -\frac{1}{2} \frac{\partial}{\partial \mu_o} \int \frac{d^3k}{(2\pi)^3} \epsilon_x(k) \frac{\partial^2 f}{\partial \mu_o^2} \\ &= -\left(1/4\pi^2\right) \int_0^\infty dk \frac{\partial f}{\partial \mu_o} \frac{\partial P}{\partial k} = -e^2/4\pi^3 k_F^2,\end{aligned}\quad (A23)$$

using Eqs.(A 15--18),

$$I_2 = \frac{1}{9} \frac{\partial}{\partial \mu_o} \int \frac{d^3k}{(2\pi)^3} \epsilon_x(k) \cdot \epsilon_o(k) \frac{\partial^3 f}{\partial \mu_o^3} = -\frac{4}{9} I_1, \quad (A24)$$

$$\text{and, } I_3 = \frac{1}{6} \int \frac{d^3k}{(2\pi)^3} \epsilon_x(k) \frac{\partial^3 f}{\partial \mu_o^3} = e^2/8\pi^3 k_F^2. \quad (A25)$$

$$\text{Hence, } \tilde{\chi}_1^{(2)} = 5e^2/72\pi^3 k_F^2, \quad (A26)$$

$$\text{and } \gamma_2 = -7e^2/216\pi (3\pi^2)^{1/3}, \quad (A27)$$

which is evaluated in Column I of Table I.

The calculation of γ_2 here is similar to the calculation of Ma and Brueckner for the corresponding term in the correlation energy.

I wish to thank Dr. S.K.Ma, Dr. F. Herman and Dr. I.B. Ortenburger for helpful conversations.

REFERENCES AND FOOTNOTES

- * Supported in part by the National Science Foundation Grant No. GP-14912.
- 1. P. Hohenberg and W. Kohn, Phys. Rev. 136, B864 (1964).
- 2. W. Kohn and L. J. Sham, Phys. Rev. 140, A1133 (1965).
- 3. Note the factor of a half in front of the gradient term which is absent in the definition used in Ref.1.
- 4. R. Gaspar, Acta Phys. Hung. 3, 263 (1954).
- 5. P. A. M. Dirac, Proc. Camb. Phil. Soc. 26, 376 (1930).
- 6. F. Herman, J. P. Van Dyke and I. B. Ortenburger, Phys. Rev.

- Letters 22, 807 (1969); Intern. J. Quan Chem. 3S, 827 (1969).
- 7. D. J. W. Geldart and R. Taylor, Can. J. Phys. 48, 155 (1970).
 - 8. S. K. Ma and K. A. Brueckner, Phys. Rev. 165, 18 (1968).
 - 9. See the following paper by J. Woo.
 - 10. L. J. Sham and W. Kohn, Phys. Rev. 145, 561 (1966).
 - 11. B. I. Lundqvist, Phys. Stat. Sol. 32, 273 (1969).
 - 12. L. Hedin and S. Lundqvist, Solid State Physics, 23, 1 (1969).
 - 13. D. Liberman, Phys. Rev. 171, 1 (1968).
 - 14. A. W. Overhauser, Bull. Am. Phys. Soc. 15, 344 (1970). I am grateful to Dr. Overhauser for a copy of his preprint.
 - 15. J. C. Slater, Phys. Rev. 81, 385 (1951).

SOME REMARKS ON EXCHANGE INHOMOGENEITY CORRECTIONS IN MANY-ELECTRON SYSTEMS

Irene B. Ortenburger and Frank Herman

IBM Corporation, Monterey & Cottle Roads, San Jose,
California 95114

In the preceding paper Sham pointed out that a number of people have by now calculated the first principles value of the second order exchange inhomogeneity parameter β . Ma and Ortenburger¹ find a value of 0.0011 (in the units of HVDO), Liberman² finds 0.0016, while the nomenclal work of Geldart and Taylor³ would seem to support the smaller value. In any case the theoretical value for β is considerably smaller than the values obtained empirically by Herman, Van Dyke, and Ortenburger^{4,5} from self-consistent calculations for a number of representative atomic systems. In the work of HVDO the optimum value of β is found to be about 0.005; the inclusion of exchange inhomogeneity corrections with this coefficient leads to a considerable improvement in the total energy: this is lowered about half of the distance between the best energy obtained ignoring these corrections. and the exact Hartree-Fock total energy.

On the other hand, if one requires the total energy of an atom, as given by the statistical field approximation including inhomogeneity terms, to be equal to the corresponding Hartree-Fock total energy, the value of β is found to be about 0.003. The question is: why are these empirical values for β larger than the first-principles value of about 0.001?

A natural answer to this question is that the empirical coefficient for the second order exchange inhomogeneity terms, as obtained by HVDO, accounts for higher-order terms as well as the explicit second order terms that appear in their formalism. This would be the case if the higher order terms had the same general form as the second order terms, so that the explicit second order

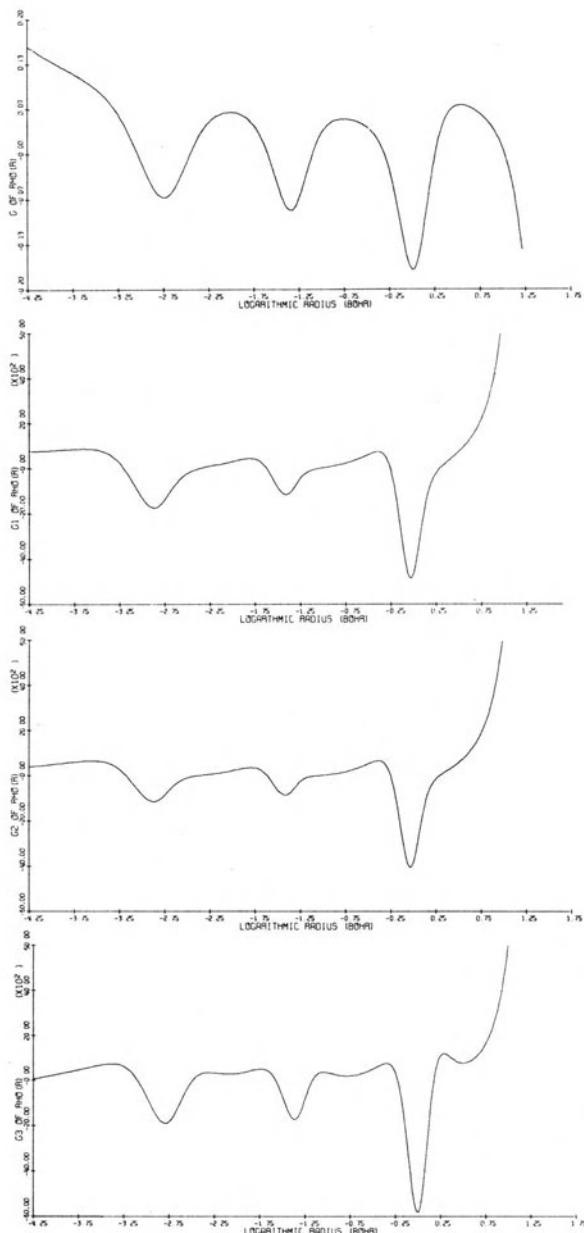


Figure 1. The second order exchange inhomogeneity correction $G(\rho)$ and the threefourth order terms $G_1(\rho)$, $G_2(\rho)$, and $G_3(\rho)$, plotted against $\log r$, in atomic krypton. Small- r and large- r convergence factors have not been used.

terms would be representing the higher order terms to all intents and purposes. For example, suppose that each of the three fourth-order terms has a magnitude roughly $1/2$ as large as the second order term. Then the second order term appearing alone with a coefficient $(1 + 3/2) = 5/2$ as large as its first-principles value would actually be taking the fourth as well as the second order terms into account, and similarly if still higher order terms are brought into play. For this argument to be meaningful, we must be able to show that the fourth-order terms (for example) are actually quite similar in form to the second order term.

The gradient expansion of the exchange contribution to the potential including fourth-order terms is given by

$$V_{x\alpha\beta\gamma}(\rho) = -6 \left(\frac{3}{8\pi} \rho \right)^{1/3} [\alpha + \beta G(\rho) + \gamma_1 G_1(\rho) + \gamma_2 G_2(\rho) + \gamma_3 G_3(\rho)] \quad (1)$$

Extending our earlier work on atomic krypton, we have calculated the form of the three fourth-order exchange inhomogeneity corrections. These are shown in Fig. 1, as is the second order term. It will be seen that the three fourth-order terms are nearly identical, and that they have the same overall character as the second order term although differing from it in the finer details. Their main effect will be to shift the maxima in $G(\rho)$ slightly to the right while leaving the zeros relatively unmoved.

We feel that this comparison lends some support to the above argument, and accounts for the fact that the empirical value for β is found to be larger than the first-principles value.

It will be interesting to see whether we can obtain still better solutions for atomic systems (still lower total energies) by setting $\alpha = 2/3$ (the Kohn-Sham value), $\beta = 0.001$ (the first principles value), and treating the three fourth-order parameters γ_1 , γ_2 , and γ_3 as variational parameters. It would also be interesting to have first-principles values for these fourth-order parameters.

Finally, we would like to make a brief remark concerning the value of numerical computations in the present context. Many theoretical papers have appeared in which exchange and correlation inhomogeneity effects are discussed and dismissed because of mathematical difficulties at the origin (in atomic systems). In spite of these theoretical objections, HVDO proceeded to incorporate the exchange inhomogeneity corrections into numerical atomic structure calculations, with the gratifying result that half of the discrepancy in total energy between the optimum $X\alpha$ solution and the exact Hartree-Fock solution was removed. This result could hardly have been anticipated from the purely theoretical papers. In effect,

the HVDO calculations can be regarded as an experiment which will hopefully stimulate further theoretical progress. We hope that future improvements in the treatment of inhomogeneity effects will reduce still further the discrepancy between optimized statistical exchange calculations and Hartree-Fock calculations. Of course, there still remains the problem of getting theoretical energies to agree with experiment, and this obviously requires the inclusion of relativistic and correlation effects. However, from our experience with exchange inhomogeneity corrections, we are encouraged to reopen the question of correlation inhomogeneity corrections, both theoretically and numerically.

REFERENCES

- [1] S. Ma and I.B. Ortenburger, unpublished.
- [2] D.A. Liberman, private communication.
- [3] D.J.W. Geldart and R. Taylor, Can. J. Phys. 48, 155 (1970).
- [4] F. Herman, J.P. Van Dyke, and I.B. Ortenburger, Phys. Rev. Letters 22, 807 (1969).
- [5] F. Herman, I.B. Ortenburger, and J.P. Van Dyke, Intern J. Quantum Chem. IIIS, 827 (1970).

DIELECTRIC FUNCTION OF UNIFORM ELECTRON GAS

J. W. F. Woo

IBM Thomas J. Watson Research Center
Yorktown Heights, New York

The dielectric function of the uniform electron gas at zero temperature has been calculated in the self-consistent Hartree-Fock (SHF) approximation.

The approximation used is equivalent to the following. The system considered is described by the Hartree-Fock Hamiltonian with the exchange interaction replaced by a Yukawa potential with a Thomas-Fermi screening length. Finding the linear response of this system to an external field of momentum q and frequency ω leads to the following expression for the dielectric function ϵ .

$$\begin{aligned}\epsilon(q, \omega) &= 1 - \frac{4\pi e^2}{q^2} L(q, \omega) \\ L(q, \omega) &= 2 \int \frac{d^3 k}{(2\pi)^3} \tilde{L}(k, q, \omega) \\ \tilde{L}(k, q, \omega) &= \frac{f(E_{k+q}) - f(E_k)}{E_{k+q} - E_k - \hbar\omega} \left\{ 1 - \int \frac{d^3 k'}{(2\pi)^3} \frac{4\pi e^2}{|k-k'|^2 + k_{FT}^2} \tilde{L}(k', q, \omega) \right\}\end{aligned}$$

k_{FT}^{-1} is the Thomas-Fermi screening length, f is the Fermi function and E_k is the single particle energy in the Hartree-Fock approximation. Our calculation satisfies particle conservation.

The integral equation has been solved by iteration. Delta function and principal value integrals are involved. The former is done by explicitly finding the zero of the argument of the δ

function. The principal value integral is calculated by approximating the singular part of the integrand with a Lorentzian of width γ and finding the value as γ approaches zero.

The real (ϵ_1) and imaginary (ϵ_2) parts of the dielectric function for $r_s = 3$ and $q = k_F$ are shown on the graphs. E_F^0 is the Fermi energy of the noninteracting gas. For comparison, the RPA and the Hubbard dielectric functions for the same q and r_s are shown. The screened potential used in the present calculation is also used to evaluate $\epsilon_{\text{Hubbard}}$.

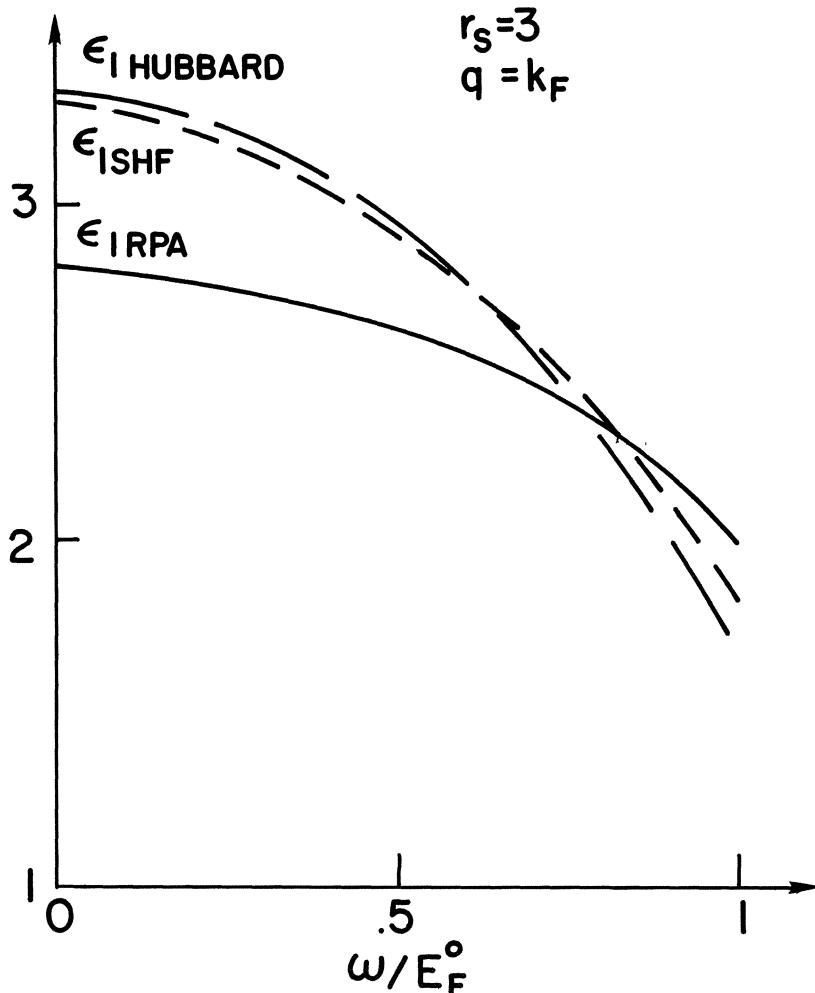


Figure 1. The real part of the dielectric function as a function of frequency.

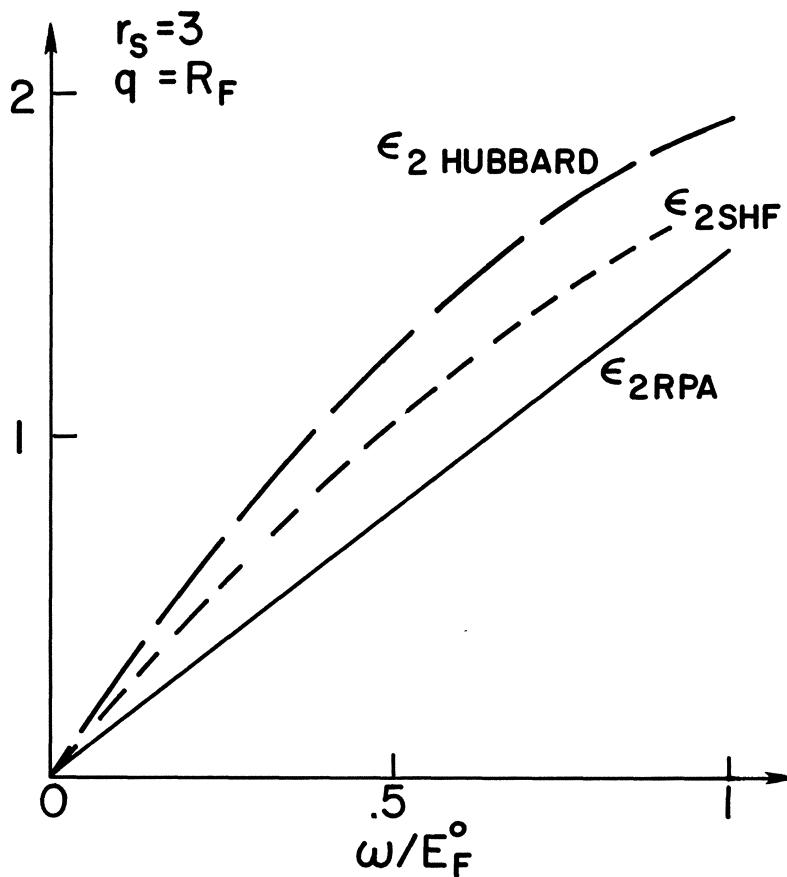


Figure 2. The imaginary part of the dielectric function as a function of the frequency.

KOHN-SHAM SELF-CONSISTENT SCHEME APPLIED TO THE
CALCULATION OF ATOMIC SYSTEMS AND METALLIC SODIUM[†]

B.Y. Tong

Department of Physics, University of Western
Ontario, London, Ontario.

INTRODUCTION

Several years ago Kohn and his collaborators¹⁻⁵ proposed a new approach to study the properties of an interacting inhomogeneous electron system. It is similar in spirit to the Landau theory of Fermi liquids. Both theories try to study certain physical phenomena of an interacting Fermi system in terms of a few parameters. In a way, the two theories supplement each other. The Landau theory is after the single particle excitations, whereas the Kohn theory emphasizes the calculation of static properties and the total energy of the whole system. The basic quantity involved in the latter theory is the total electron density $n(\underline{r})$. A review of this theory is given by Sham⁶ at this conference, and we shall refer the readers to his report, and to the references¹⁻⁵ for the detailed description of this new approach.

In essence, the Kohn theory is an exact, formal variational principle for the total energy of an interacting inhomogeneous electron gas, in which the total electron density $n(\underline{r})$ is the variable function. Into this principle enters a universal functional $F[n(\underline{r})]$ which governs the ground state of all electronic systems, no matter what external potential the electrons are moving in. Once this functional is specified, the problem is

[†] Work completed at Physics Department, University of California, San Diego in 1967. It was partially supported by the U.S. Office of Naval Research.

reduced to solve several equations self-consistently. The Hohenberg-Kohn¹ formulation has been extended to study problems at finite temperatures by Mermin² and to include properties of elementary excitations by Sham and Kohn⁴.

To develop these ideas for practical calculations, Kohn and Sham³ formulated a self-consistent scheme, making full use of our present day knowledge of an interacting homogeneous electron gas. Their original scheme is formally exact. In our present report, we describe how the scheme can be used to study real physical systems: atoms and metallic sodium, by choosing a suitable functional $F[n]$. An approximate version of the scheme known as the local effective potential is used. To stay within the scope of this conference, we shall only emphasize the calculation procedures and discuss their advantages and short-comings. We shall leave the detailed discussion of the results and their interpretation to publications elsewhere^{7,8}.

I ATOMIC SYSTEMS

The second Hohenberg-Kohn theorem^{2,8} states that for a given external potential $v(\underline{r})$, the energy functional $E_v[n]$, defined as

$$E_v[n] = \int v(\underline{r}) n(\underline{r}) d\underline{r} + F[n] \quad (1)$$

attains its minimum for the correct density $n(\underline{r})$, and this minimum is equal to the ground state energy E . $F[n]$ is a unique functional of $n(\underline{r})$, and is a universal functional since it is the same for any number of particles and any external potential. It characterizes the whole system. For interacting electron systems as the case of an atom, it is convenient to write $F[n]$ formally as

$$F[n] = T_s[n] + \frac{1}{2} \iint \frac{n(\underline{r})n(\underline{r}')}{|\underline{r}-\underline{r}'|} d\underline{r} d\underline{r}' + E_{xc}[n] \quad (2)$$

where $T_s[n]$ is defined to be the kinetic energy of a non-interacting electron gas of density $n(\underline{r})$. It is not the kinetic energy of the actual system. The second term is the classical Coulomb energy. $E_{xc}[n]$ is what is left-over from the first two terms and is defined to be the exchange and correlation energy of the interacting electron gas. It is also a universal functional. The energy functional then takes the form

$$\begin{aligned} E_v[n] &= T_s[n] + \int v(\underline{r}) n(\underline{r}) d\underline{r} + \frac{1}{2} \iint \frac{n(\underline{r})n(\underline{r}')}{|\underline{r}-\underline{r}'|} d\underline{r} d\underline{r}' \\ &+ E_{xc}[n] \end{aligned} \quad (3)$$

By the stationary nature of the true ground state subject to the condition

$$\int \delta n(\underline{r}) d\underline{r} = 0 \quad (4)$$

Kohn and Sham³ obtained their formally exact self-consistent scheme. It is found that instead of solving complicated equations of the actual many-electron system, we have only to solve the problem of an equivalent non-interacting electron system moving in a single-particle effective potential V_{eff} :

$$\{-\frac{1}{2}\nabla^2 + V_{\text{eff}}(\underline{r})\} \psi_i(\underline{r}) = \epsilon_i \psi_i(\underline{r}), \quad (5)$$

where

$$V_{\text{eff}}(\underline{r}) = V(\underline{r}) + \int \frac{n(\underline{r}')}{|\underline{r}-\underline{r}'|} d\underline{r}' + V_{\text{xc}}(\underline{r}),$$

$$V_{\text{xc}}(\underline{r}) = \frac{\delta E_{\text{xc}}[n]}{\delta n(\underline{r})}. \quad (6)$$

The total density of the many-electron system is given by

$$n(\underline{r}) = \sum_{i=1}^N |\psi_i(\underline{r})|^2, \quad (7)$$

and the total energy is given by equation (3). The procedure to find the correct density and energy of the system is to solve equations (3) to (7) self-consistently.

The Kohn-Sham self-consistent scheme is formally exact. The exact knowledge of $E_{\text{xc}}[n]$ gives us the exact $F[n]$ and thus the exact solution. In practice, we do not know the exact functional $E_{\text{xc}}[n]$. One great merit of this scheme comes from the equivalent single-particle picture. It appeals to physical intuition and allows us to approximate the correct $E_{\text{xc}}[n]$. In the present report, we shall restrict ourselves to the local effective potential approximation³. Here, $E_{\text{xc}}[n]$ is defined to be

$$E_{\text{xc}}[n] = \int \epsilon_{\text{xc}}(n(\underline{r})) n(\underline{r}) d\underline{r} \quad (8)$$

where $\epsilon_{\text{xc}}(n(\underline{r}))$ is the exchange and correlation energy per electron of a uniform electron gas of density n . The effective potential is then

$$V_{\text{eff}} = V(\underline{r}) + \int \frac{n(\underline{r}')}{|\underline{r}-\underline{r}'|} d\underline{r}' + \mu_{\text{xc}}(n) \quad (9)$$

$$\text{where, } \mu_{\text{xc}}(\underline{r}) = \frac{d}{dn}(n\epsilon_{\text{xc}}) \quad (10)$$

is the exchange and correlation part of the chemical

potential of a uniform gas at density n . From our knowledge of a homogeneous interacting electron gas, we can split μ_{xc} further into the exchange and correlation parts:

$$\mu_{xc}(n) \equiv \mu_x(n) + \mu_c(n) \quad (11)$$

where

$$\mu_x = - \left(\frac{3n}{\pi} \right)^{1/3}. \quad (12)$$

For μ_c , we use a interpolation between the low density Wigner correlation formula and the high density expression of Gell-Mann and Brueckner. The actual interpolated curve for μ_c as a function of density n is given in reference 7.

For atomic systems, the external potential is the Coulomb potential of the nucleus:

$$V(\underline{r}) = - \frac{Z}{\underline{r}} \quad (13)$$

The boundary condition is $\psi_i(\infty) = 0$. The total energy (equation (3)) then becomes

$$E = \sum_{i=1}^N \varepsilon_i - \frac{1}{2} \iint \frac{n(\underline{r})n(\underline{r}')}{|\underline{r}-\underline{r}'|} d\underline{r} d\underline{r}' \\ + \int n(\underline{r}) \{ \varepsilon_{vc}(n(\underline{r})) - \mu_{xc}(n(\underline{r})) \} d\underline{r}. \quad (14)$$

These two equations together with equations (5) to (7) form a self-consistent set for the atomic calculation.

Thus we see how we have reduced a complicated problem of many-electron system of inhomogeneous density transformed to a relatively simple problem. It involves of solving equations having the analogs of single-particle picture. This, in a way, is similar to the conventional Hartree-Fock formulation. However, the single-particle equations have different physical meaning. For example, Koopman's theorem⁹ which can be applied in the Hartree picture of atomic system is no longer valid in our case.

The results of our computation on atomic systems have been published and we refer our reader to it for detailed discussion of the results and their comparison with other schemes like the Hartree-Fock method, and the Slater approximation⁷. Here, we shall only summarize the advantages and disadvantages of this local potential approximation of the Kohn-Sham self-consistent scheme as applied to the study of properties of atomic systems.

The local potential approximation is extremely easy

to use as compared to the usual schemes like the Hartree-Fock method. The effective potential V_{eff} (equation (9)) is the same for all electrons $i = 1 \dots N$. Therefore there is no problem of the orthogonality of the wave-functions.

It has also the advantage of being able to incorporate the exchange and correlation effects in a natural way. They can be separated if one wishes to estimate their individual contributions.

The results reported in reference 7 show that the total energies calculated with such a scheme is fairly good, but excellent total electron densities are obtained with this relatively much simpler computation. Theoretical estimates show that in a slowly varying density electron gas, this scheme gives electron densities accurate to $O(|\nabla n|^4)$ as compared to total energy which is accurate to $O(|\nabla n|^2)$. (The Slater scheme gives densities to $O(|\nabla n|^2)$.)

The approximate expression for $E_{xc}[n]$ is obtained from a uniform many-electron gas system of density n . It therefore follows that our result should be better the larger the atomic system. This is borne out from the results. Also it is seen from Table V of reference 7 that the major errors in the ionization energies come from the innermost electrons. The smallness of the atomic system is also the cause of the factor of two over-estimate of the correlation contribution reported previously⁷. These errors do not trouble us when we study metallic sodium, because the latter is a many-electron system. For atomic systems, they can be eliminated if we include the non-local potential. We have not carried out such computation.

Our simple computation yields an electron density comparable to that of a sophisticated Hartree-Fock scheme. For example, our Ne curve agrees well with the Hartree-Fock calculation by Worsley (1958)¹⁰, but it differs from the earlier calculation of Brown (1933)¹¹. The calculation of Kr is in much closer agreement with a recent Hartree-Fock calculation by Mann (1967)¹², than the earlier Worsley calculation (1958)¹³.

To solve the equation (5) we use the Herman and Skillman program¹⁴, which is well-known for its speed and accuracy. It involves only a minor modification of the program to insert the contribution from μ_{xc} in the

potential. The unmodified tail boundary condition, i.e. $\psi_i(\infty) = 0$, is used. A typical calculation of an atom to attain self-consistency to six significant figures on CDC 3600 is around one minute or less.

III METALLIC SODIUM

We shall sketch here the method used in the computation of several properties of metallic sodium, namely the equilibrium lattice separation, compressibility, cohesive energy, Knight shift, etc., but shall leave the detailed description and discussion to future publication elsewhere.

The Kohn-Sham self-consistent scheme is used. As emphasized above, the nature of the single particle Schrödinger equations (5) in this scheme is somewhat different from the conventional Hartree-Fock picture. Here we also speak of "core" and "valence" electrons, but always in the sense described in the previous section. In the present calculation the quantities we obtain from self-consistency are the total energy and the total electron density of the whole system. Since the number of electrons in a metallic sodium is large, the difficulty due to a small electron system mentioned in the last section does not appear here.

Sodium has practically a spherical Fermi surface. As suggested by Wigner and Seitz, we divide the metallic sodium crystal into cells and use their spherical approximation. This means that to a first approximation the energy of an atom in metallic sodium depends only on the atomic volume, and not on the coordinate number. It is supported by the experimental fact that energy remains practically unchanged when the coordinate number is changed, e.g., in the martensitic phase transformation of sodium from body-centered cubic to the hexagonal close-packed structure¹⁶, and the theoretical study of Hughes and Callaway¹⁷. The cells are neutral because the ground state of a many-electron system does not allow an excess accumulation of charges in a particular cell.

The Kohn-Sham self-consistent scheme can now be written as

$$\begin{aligned} E[n] &= T_s[n] - \int \frac{Zn(\underline{r})}{\underline{r}} d\underline{r} \\ &+ \frac{1}{2} \iint \frac{n(\underline{r}') n(\underline{r})}{|\underline{r}-\underline{r}'|} d\underline{r}' d\underline{r} + \int \epsilon_{xc}(n(\underline{r})) n(\underline{r}) d\underline{r} \end{aligned} \quad (15)$$

where $E[n]$ is now the total energy functional of all electrons in a single cell, and $Z = 11$ for sodium. All integrations in this equation are limited to the cell. The total electron density $n(\underline{r})$ is given by the sum of $n_c(\underline{r})$ and $n_v(\underline{r})$ respectively:

$$\begin{aligned} n(\underline{r}) &= n_c(\underline{r}) + n_v(\underline{r}) \\ n_c(\underline{r}) &= \sum_{\alpha} |\psi_{\alpha}(\underline{r})|^2 \\ n_v(\underline{r}) &= \sum_{\mathbf{k}} |\psi_{\mathbf{k}}(\underline{r})|^2 \end{aligned} \quad (16)$$

Here α runs over all the core levels, and n_v is the valence electron density obtained by summing over all the \mathbf{k} values below the Fermi surface. $T[n]$ is the single particle kinetic energies of the homogeneous electron gas in an effective potential V_{eff} satisfying the equations

$$\left\{ -\frac{\nabla^2}{2} + V_{\text{eff}}(\underline{r}) \right\} \psi_i = \varepsilon_i \psi_i, \quad i = \text{set } \{\alpha, \mathbf{k}\} \quad (17)$$

under the appropriate boundary conditions of the valence and core electrons. As before,

$$V_{\text{eff}}(\underline{r}) = -\frac{Z}{r} + \int \frac{n(\underline{r}')}{|\underline{r}-\underline{r}'|} d\underline{r}' + \mu_{xc}(n(\underline{r})) \quad (18)$$

The boundary condition for the core electrons is taken to be $\psi_{\alpha}(\underline{r})|_{\underline{r}=r_s} = 0$. r_s is the Wigner-Seitz cell radius, and is $(\frac{3}{4\pi n})^{1/3}$. The valence electron satisfies periodic boundary condition. It is found that the variational method of Kohn¹⁸ in solving such problems can be easily incorporated into this scheme. Computation-wise, it involves again repeated use of the modified Herman and Skillman program with some additional averaging process and boundary condition tests sketched below. This combination of Kohn-Sham scheme, Kohn variation method, and Herman and Skillman program preserves the two characteristics required of the calculation: accuracy and speed. Accuracy is of extreme importance in the present calculation. The total energy E is of the order -323 ry., whereas the quantities which we want to evaluate are extremely small in comparison with E , e.g. the bulk modulus K is given by

$$K = \frac{1}{12\pi r_s^2} \frac{\partial^2 E}{\partial r_s^2} \quad (19)$$

in the spherical approximation. For r_s close to the equilibrium r_s^0 , the change in E are small (~ 0.01 ry. for a charge of about .12 in r_s).

The above equations (15) to (18) form a self-consistent set. The boundary value problem of the valence electron in the spherical approximation (with a spherical symmetrical potential) can be written as:

$$\begin{aligned} (-\frac{1}{2}\nabla^2 + V_{\text{eff}}(r) - \varepsilon_k) \psi_k(\underline{r}) &= 0, \quad r < r_s ; \\ \psi_k(\underline{r}) &= \exp[2ikr_s \cos\theta] \psi_k(-\underline{r}), \quad r = r_s ; \\ \frac{\partial \psi_k(\underline{r})}{\partial r} &= -\exp[2ikr_s \cos\theta] \frac{\partial \psi_k(-\underline{r})}{\partial r}, \quad r = r_s ; \end{aligned} \quad (20)$$

where k is the magnitude of the wave-vector assumed in the z -direction, and θ is the angle between \underline{k} and \underline{r} . Assume that $\psi_k(\underline{r})$ can be expanded in terms of spherical harmonics:

$$\psi_k(\underline{r}) = \sum_l n_l C_l P_l(\cos\theta) \frac{R_{k,l}(r)}{r} \quad (21)$$

where n_l is defined as

$$n_l = \begin{cases} 1 & \text{when } l \text{ is even} \\ i & \text{when } l \text{ is odd} \end{cases}, \quad (22)$$

and C_l are real.

The variation principle of Kohn¹⁶ states that C_l are determined by the condition that the surface integral

$$K \equiv \int_S \frac{\partial \psi_k(\underline{r})}{\partial r} \psi_k^*(-\underline{r}) \exp[-2ikr_s \cos\theta] \sin\theta d\theta \quad (23)$$

be stationary. Substituting (21) into this, and normalizing the $R_{k,l}$ such that $R_{k,l}(r_s) = r_s$, we get the secular equation

$$\det |n_l n_m (L_l + L_m) I_{lm}| = 0 \quad (24)$$

where

$$L_l = (\frac{1}{R_l} \frac{dR_l}{dr} - \frac{1}{r}) \Big|_{r=r_s} \quad (25)$$

and

$$\begin{aligned} I_{lm} &= \int_0^\pi \exp[-2ikr_s \cos\theta] P_l(\cos\theta) P_m(\cos\theta) \\ &\quad \cdot \sin\theta d\theta . \end{aligned} \quad (25)$$

Here the index k is suppressed. The valence electron density is then given by taking the sum overall \underline{k} -points up to the Fermi surface,

$$n_v(\underline{r}) = \sum_{\underline{k}} |\psi_{\underline{k}}(\underline{r})|^2 . \quad (26)$$

With $\int \psi_k^* (\underline{r}) \psi_k (\underline{r}) d^3 r = 1$, we get

$$n_v (\underline{r}) = \frac{3}{4\pi r^2 k_F^3} \int_0^{k_F} \frac{\sum_{-\ell} C_\ell R_{k,\ell}^2 (r)}{\sum_{\ell} \frac{C_\ell^2}{2\ell+1} \int R_{k,\ell}^2 (r) dr} k^2 dk \quad (27)$$

and

$$\epsilon_v = \frac{3}{k_F^3} \int_0^{k_F} \epsilon_k k^2 dk . \quad (28)$$

Since the valence band of sodium is nearly a parabola, we find it quite sufficient to evaluate only seven k -points equally spaced from 0 to k_F . Notice that once k is specified, $I_{\ell,m}$ can be evaluated for all wave-functions $R_{\ell,m}$. It saves us from evaluating $I_{\ell,m}$ for each trial wave-function $R_{\ell,m}$. It can also be seen that $I_{\ell,m}$ is related to the 3-j symbols:

$$I_{i,j} = \sum_{\ell=0}^{\infty} 2(2\ell+1)(-i)^{\ell} j_{\ell}(\rho) \begin{pmatrix} \ell & i & j \\ 0 & 0 & 0 \end{pmatrix}^2, \quad \rho = 2kr_s \quad (29)$$

In practice the expansion (21) in ℓ is terminated at $\ell \leq 6$. Test calculations show that this truncation is sufficient to secure six significant figures. The rapid convergence can be seen from the coefficients C_{ℓ} of the expansion.

To test that the solution is the correct one, satisfying equation (24), we also employ an extra check. It is shown by Kohn¹⁸ that the correct solution satisfies

$$\frac{C_1 P_1(\cos\theta) + C_3 P_3(\cos\theta) + \dots}{C_0 P_0(\cos\theta) + C_2 P_2(\cos\theta) + \dots} = \tan(k r_s \cos\theta) \quad (30)$$

In the actual calculation, an iterative process is used to find the correct wave-functions and ϵ_k .

Because of the large scale of self-consistent iteration involved in the complete calculation, we separate the complete calculation into two parts.

First we choose a $r_s = r_a$ slightly smaller than the equilibrium lattice parameter r_a^0 . The periodic boundary condition is used for the valence electron, and $\psi_{\alpha}(r)|_{r=r_a} = 0$ is used for the core electrons. The calculation is cycled until self-consistency up to seven significant figures is obtained.

We calculate all other $r_s > r_a$, by fixing the core

density $n_s(r)$ to be the same as $n^0(r)$. Self-consistency is attained for the valence electron. In doing this we have made the heuristic assumption that the single particle kinetic energy $T_s[n]$ can be split into the core and the valence parts depending on their corresponding densities. The validity of the assumption is supported by the fact that the final results of two different r_a calculations are practically the same.

Computations were carried out on digital computer CDC 3600. Each cycle in producing a new improved potential involves of solving nearly a hundred Schrödinger equations. Most of these are related to finding self-consistently the k -dependent valence solution. Speed, therefore, is an essential factor in the computation. Together with various boundary condition tests, determinant evaluations and various integrations, the total time taken in one cycle amounts to about one minute. First part of the calculation may take as long as 30 minutes if the input is not close to the true solution. The second part of the calculation usually takes about five to seven minutes for each value of r_s .

We did two sets of calculations of E as functions of r_s using two different r_a 's. Each set consists of four points only. The results are summarized in Table I. Previous calculations of the same quantities are also listed for comparison. However, owing to limited space here, we shall not go into details.

Briefly, we see that here we have an apriori calculation starting from no other information than $Z = 11$. Most other computations use constructed single-particle potentials based on some experimental data, e.g., the quantum defect calculations. In many cases, the core electrons are taken to be the same as that of the atoms. These effective potentials have a discontinuity at r_a . Our core electrons do not have this difficulty, and they satisfy the correct boundary condition. It is very nearly the true self-consistent value, since r_a is close to r^0 , and all calculated r_s are in the near neighbourhood of r_a .

Exchange and correlation effects are taking into account for both the core and the valence electrons. Previous apriori calculations, e.g. Hartree-Fock core density, usually fail to take care of both contributions in a natural way.

The scheme is a fairly large scale self-consistent

Table I. Equilibrium Lattice Parameter, Bulk Modulus K, and Cohesive Energy of Metallic Sodium at 0°K

Author or Method	Lattice Constant (Å)	K (10^{11} dynes/cm 2)	Cohesive Energy (kcal/mol)
1. Experiment	4.225^a (at $T=5^\circ\text{K}$)	.85 ^b 5 .57 ^c (at $T=293^\circ\text{K}$)	26.0
2. Modified Kohn-Sham scheme ^c			
r_a^1	4.103 ₀	.91	(33.8)
r_a^2	4.102 ₇	.89	(34.0)
3. (free ion core)	4.115 ₇	.87	(32.2)
4. Wigner & Seitz (1934)	4.75		23.2
5. Fuchs (1936)		.87 ₈	
6. Bardeen (1938)	4.53	.83	23.0
7. Kuhn & Van Vleck (1950)	4.14	.78	25.9
8. Brooks (1953)	4.26	.61 ₇ ^d (at $T=293^\circ\text{K}$)	26.3

^aQuoted by C.S. Barrett. (1956) *Acta Cryst.* 9, 671.

^bQuoted by Fuchs (1936).

^cThe parameter r_a is chosen in the two cases as $r_a^1 = 3.65445442$ and $r_a^2 = 3.71814866$.

^dVan Vleck, J.H. (1954) Proc. 1953 intern. conf. of theor. phys., Kyoto and Tokyo. Page 641.

References:

- Bardeen, J. (1938) *J. Chem. Phys.* 6, 367.
- Brooks, H. (1953) *Phys. Rev.* 91, 1027.
- Fuchs, K. (1936) *Proc. Roy. Soc. A* 157, 444.
- Kuhn, T.S. and J.H. Van Vleck. (1950) *Phys. Rev.* 79, 382.
- Wigner, E. and F. Seitz. (1934) *Phys. Rev.* 46, 509.

calculation. We did only four points for each r_a . We would like to do a more thorough computation, probably without splitting the procedure into two parts as described above.

The final V_{eff} obtained can in a way be regarded as another form of pseudopotential. Pseudopotentials are usually obtained from fitting some experimental data. By this method we have a strictly apriori calculated pseudopotential.

The results (2) in Table I are reasonably good, especially the equilibrium lattice constant, and compressibility. Cohesive energies from our calculations are put in brackets because they are less dependable numbers. They are obtained by subtracting energy of the isolated atom from the energy per atom of metallic sodium. These are two large and nearly equal numbers, and our electron-gas type of theory is not sufficiently accurate for the atomic calculation to yield a dependable difference.

We saw that the Kohn-Sham scheme gives excellent electron densities even in atomic systems. It is therefore expected we can make a good Knight shift calculation. This is being carried out.

ACKNOWLEDGEMENTS

I would like to express my gratitude to Professor W. Kohn who suggested, and directed me throughout the course of this work, and to Dr. L.J. Sham for his continuous guidance and comments. I also would like to thank Mrs. C. Troop for the excellent typing of this manuscript.

REFERENCES

1. P. Hohenberg and W. Kohn, Phys. Rev. 136, B 864 (1964).
2. D. Mermin, Phys. Rev. 137, A1441 (1965).
3. W. Kohn and L.J. Sham, Phys. Rev. 140, A1133 (1965).
4. L.J. Sham and W. Kohn, Phys. Rev. 145, 561 (1966).

5. W. Kohn, Many-body Theory, 1965 Tokyo Summer Lectures in Theoretical Physics, Part I, page 73-85. New York, W.A. Benjamin, 1966.
6. L.J. Sham, report at this conference.
7. B.Y. Tong and L.J. Sham, Phys. Rev. 144, 1 (1966), and manuscripts under preparation.
8. B.Y. Tong, Thesis, University of California, San Diego, (1967).
9. T.A. Koopman, Physica 1, 104 (1933).
10. B.H. Worsley, Can. J. Phys. 36, 289 (1958).
11. F.W. Brown, Phys. Rev. 44, 214 (1933).
12. J.B. Mann, private communication.
13. F.W. Worsley, Proc. Roy. Soc. (London) A247, 390 (1958).
14. F. Herman, and S. Skillman, Atomic Structure Calculation, New Jersey, Prentice-Hall, 1963.
15. E.P. Wigner and F. Seitz, Solid State Physics 1, 97 (1955).
16. C.S. Barrett, Acta Cryst. 9, 671 (1956).
17. A.J. Hughes and J. Callaway, Phys. Rev. 136, A1390 (1964).
18. W. Kohn, Phys. Rev. 87, 472 (1952); Ibid. 96, 590 (1954).

A POTENTIAL FUNCTION FOR BAND STRUCTURE CALCULATIONS

David A. Liberman

University of California, Los Alamos Scientific
Laboratory, Los Alamos, New Mexico

Since 1951, when Slater¹ proposed the local $\rho^{1/3}$ approximation for the exchange potential, it has been much used in atomic and solid state band structure calculations. Recently there have been many attempts at modifying the method. Most of these efforts come from solid state physicists which is fair evidence, I think, that the method suffers from real difficulties when used for solids.

The nature of the difficulty can be seen by looking at some simple cases. The first one is a lattice of hydrogen atoms spaced far apart. In any cell of the lattice the potential function is $-e^2/r$, where r is measured from the nucleus in that cell. But the $\rho^{1/3}$ method gives two more terms--the Coulomb potential of the electron charge distribution and the approximate exchange potential. If these last two terms cancelled each other a correct calculation of the energy would be obtained, but they don't--not even approximately. This failure of the $\rho^{1/3}$ exchange potential to accurately cancel the self-interaction part of the Coulomb potential is the main difficulty, I think.

It might be hoped that a correlation potential would help. To show that the same difficulty persists if correlation is represented by a simple function of density, consider the case of a low density electron gas. Wigner² showed that the electrons will arrange themselves on lattice. In any given cell the potential arises solely from the positive charge background in that cell. In Kohn and Sham's³ version of the $\rho^{1/3}$ method there is, however, also the Coulomb potential of the electron charge distribution,

the $\rho^{1/3}$ exchange potential, and the correlation potential. From Wigner's work it is known that the correlation energy in the low density limit is 0.96 times the exchange energy. If the variational method is then used to derive an exchange-correlation potential, it is found to have the $\rho^{1/3}$ form and to have a strength 1.3 times Slater's exchange potential. Again it may be asserted that there is inadequate cancellation of the self-interaction part of the Coulomb potential by the exchange-correlation potential.

Finally, consider a more realistic example--an alkali metal at normal density. Wigner and Seitz⁴ gave a prescription for treating the alkali metals which certainly works even if it is not the only one that will. First they chose to ignore the interactions of the valence electrons with each other except insofar as they cancel the potentials of charges outside a given cell. Using only an ion core potential they obtained very good estimates of the cohesive energies. Then they treated the inter-electron terms by perturbation means and found that they very nearly cancelled one another. If they had put these terms into the potential function according to the $\rho^{1/3}$ method, they would have obtained very poor results in their calculation of lattice parameters and cohesive energies just as in the previous examples. This is not obvious but it is true, as I found out by doing the self-consistent field band structure calculation for lithium. With the Kohn-Sham exchange potential and a normal density lattice parameter the pressure came out to be +105 kbar. Since the bulk modulus is about 120 kbar, a zero pressure density about half as large as the observed one is indicated. Commonly used correlation terms are too small by a factor of 5 or 10 to correct this.

How can the $\rho^{1/3}$ method be modified so that it will work for the examples I have used above and, hopefully, for solids in general? I offer the following prescription which avoids the problems arising from the self-interaction part of the Coulomb potential, and has already given good results for lithium, beryllium, aluminum, and iron:

Imagine that a cell of a simple solid is divided into two parts by a sphere centered on the atomic nucleus as indicated in the figure. The region inside the sphere is called S, the region outside T, and the sphere radius R. R is defined by the condition that region T contains, on the average, just one electron in it:

$$\int_T \rho(\vec{x}) d\vec{x} = 1. \quad (A)$$

The electron-electron potential energy for the charge inside the

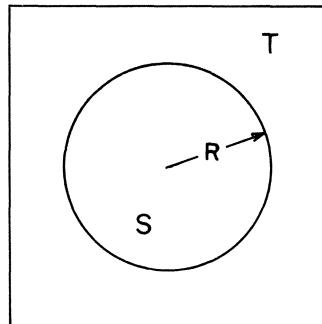


Fig. Schematic diagram of a cell in a crystal lattice. The region T contains one electron. The charge density is assumed to be spherically symmetric.

sphere is approximated by a Coulomb and an exchange term exactly as suggested by Kohn and Sham:

$$U = \frac{e^2}{2} \int_S \int_S \frac{\rho(\tilde{x}_1)\rho(\tilde{x}_2)}{|\tilde{x}_1 - \tilde{x}_2|} d\tilde{x}_1 d\tilde{x}_2 - \frac{3e^2}{4\pi} \int_S \rho(\tilde{x})(3\pi^2\rho(\tilde{x}))^{1/3} d\tilde{x}. \quad (B)$$

(A density dependent correlation term may be added here if desired.) Since, on the average, there is only one electron in T, there is no contribution to the inter-electron potential energy from the charge in T interacting with itself. Finally there is the potential energy of the electron in T interacting with those in S:

$$V = e^2 \int_S \int_T \frac{\rho(\tilde{x}_1)\rho(\tilde{x}_2)}{|\tilde{x}_1 - \tilde{x}_2|} d\tilde{x}_1 d\tilde{x}_2. \quad (C)$$

There can be no exchange term here because the "hole" in the two-particle distribution function has been "used up" in assuming no self-interaction of the charge in T. Also it is assumed that because of electrical neutrality, there is no interaction between the charge in one cell and that in another.

The expressions for kinetic energy and electron-nucleus potential energy are as usual. The variational method may now be used to find stationary values of the total energy subject to the usual normalization condition and to the condition (A) above. This second constraint on the variation may be introduced with a second Lagrange multiplier. Variations of the sphere radius, R, must be considered as well as those in the wave function. The final result of some fairly obvious manipulation is the one-electron potential

TABLE. BULK PROPERTIES OF SOLIDS AT ZERO PRESSURE AND TEMPERATURE

		Lithium	Beryllium	Aluminum	Iron
Wigner-Seitz cell radius (Bohr radii)	calc. exp. ^a	3.18 3.24	2.40 2.35	3.06 2.99	2.69 2.66
Cohesive energy (kcal/mole)	calc. exp. ^b	40 38	59 78	61 77	160 182
Bulk modulus (kbar)	calc. exp. ^c	131 123	1040 1100	930 780	1750 1690

^aThe lithium cell radius is from C. A. Swenson, J. Phys. Chem. Solids 27, 33 (1966). The others are from Crystal Data, J. D. H. Donay, Editor, Amer. Crys. Assn. (1963).

^bJANAF Thermochemical Tables, D. R. Stull, Editor, Clearinghouse for Federal Scientific and Technical Information, Springfield, Va. (1965). The experimental cohesive energy of iron has been increased by 83 kcal/mole so as to represent the value relative to the average over the ground state configuration for the iron atom as the calculation of the iron atom is for this average. This value is based partly on observed energy levels in iron and partly on a configuration-splitting calculation by R. D. Cowan (unpublished).

^cThe lithium bulk modulus is Swenson's (Ref. a). The others are from velocities of sound quoted in the Handbook of Chemistry and Physics, R. C. Weast, Editor, The Chemical Rubber Co., Cleveland, Ohio (1967).

function:

$$v(r) = -Ze^2/r + e^2 \left\{ \begin{array}{l} \int_{S+T} \rho(x)/|r-x| dx - \frac{1}{\pi} (3\pi^2 \rho(r))^{1/3} + \lambda, \quad r \in S \\ \int_S \rho(x)/|r-x| dx \end{array} \right. , \quad r \in T \quad (D)$$

where

$$\lambda = - \int_T \rho(x)/|R-x| dx + \frac{3}{4\pi} (3\pi^2 \rho(R))^{1/3}.$$

With this potential function self-consistent field band structure calculations have been done for lithium, beryllium, aluminum, and iron. The calculated bulk properties of these metals and the corresponding observed values are shown in the table. The assumptions used in constructing the inter-electron potential energy functional--U+V--more or less guarantee good results for valence bands with one electron per atom. It is gratifying that good results are also obtained in the more complicated cases.

REFERENCES

1. J. C. Slater, Phys. Rev. 81, 385 (1951).
2. E. Wigner, Phys. Rev. 46, 1002 (1934); Trans. Faraday Soc. 34, 678 (1938).
3. W. Kohn and L. J. Sham, Phys. Rev. 140, A1133 (1965).
4. E. Wigner and F. Seitz, Phys. Rev. 43, 804 (1933); 46, 509 (1934).
5. D. A. Liberman, Phys. Rev. B (in press).

Comment by L. J. Sham

I believe that the large pressure at the experimental lattice constant Liberman obtained is due to his inclusion of exchange without correlation. Since, for a constant density appropriate for sodium, exchange energy and correlation energy per electron are comparable, it is difficult to understand why, for alkali metals, Liberman found the "commonly used correlation term" to be too small by a factor of 5 or 10. Tong reports elsewhere in this proceeding that he has applied the method which Kohn and I proposed to sodium starting with the Prokofiev potential for the ion and obtained satisfactory results for total energy per electron, equilibrium lattice constant and bulk modulus.

We expect our method to be valid when the Wigner-Seitz method is successful because the latter can be put within our framework.

Write the total energy as

$$E = T_I[n] + \int dr V(r) n(r) + \frac{1}{2} \iint dr dr' n(r) \frac{1}{|r-r'|} n(r')$$

$$+ \int dr n(r) \epsilon_{xc}(n(r))$$

where T_I is the kinetic energy of non-interacting Fermions with density distribution $n(r)$ and $\epsilon_{xc}(n)$ is the exchange and correlation energy per electron. Wigner-Seitz's method amounts to dividing the solid into cells, treating the last two terms using constant density only and calculating the first two terms using the single particle prescription with the bare ion potential $V(r)$ for each cell.

Answer by D. A. Liberman

The contribution to the pressure in lithium at normal density from correlation is -21 kbar according to the familiar expression of Gell-Mann and Brueckner. The Pines-Nozières calculation of the correlation energy gives a pressure term half as large. These fall short by factors of 5 and 10 respectively of reducing to zero the pressure (+105 kbar) I obtained using the Kohn-Sham exchange potential.

Sham notes that the Wigner-Seitz calculations for the alkali metals can be brought within the framework of the Hohenberg-Kohn-Sham method if $V(r)$ in his formula is an ion core potential and the last two terms are treated as perturbations. But if these last terms are not treated as a perturbation and are included in a straightforward fashion, they will yield additional terms in the potential function which in effect screen it strongly at large radii. As a result the wave function is modified and there is a redistribution of charge. The net effect, if my calculations are to be believed, is a weaker binding of the valence electrons to the cores and a higher pressure.

CRYSTAL POTENTIALS USED IN ENERGY BAND THEORY

A. M. Boring and E. C. Snow

University of California
Los Alamos Scientific Laboratory
Los Alamos, New Mexico 87544

We have begun a study which has as its goal the determination of a good, one-electron, exchange-correlation operator for crystals. This paper contains preliminary results of eight, self-consistent field, APW calculations performed for copper in the course of this work. We chose Cu because its electronic structure has been determined by several energy band methods, and also because relativistic and non-muffin tin effects are small in the crystal. In these computations all electronic states were calculated at each iteration. The core states (1s, 2s, 2p) were determined in the crystal field by a Herman-Skillman type program with boundary conditions, written by Snow. The innerband states (3s, 3p) were calculated at 32 points in the Brillouin zone via the APW program, and the band states (3d, 4s) were similarly calculated at 2048 points. Our self-consistent criterion was a maximum energy difference in the one-electron eigenvalues of 0.002 ry from one iteration to the next.

In these calculations we have used an exchange-correlation operator that is a product of a local operator L times an operator N which is either local or non-local. In all the cases considered to date the local operator L is the Slater electron gas exchange operator, i.e.,

$$L = V_{SX}(r) = - 3/2(2/\pi)k_F(r) = - 6(3/8\pi\rho(r))^{1/3} \text{ ry.}$$

The operator N was chosen to be one of the following: local, only slightly non-local, and very non-local. These three cases will be described separately.

Local: In five of the calculations N is a constant (α). Three of these calculations had been done earlier by Snow,¹⁻² when he was trying to determine whether the Slater exchange-correlation ($\alpha=1$) or the Kohn-Sham-Gasper exchange-correlation ($\alpha=2/3$) gave the

better self-consistent results as compared with experimental data. Since it appeared that a value of α somewhere between these two limits would do the best job, the $\alpha=5/6$ calculation was performed.² These calculations were not repeated in this study but the results were rechecked. The $\alpha=0.721$ calculation was performed because that value was suggested by Slater³ using first principles arguments as a guide line. Specifically, $\alpha=0.721$ is the value for which the virial theorem is satisfied in the Cu atom when the energies are computed via Hartree-Fock theory using X_α orbitals. Since Kmetko⁴ has determined the α 's that minimize* the total energy for the whole periodic table, one immediately has a good exchange-correlation operator for any crystal, if this operator does in fact give the best overall results. The $\alpha=0.300$ calculation was undertaken in an attempt to reproduce the results obtained by the non-local $\alpha(\eta, r)$ exchange-correlation operator (discussed below).

Only slightly non-local: Two calculations were performed in which an approximation to the Hedin-Lundqvist type exchange-correlation operator⁵ was used. (To indicate the merits of our self-consistent solutions relative to the exact solutions, we remark that if we had used the Hedin-Lundqvist operator without approximation we would have obtained self-consistent solutions of the Dyson equation which would have been valid only in the Random Phase Approximation.) For the electron gas problem, Hedin and Lundqvist have plotted their operator as functions of the wave vector k of the one-electron state under consideration and of the density parameter r_s . Thus, the Hedin-Lundqvist operator is:

$$\alpha(\eta, r_s) ; \quad \eta = k/k_F .$$

We have approximated this operator by

$$\alpha(\eta, r) = C + C'\eta(r) .$$

In this operator the density dependence appears in two ways: (1) the parameter C is determined from the value of the density, and (2) $\eta(r)$ can be density dependent. We chose to use the average electron density of Cu in determining C so that it is a constant, i.e., $C=0.793$. The parameter C' is just the slope of the Hedin-Lundqvist function which for $\eta \lesssim 2$ is also a constant and equal to 0.036. With these parameters chosen, we did one calculation in which the operator η was density dependent and one in which it was not. Note that η always depends upon the wave vector of the state under consideration. For the density dependent operator η , k and k_F were determined by a scheme given by Liberman,⁶ i.e.,

$$k_F(r) = (3\pi^2\rho(r))^{1/3} ,$$

$$k(r) = \sqrt{E - V_{TOT}(r)} ,$$

$$V_{TOT}(r) = V_C(r) + \alpha(\eta, r)V_{XS}(r) .$$

For the non-density dependent operator, k and k_F were determined

*The α that satisfies the virial theorem for a given atom is very near that which minimizes the total energy.

from the equations $k = \sqrt{E - E(\Gamma_1)}$ and $k_F = \sqrt{E_F - E(\Gamma_1)}$, so that the non-local nature (k -dependence) of the operator would be determined from the energy (E) of the wave vector relative to the lowest occupied band state Γ_1 .

Very non-local: In the final calculation, the exchange-correlation operator was $V_{XC}(\eta, r) = F(\eta, r)V_{SX}(r)$, where

$$F(\eta, r) = 1/2 + \frac{1-\eta^2}{4\eta} \ln \left[\frac{|1+\eta|}{|1-\eta|} \right]$$

is the Liberman version of the Fock electron gas exchange operator in which $\eta(r)$ is determined as in Ref. 6 (as above). This exchange-correlation operator has been used by Slater, Wilson, and Wood⁷ in atomic calculations and gives results that are in good agreement with Hartree-Fock results. This has led us to believe that the results obtained with this operator for the solid will have many features in common with the results that would be obtained if the true Hartree-Fock operator were used. We note in passing that this operator gives the same general result in a rigorous band calculation as does the Hartree-Fock operator for the electron gas; i.e., the band widths are about twice as wide as those obtained both from experimental data and from many other calculations.

In Table I are given the results of our eight calculations that involved use of the exchange-correlation operators mentioned above. Also shown are the photoemission data on Cu and the non-self-consistent AFW calculation by Burdick,⁸ who used the Chodorow potential. Concerning the energy differences given in the table, $\Gamma_{25}-\Gamma_1$ gives the relative position of the sp- and d-bands at $\vec{k}=0$, $X_5-\Gamma_1$ gives the top of the d-band relative to the bottom of the sp-band, X_5-X_1 gives the width of the d-band, $X_4-\Gamma_1$ gives the width of the sp-band, and the last three differences (E_F-X_5 , E_F-L_3 , and E_F-L_2) were included because they have been determined experimentally. From the data presented in Table I, we are reluctant to choose any one of the exchange-correlation operators as giving the "best" results, although the $\alpha=5/6$ calculation seems to do well for the photoemission data. It is with this in mind that we compare the calculational results in Table I.

The first interesting feature is that the local operators do not seem to have any effect on the width of the sp-band.* The $\alpha(\eta, r)$ operator, however, caused the sp-band to narrow slightly and the $F(\eta, r)$ operator caused it to widen considerably. The width of the d-band is very sensitive to both local and non-local forms of the exchange-correlation operator. For the local operators,

*The width of the sp-band for $\alpha=2/3$ is incorrect because the state at Γ_1 was calculated incorrectly, the correct value is closer to 0.800.

Table I. Energy differences indicating widths and positions of bands.^a

	Experimental						
	$\Gamma_{25}' - \Gamma_1$	$X_5 - \Gamma_1$	$X_5 - X_1$	$X_4 - \Gamma_1$	$E_F - X_5$	$E_F - L_3$	$E_F - L_2'$
			0.250		0.147	0.162	0.026
APW Calculations ^b							
Burdick ^f	0.399	0.512	0.249	0.804	0.143	0.154	0.045
Snow ($\alpha=1$)	0.278	0.368	0.189	0.793	0.223	0.232	-0.005
Snow ($\alpha=5/6$)	0.375	0.477	0.244	0.794	0.149	0.159	0.029
Snow ($\alpha=2/3$)	0.392	0.508	0.262	0.720	0.111	0.123	0.102
$\alpha=0.721$	0.426	0.535	0.243	0.796	0.118	0.129	0.057
$\alpha=0.300$	0.525	0.690	0.398	0.799	0.056	0.077	0.201
$F(\eta, r)$	0.791	0.958	0.395	1.253	0.241	0.257	0.091
$\alpha=C+C'\eta(r)$	0.371	0.474	0.225	0.774	0.140	0.150	0.033
$\alpha=C+C'\eta$	0.348	0.436	0.196	0.749	0.143	0.151	0.021

^aenergies are given in ry.

^bSelf-consistent, nonrelativistic calculations at 2048 points.

^cC. W. Berglund and W. E. Spicer, Phys. Rev. 136, A1044 (1964).

^dA. H. Lettington, Phys. Letters 9, 98 (1964).

^eD. E. Eastman and J. K. Cashion, Phys. Rev. Letters 24, 310 (1970).

^fNon-self-consistent calculation, Ref. 8.

the results indicate that decreasing α increases the width of the d-band in a somewhat linear fashion. Again, the $\alpha(\eta, r)$ operator slightly decreased the band width while the $F(\eta, r)$ caused it to widen. Other comparisons of interest in Table I are found in the $F(\eta, r)$ and in the $\alpha=0.300$ calculations. The value of 0.300 resulted from a trial and error attempt to find an α that would give the same width for the d-band as would the $F(\eta, r)$ operator. While the $\alpha=0.300$ operator gave the same d-band width, it did not reproduce any of the other features of the $F(\eta, r)$ calculation.

In future work we plan to compare the results of these eight calculations with Fermi surface data. Next we will attempt some cohesive energy calculations, which should indicate whether or not a single exchange-correlation operator can provide good results both on bulk properties and on single excitation features.

REFERENCES

1. E. C. Snow and J. T. Waber, Phys. Rev. 157, 570 (1967).
2. E. C. Snow, Phys. Rev. 171, 785 (1968).
3. J. C. Slater, "Present Status of the X_α Statistical Exchange", Semi-Annual Progress Report No. 71 S.S.M.T.G,M.I.T. p. 13-14

- (1969).
- 4. E. Kmetko, Phys. Rev. A 1, 37 (1970).
 - 5. L. Hedin and S. Lundqvist, Solid State Physics 23, 1 (1969).
 - 6. D. Liberman, Phys. Rev. 171, 1 (1968).
 - 7. J. C. Slater, T. M. Wilson, and J. H. Wood, Phys. Rev. 179, 28 (1969).
 - 8. G. Burdick, Phys. Rev. 129, 138 (1963).

SINGLE-PARTICLE STATES IN MANY-BODY SYSTEMS^{*}

Seb Doniach

Department of Physics, Stanford University

Abstract

Relaxation of the wavefunctions of electrons in the region of a hole formed in an interband transition in a solid leads to a shift of the interband transition energy relative to that which would be predicted by a band structure calculation fit to fermi surface parameters. The theory of this relaxed orbital correction (ROC) is set up in terms of a Slater Koster model in which relaxed atomic and molecular cluster energies are introduced as parameters. Using this model a sum rule is found which gives the effect of band structure on the ROC for a core state hole. A perturbation theory is made of the effects of hole recoil for valence or d-band holes.

1. Introduction

There are currently two major classes of experiment used to test computed band structures in solids. These are fermi surface measurements in metals (e.g. the de Haas van Alphen effect), which essentially probe the ground state of the metal, and optical and photoemission experiments which involve interband transitions and consequently study properties of excited states.

The purpose of this paper is to point out that the essential physics leading to effective one-particle energies is different in these two types of experiment, so that band structure parameters determined by fitting to the two classes of experiment should not, in principle, be expected to agree. The difference arises as a result of the relaxation of the many electron system about the hole which is formed in an interband or photoemitting transition. This is not a new idea and has been quantitatively studied in the case of transitions between atomic levels in many electron atoms¹⁾. However

the effect has not been much discussed for energy band in solids²⁾¹²⁾

In this paper we present two new theoretical results directed at the extension of calculations of this type from atomic systems, with discrete 1-particle energy levels, to solids in which these levels are broadened into bands.

The finite band widths in a solid have two effects on the above relaxation effect, or "relaxed orbital correction" (ROC):

1) Suppose the hole state in question belongs to an inner core level. Then the relaxation of the outer (valence) electrons will be modified from that occurring in the atom by their finite bandwidth (i.e. by overlap with neighbors). To handle this we set up an effective 1-particle "impurity potential" for the solid (within the Hartree-Fock approximation) containing an atom with an ionized core state. In the context of this model Hamiltonian we show that the relaxed orbital correction may be exactly expressed in terms of an integral over the phase shifts for the valence (or conduction) electrons interacting with the impurity (analogous to the Friedel sum rule).

2) If the hole state is in a valence or inner band (such as a d-band) then the relaxation of the other electrons in the solid will be modified by the fact that the hole can now move through the solid so that the relaxation is no longer localized about a particular atom. If the coherence of this motion is neglected one reaches a limit, which has been discussed by Slater³⁾, in which $(1/N)^{th}$ of a hole is created on each atom in the crystal. In order to discuss effects of coherence, or partial localization, we develop a perturbation theory for the energy shift leading to an explicit formula for the dependence of the ROC on hole mass and hence allowing us to calculate the effects of coherence of the hole wave packet on the relaxation all the way from the infinite mass (core state) to the zero mass (incoherent) limits. (It turns out that the usual Breuckner-Goldstone form of perturbation theory fails in this case.)

2. Model Hamiltonian for Core State Hole

We set up the effects of the hole on the remaining electrons in the crystal in terms of the Hartree-Fock theory. Initially the ground state energy of the crystal may be written as a sum over the ground state configuration of Hartree-Fock energy levels, which we denote by $[N]$, as

$$E = \sum_N E_\mu[N] n_\mu \quad (1)$$

Here $E_\mu [N]$ are effective 1-particle energies (depending on the ground state configuration).

$$E_\mu [N] = \epsilon_\mu + \frac{1}{2} \sum_N V_{\mu\nu} n_\nu \quad (2)$$

where

$$\epsilon_\mu = \int d\mathbf{z}_1 \psi_\mu^*(\mathbf{z}_1) \left(-\nabla_1^2 + \sum_n \frac{2Z_n}{r_{1n}} \right) \psi_\mu(\mathbf{z}_1) \quad (3)$$

is the energy of the μ 'th electron moving in the field of the ions and

$$\begin{aligned} V_{\mu\nu} = & \int d\mathbf{z}_1 d\mathbf{z}_2 \left\{ \psi_\mu^*(\mathbf{z}_1) \psi_\mu(\mathbf{z}_1) \frac{1}{r_{12}} \psi_\nu^*(\mathbf{z}_2) \psi_\nu(\mathbf{z}_2) \right. \\ & \left. - \psi_\mu^*(\mathbf{z}_1) \psi_\nu(\mathbf{z}_1) \frac{1}{r_{12}} \psi_\nu^*(\mathbf{z}_2) \psi_\mu(\mathbf{z}_2) \right\} \end{aligned} \quad (4)$$

is the sum of coulomb and exchange potential energies. In a practical case, the exchange term would be approximated by something like a Kohn-Sham⁴⁾ potential.

Now suppose one of the core electrons on a given atom is removed. The system then adopts a new set of relaxed Hartree-Fock wavefunctions $\tilde{\psi}_\mu(\mathbf{z})$ with a new configuration $[N']$ (corresponding to the missing hole) and new ground state energy

$$\tilde{E} = \sum_{N'} \tilde{\epsilon}_\mu n_\mu + \frac{1}{2} \sum_{N'} V_{\mu\nu} n_\mu n_\nu \quad (5)$$

The resulting change in energy of the system is

$$\begin{aligned} \Delta E = E - \tilde{E} = & \sum_{N'} (\epsilon_\mu - \tilde{\epsilon}_\mu) n_\mu + \epsilon_\lambda \\ & + \frac{1}{2} \sum_{N'} (V_{\mu\nu} - \tilde{V}_{\mu\nu}) n_\mu n_\nu + \sum_N V_{\lambda\nu} n_\nu \end{aligned} \quad (6)$$

where λ refers to the particular hole state which has been removed. If the Hartree-Fock eigenvalues are denoted by E_μ^{HF} where

$$E_\mu^{\text{HF}} = \epsilon_\mu + \sum_N V_{\mu\nu} n_\nu \quad (7)$$

then (6) may be written as

$$\Delta E = E_{\lambda}^{\text{HF}} + E_{\text{ROC}} \quad (8)$$

where E_{ROC} is the correction to Koopman's theorem resulting from the orbital relaxation⁵⁾ and is given by

$$E_{\text{ROC}} = \sum_{N'} (\epsilon_{\mu} - \tilde{\epsilon}_{\mu}) n_{\mu} + \frac{1}{2} \sum_{N'} (V_{\mu\nu} - \tilde{V}_{\mu\nu}) n_{\mu} n_{\nu} \quad (9)$$

In order to make progress in estimating E_{ROC} for a solid we reexpress the initial ground state (1) and the final ground state (less the Hartree-Fock eigenvalue E_{λ}^{HF}) in terms of a Wannier function basis set leading to a Slater-Koster model Hamiltonian⁷⁾ for the initial crystal

$$H = \sum_{i,n} E_n c_i^+ c_i + \sum_{ij} c_{in}^+ c_{jm} H_{ij} \quad (10)$$

where i is a site suffix and n an orbital suffix. For insulators H_{ij} is (presumably) a reasonably short range overlap or tunnelling matrix element. In the final, relaxed state, of the crystal then, at least for insulating materials, only a few of the overlap matrix elements (for atoms in the neighborhood of the ionized atom) will be expected to change appreciably. The model Hamiltonian in the relaxed state now becomes

$$\tilde{H} = H + \sum_{in} \Delta_n c_i^+ c_i + \sum_{nn'} \Delta_{nn'} c_{en}^+ c_{e'n'} \quad (11)$$

where

$$\Delta_n = \tilde{E}_n - E_n, \quad \Delta_{nn'} = \tilde{H}_{nn'} - H_{nn'}$$

Of course writing this down really involves all the physics of the problem. However, if we make one key assumption, namely that the energy shifts Δ_n and the overlap shifts $\Delta_{nn'}$ can be reasonably calculated by solving the few atom problem of the "molecular cluster" surrounding the atom with the ionized core state, then it is possible to calculate the effect of all the remaining atoms of the crystal on the levels of the cluster. In this way the energy shift due to the relaxation of orbitals within the cluster is modified as a result of its effect on the one-electron levels within the energy band of the whole crystal.

The way this goes is by analogy with the Friedel sum rule. The shift terms in equation (5) act as a localized impurity potential (of matrix character) on the one-electron energy band states. Since the potential is localized the energy shift of each band state $\epsilon_{k,v}$ is of order $1/N$ (N the number of atoms in the crystal). However, since there are N levels in the band the total shift is of order unity. We show in the appendix that the sum of these shifts can be expressed in general in terms of a matrix

$$D(\epsilon)_{nn'} = \sum_n G^0(\epsilon)_{nn'} \Delta_{n,n'} \quad (12)$$

where G^0 is the 1-electron Green's function of the unperturbed Hamiltonian (4) and n refers to both site and band-index suffixes. Then the total energy shift on going from the ground state of H to the ground state of H' (eq. 5) is

$$E_{Roc} = \frac{1}{2\pi i} \int_{-\infty}^{\epsilon_F} \epsilon \frac{\partial}{\partial \epsilon} [\text{Disc} \left\{ \text{Trace} \log (1 - \frac{1}{\epsilon}) \right\}] \quad (13)$$

where Disc is the discontinuity for ϵ crossing the real energy axis.

The relation to the Friedel sum rule is seen by expressing (13) in terms of the phase shifts for band electron scattering from the cluster of atoms with relaxed orbitals, treating the band electrons as free electron-like. (As in the Friedel case). We show in the Appendix that (13) may reasonably be rewritten - the proof is actually only given for s-wave scattering -

$$E_{Roc} = \sum \epsilon_B + \frac{1}{\pi} \int_{-\infty}^{\epsilon_F} d\epsilon \sum_l (2l+1) \epsilon \frac{\partial \eta_l(\epsilon)}{\partial \epsilon} \quad (14)$$

where ϵ_B are any bound states which may be pushed right out of the band and $\eta_l(\epsilon)$ are the l-wave phase shifts.

How big is the effect of the band on the atomic shifts Δ ? To see this we take a very simple case in which only the relaxation of the levels on the central (hole) atom is considered. Then the scattering is that of the simplest Slater-Koster impurity problem⁸. The phase shift is pure s-wave and is given by

$$\delta_0(\epsilon) = \tan^{-1} \left[\frac{-\pi \Delta_0 N(\epsilon)}{1 - \Delta_0 F(\epsilon)} \right] \quad (15)$$

where $N(\epsilon)$ is the band density of states and $F(\epsilon)$ is its Hilbert transform. Then (6) can be estimated in two limits - that Δ_0 is either large or small compared to the band width. For $\Delta_0 \gg (\epsilon_{\max} - \epsilon_{\min})$ the main contribution comes from the bound state. (In this limit it is easy to show that this has energy = $\Delta_0 + \epsilon_{\min}$). The band contribution can be got on integration by parts as

$$E_{Roc}^{\text{band}} = \left[\epsilon \frac{\delta(\epsilon)}{\pi} \right]_{\epsilon_{\min}}^{\epsilon_{\max}} - \frac{1}{\pi} \int_{\epsilon_{\min}}^{\epsilon_{\max}} d\epsilon \tan^{-1} \left[N(\epsilon) \frac{\pi}{F(\epsilon)} \right] \quad (16)$$

The integral is zero for a symmetric band, while $\delta(\epsilon_{\max}) - \delta(\epsilon_{\min}) = \pi$ for one bound state. So the correction to Δ_0 is of order $(\epsilon_{\max} - \epsilon_{\min})$. For Δ_0 small, (6) can be expanded to give

$$\delta_0(\epsilon) \cong -\pi \Delta_0 N(\epsilon) - \pi \Delta_0^2 N(\epsilon) F(\epsilon) \quad (17)$$

from which

$$E_{Roc} \cong \Delta_0 \int_{-\infty}^{\epsilon_F} d\epsilon N(\epsilon) + \Delta_0^2 \int_{-\infty}^{\epsilon_F} d\epsilon d\epsilon' \frac{P}{\epsilon - \epsilon'} N(\epsilon) N(\epsilon') \quad (18)$$

For a simple band $\int N(\epsilon) d\epsilon$ is unity so for a full band

$$E_{Roc} \cong \Delta_0 [1 + \Delta_0 \langle N \rangle] \quad (19)$$

where

$$\langle N \rangle = \int_{-\infty}^{\epsilon_F} \frac{P}{\epsilon - \epsilon'} N(\epsilon) N(\epsilon') d\epsilon d\epsilon' \quad (20)$$

Thus in this limit of shift Δ_0 small compared to band width the atomic shift is modified by the coupling to the band through the factor $(1 + \Delta_0 \langle N \rangle)$.

3. Effect of Hole Recoil

The above discussion has been directed to the relaxed orbital correction resulting when a core state is ionized. If a band state is ionized, the argument will not work as it stands, as the change in band parameters in the region of the ionized atom is no longer localized at a given atomic site but will move through the lattice.

Even in the core state case it may be argued that an ionization event will couple to a Bloch state of the core level, rather than to a local state, as the electromagnetic field is essentially uniform in space when measured on the atomic scale.

In this section we show that (a) for the case of a deep core state the localized description is correct and (b) for the case of finite hole band width the perturbation theory of the relaxed orbital correction must be modified by the inclusion of hole recoil energy in the energy denominators, also that this perturbation theory contains anomalous terms (in the sense of Kohn and Luttinger).

The coupling to the electromagnetic field leads to an absorption rate given in terms of the current-current correlation function

$$R(t) = \langle j_q(t) j_q(0) \rangle \quad (21)$$

where q is the wave vector of the light (tends to zero). Writing

$$j_q = \sum_{p,\nu\nu'} j_{p,\nu\nu'} c_{p+q,\nu}^+ c_{p',\nu'}^- \quad (22)$$

$R(t)$ can be factored into a product of electron and hole propagators provided electron-hole interactions (exciton effects) are neglected⁹). For the core state we then have to deal with the core-hole propagator

$$\begin{aligned} g_{\text{core}}(t) &= \langle c_{p,\text{core}}^+(t) c_{p,\text{core}}(0) \rangle \\ &= \frac{1}{N} \sum_{ij} e^{i\vec{k} \cdot (\vec{R}_i - \vec{R}_j)} \langle c_{i,\text{core}}^+(t) c_{j,\text{core}}(0) \rangle \end{aligned} \quad (23)$$

where i, j are site suffixes. However in the limit that the core state band width tends to zero it is clear that the hole cannot propagate from one core to another (provided we neglect lifetime effects due to x-ray emission or Auger effect). Hence

$$\begin{aligned} \langle c_i^+(t) c_j(0) \rangle &= \delta_{ij} \langle c_i^+(t) c_i(0) \rangle \\ &= g_{\text{core}}(t) \end{aligned} \quad (24)$$

In real life the emission and re-absorption of virtual x-rays or Auger 2-electron excitations will allow the core-hole to diffuse through the crystal. However it seems likely that this diffusion

will be incoherent in nature. It will lead to a broadening of the absorption line in addition to the shift we are studying here. For the purpose of the present discussion these effects will be neglected.

In the finite band width case (finite hole mass) we have

$$g_{p,v}(t) = \langle c_{p,v}^+(t) c_{p,v}(0) \rangle$$

which can be expanded in a complete set of N-electron and N-1 electron states

$$g_{p,v}(t) = \sum_{N,N-1} \left| \langle N | c_{p,v}^+ | N-1 \rangle \right|^2 e^{i(E_N - E_{N-1})t} \quad (25)$$

In the case of a hole at a band maximum or minimum ($p=0$) the spectral density¹⁰ of (25) will cut off at the ground state of the (N-1) electron system with the hole frozen in the v-band giving

$$E_{\text{threshold}} = (E_N^G - E_{N-1}^G) \quad (26)$$

Strictly speaking the (N-1) electron system will not be stable with the hole in the v-band, but we neglect lifetime effects for the moment.

$E_{N-1}^G(v)$ may be calculated by using the linked cluster expansion at finite temperature. We just write down the first two terms for a model in which the energy levels of a single valence or conduction band relax in the presence of the hole. The relaxed orbital correction then becomes to second order in Δ_0

$$\begin{aligned} E_{\text{ROC}} &= \frac{\Delta_0}{N} \sum_p \int_c d\varepsilon G_c(\varepsilon) + \left(\frac{\Delta_0}{N} \right)^2 \sum_{pq} \int d\varepsilon G_p(\varepsilon) G_q(\varepsilon + E_q) \\ &= \Delta_0 + \left(\frac{\Delta_0}{N} \right)^2 \sum_{pq} \left\{ \frac{f_p(1-f_{p+q})}{\varepsilon_p - \varepsilon_{p+q} - E_q} + \frac{f_{p+q}(1-f_p)}{\varepsilon_p - \varepsilon_{p+q} + E_q} \right\} \end{aligned} \quad (27)$$

where E_q is the energy of hole in state q corresponding to the diagrams in Fig. 1.

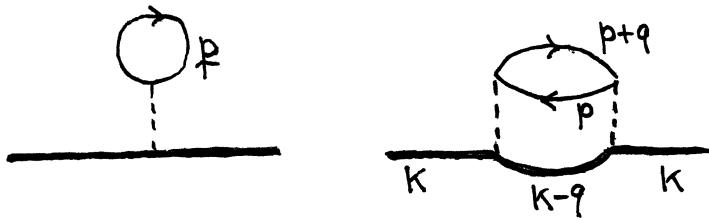


Fig. 1: Lowest order diagrams for relaxation of electron states in presence of hole.

Note: These are Feynman type diagrams (all time orderings are to be included).

Equation (27) may be rearranged to give

$$E_{Roc} = \Delta_0 + \left(\frac{\Delta_0}{N}\right)^2 \sum_{pp'} f_p \left[\frac{\epsilon_p - \epsilon_{p'}}{(\epsilon_p - \epsilon_{p'})^2 + E_{p-p'}} \right] \quad (28)$$

In the limit of zero hole band width this reduces to the weak coupling formula (18). This result is an extension to the solid of the result of Hedin and Johannson.

Formula (28) shows that as the hole band width becomes larger than that of the band which is relaxing, the band correction to E_{Roc} is reduced. In the limiting case of $E_q \gg (\epsilon_p - \epsilon_{p+q})$ we can expand to give

$$E_{Roc} \approx \Delta_0 \left(1 + \frac{\Delta_0}{N} \sum_{q,p} \left(\frac{\epsilon_p - \epsilon_{p+q}}{E_q^2} \right) \right) \quad (29)$$

Using an effective mass approximation for E

$$E_q = q^2/2M_{\text{hole}}$$

we can estimate by expanding the numerator in powers of q to give

$$\begin{aligned} (E_{\text{ROC}} - \Delta_0) &\cong \frac{\Delta_0}{N} \sum_q \left(\frac{2M_{\text{hole}}}{q^2} \right)^2 \left(\frac{\partial^2 E_p}{\partial p^2} \right) \quad (30) \\ &\cong \left(\frac{\Delta_0}{W_{\text{hole}}} \right) \left(\frac{M_{\text{hole}}}{m^*_{\text{elec}}} \right) \Delta_0 \end{aligned}$$

where $W_{\text{hole}} = \frac{q_{\text{max}}^2}{2M_{\text{hole}}}$ is of order of the hole band width. Thus as $M_{\text{hole}}/m^*_{\text{elec}}$ gets smaller, i.e. as the hole starts to recoil faster relative to the tunnelling rate of the relaxing electron band, the band correction to the atomic value Δ_0 of the relaxed orbital shift, gets small.

Acknowledgments I am very grateful to Frank Hermann for a most useful discussion and to Walt Harrison for some helpful comments.

Appendix: derivation of sum rule formula.

The change in ground state energy induced by the orbital relaxation effect is expressed in equation (11) in terms of a localized shift of Slater-Koster parameters about the ionized atom. Introducing a Green's function (in Wannier representation) for the shifted Hamiltonian

$$\tilde{G}_{nn'}(t) = -i \langle T \{ \tilde{c}_n(t) \tilde{c}_{n'}^\dagger(0) \} \rangle \quad (A1)$$

where $\tilde{c}_n(t) = e^{i\tilde{H}t} c_n e^{-i\tilde{H}t}$, we have

$$E_{\text{ROC}} = \langle \tilde{H} - H \rangle = \lim_{t \rightarrow 0} -i \left[\sum_{ll'} t_{ll'} (\tilde{G}_{ll'} - G_{ll'}) + \sum_{nn'} \Delta_{nn'} \tilde{G}_{nn'} \right] \quad (A2)$$

where we have rewritten the unperturbed model Hamiltonian as

$$H = \sum_{ll'} t_{ll'} c_l^\dagger c_{l'} \quad (A3)$$

In terms of the Fourier transform of G this may be written

$$E_{ROC} = \frac{1}{2\pi i} \int_C d\varepsilon \left\{ \sum_{ij} t_{ij} G_{in} (1 - \frac{\Delta}{D})_{nn'}^{-1} \Delta_{n'n''} G_{n''j} + \sum_{nn'} \Delta_{nn'} (1 - \frac{\Delta}{D})_{nn'}^{-1} G_{n''n'} \right\} \quad (A4)$$

where D is defined in equation (12) and C is a contour surrounding energy states below the fermi level. Introducing a complete set of Bloch states for H we have (k denotes both wave vector and band suffixes)

$$G_{ll'}(\varepsilon) = \sum_k \langle l|k\rangle \frac{1}{\varepsilon - \varepsilon_k} \langle k|l'\rangle \quad (A5)$$

we find, after some algebra,

$$\begin{aligned} E_{ROC} &= \frac{1}{2\pi i} \int_C d\varepsilon \sum_{nn'k} \langle n'|k\rangle \frac{\varepsilon}{(\varepsilon - \varepsilon_k)} \langle k|n\rangle \left[(1 - \frac{\Delta}{D})_{nn'}^{-1} \right]_{nn'} \\ &= \frac{1}{2\pi i} \int_C d\varepsilon \varepsilon \sum \left(\frac{-\partial}{\partial \varepsilon} G_{nn'} \right) \left[(1 - \frac{\Delta}{D})_{nn'}^{-1} \right]_{nn'} \\ &= \frac{1}{2\pi i} \int_C d\varepsilon \varepsilon \frac{\partial}{\partial \varepsilon} \text{Trace} \left[\log (1 - \frac{\Delta}{D}) \right] \end{aligned} \quad (A6)$$

which is equivalent to (13). To see this in terms of phase shifts we go to the single atom level-shift limit of equation (15). Then (13) becomes

$$E_{ROC} = \frac{D_{disc}}{2\pi i} \int_{-\infty}^{\varepsilon_F} d\varepsilon \varepsilon \frac{\partial}{\partial \varepsilon} \left[\log (1 - \Delta_0(F(\varepsilon)) \pm i\pi N(\varepsilon)) \right] \quad (A7)$$

($\pm i\pi N(\varepsilon)$ occur above and below the energy axis)

$$\begin{aligned} &= \int_{-\infty}^{\varepsilon_F} d\varepsilon \varepsilon \frac{\partial}{\partial \varepsilon} \left[\tan^{-1} \left(-\frac{\Delta_0 \pi N(\varepsilon)}{1 - \Delta_0 F(\varepsilon)} \right) \right] \\ &= \int_{-\infty}^{\varepsilon_F} d\varepsilon \varepsilon \frac{\partial \eta_0(\varepsilon)}{\partial \varepsilon} \end{aligned} \quad (A8)$$

If a bound state occurs there will be an additional residue at the bound state energy. For general $\Delta_{nn'}$ it seems likely that

This will generalize to equation (14), provided the unperturbed Green's function G_{ij} behaves free-electron-like for large $(R_i - R_j)$. For a general energy band the phase shift approach does not really mean very much and a more general approach has to be worked out¹¹).

References and footnotes

- *) Research supported by Army Research Office, Durham and by ARPA through the Center for Materials Research at Stanford University.
- 1) L. Hedin and B. Johannson, J. Phys. C. (Phys. Soc. London) Ser. 2 Vol. 2, 1336 (1969) and references given therein.
- 2) See, however, the discussion by Prof. Slater at this conference.
- 3) J. C. Slater MIT semi-annual progress report #71 (Solid state and molecular theory group) July 15, 1969.
- 4) W. Kohn and L. J. Sham Phys. Rev. 140 A1133 (65).
- 5) These corrections should not be confused with the "Koopman's⁶) corrections" introduced by Hermann, Ortenberger and Van Dyke to account for the breakdown of Koopman's theorem on using a free-electron exchange ($\rho^{1/3}$ law) within the context of a frozen orbital treatment.
- 6) F. Hermann, I. B. Ortenburger, J. P. Van Dyke, Intl. J. Quantum Chemistry III S. 827 (70).
- 7) J. C. Slater and G. F. Koster, Phys. Rev. 94, 1498 (54).
- 8) G. F. Koster and J. C. Slater, Phys. Rev. 96, 1208 (55).
- 9) S. Doniach and M. Sunjic, J. Phys. C. (Phys. Soc. London) 3, 284 (1970).
- 10) S. Doniach, Phys. Rev., submitted for publication.
- 11) R. Harris, J. Phys. C. (Phys. Soc. London) 3, 172 (1970).
- 12) I am grateful to David Beaglehole for pointing out to me that this effect was discussed by J. Friedel, Proc. Phys. Soc. (London), B65, 769 (1952) who estimated that E_{ROC} is of order 0.5 ev for d-band holes in Cu.

TOWARDS SELF CONSISTENCY WITH THE TIGHT BINDING APPROXIMATION*

Joseph Callaway and John L. Fry

Dept. of Physics & Astronomy, Louisiana State University
Baton Rouge, Louisiana 70803

ABSTRACT

A procedure is described which should make self consistent energy band calculations possible within the framework of the tight binding approximation. The essential concept of the method is that the same set of integrals is required for construction of an iterated crystal potential as is required in the computation of matrix elements of the Hamiltonian in the band calculation.

INTRODUCTION

The tight binding method for energy band calculations has generally been regarded as suitable primarily for obtaining a simple first approximation to a complex band structure, or as an interpolation scheme to supplement more refined methods. It should be apparent from the success obtained in some applications to materials with nearly free like band structures^{1,2} that the method should also be quite powerful in quantitative calculations from first principles for a wide variety of materials. In this note, we describe a procedure by which self consistent band calculations may be performed using the tight binding method. This appears to be a practical procedure, since it is possible to obtain the Fourier coefficients of potential resulting from some stage of the iteration procedure in terms of the same integrals employed to determine the matrix elements of the Hamiltonian. If a fixed set of basis functions is employed, these integrals need to be computed only once.

PROCEDURE

We will outline the essential features of the self consistent procedure. The fundamental problem is to determine a new (or iterated) potential after a given stage of the band structure calculation has been completed. We may therefore suppose that we are in possession of a definite set of energy bands $E_n(\underline{k})$ and Bloch functions $\psi_n(\underline{k}, \underline{r})$ which have been obtained by diagonalization of the Hamiltonian and overlap matrices resulting from some assumed set of wave functions. The Bloch functions are

$$\psi_n(\underline{k}, \underline{r}) = \sum_i a_{ni}(\underline{k}) \varphi_i(\underline{k}, \underline{r}) \quad (1)$$

where the $\varphi_i(\underline{k}, \underline{r})$ are tight binding functions for wave vector \underline{k} and the a_{ni} are coefficients determined by the diagonalization process:

$$\varphi_i(\underline{k}, \underline{r}) = \frac{1}{\sqrt{N}} \sum_{\mu} e^{i \underline{k} \cdot \underline{R}_{\mu}} u_i(\underline{r} - \underline{R}_{\mu}) . \quad (2)$$

The summation runs over all N lattice sites \underline{R}_{μ} . For convenience, we have assumed a monoatomic crystal. The functions $u_i(\underline{r} - \underline{R}_{\mu})$ are localized functions centered on lattice site \underline{R}_{μ} . In order to obtain enough flexibility in the basis set so that the matrix diagonalization will yield an adequate solution of the Schrodinger equation with the assumed potential, the $u_i(\underline{r})$ will be taken to be individual Slater type orbitals (rather than complete atomic functions, which are linear combinations of several of the u_i). Gaussian orbitals may also be satisfactory.

It is not necessary to exhibit the potential in position space. Since the method of determining the matrix elements of the Hamiltonian developed by Lafon and Lin² requires the Fourier coefficients of potential, $V(\underline{K})$ [where \underline{K} is a reciprocal lattice vector] it suffices to obtain these. For $\underline{K} \neq 0$, the Fourier coefficients of the Coulomb potential are related to the corresponding coefficients of the total charge density $\rho_T(\underline{K})$ by³

$$V(\underline{K}) = -8\pi \frac{\rho_T(\underline{K})}{\underline{K}^2} \quad (3)$$

(in atomic units, with energies in Rydbergs). The Fourier coefficients $\rho(\underline{K})$ of the electron charge density are given by

$$\rho(\underline{K}) = \frac{1}{N\Omega} \int \rho(\underline{r}) e^{i \underline{K} \cdot \underline{r}} d^3 r \quad (4)$$

in which Ω is the volume of the unit cell, and

$$\rho(\underline{r}) = \sum_{n, \underline{k} \text{ occ}} |\psi_n(\underline{k}, \underline{r})|^2 . \quad (5)$$

We substitute (1), (2), and (5) into (4); and convert the sum on \underline{k} to an integral. The result is

$$\rho(\underline{k}) = \frac{1}{(2\pi)^3} \sum_{n, \underline{k}, j} \int d^3k a_{ni}^*(\underline{k}) s_{ij}(\underline{k}, \underline{k}) a_{nj}(\underline{k}) . \quad (6)$$

The integral includes that portion of the Brillouin zone in which band n is occupied. The quantities s_{ij} are generalized overlap matrix elements

$$s_{ij}(\underline{k}, \underline{k}) = \sum_{\sigma} e^{i\underline{k} \cdot \underline{R}\sigma} \int u_i^*(\underline{r}) e^{-i\underline{k} \cdot \underline{r}} u_j(\underline{r} - \underline{R}\sigma) d^3r . \quad (7)$$

Equations (3), (6), and (7) determine the required $V(\underline{k})$. The integrals which appear in (7) are exactly those which are required in the computation of the matrix elements of the Hamiltonian on the basis of the functions $\varphi_i(\underline{k}, \underline{r})$. Specifically, we have

$$\begin{aligned} & \int \varphi_i^*(\underline{k}, \underline{r}) V(\underline{r}) \varphi_u(\underline{k}, \underline{r}) d^3r = \\ &= \sum_{s, \sigma} V(\underline{k}_s) e^{i\underline{k} \cdot \underline{R}\sigma} \int u_i(\underline{r}) e^{i\underline{k}_s \cdot \underline{r}} u_j(\underline{r} - \underline{R}\sigma) d^3r \\ &= \sum_s V(\underline{k}_s) s_{i,j}(\underline{k}, -\underline{k}_s) \end{aligned} \quad (8)$$

in which $V(\underline{r})$ is the periodic crystal potential. If a fixed set of basis functions is employed, the quantities s_{ij} need be computed only once.

In the case $K = 0$, the Fourier coefficient of potential must be determined by a limiting process

$$V(0) = -8\pi \lim_{K \rightarrow 0} \frac{\rho_T(\underline{k})}{K^2} \quad (9)$$

The limit exists, and may be expressed as

$$V(0) = \frac{1}{6\pi^2} \sum_{nij} \int a_{ni}^*(\underline{k}) s_{ij}^{(2)}(\underline{k}) a_{nj}(\underline{k}) d^3k \quad (10)$$

where

$$s_{ij}^{(2)}(\underline{k}) = \sum_{\sigma} e^{ik \cdot R_{\sigma}} \int u_i^*(\underline{r}) r^2 u_j(\underline{r-R_{\sigma}}) d^3 r. \quad (11)$$

The exchange potential presents more difficulties. We suppose that the Slater "Xα" approach is used.⁴ The exchange potential is then proportional to $\rho^{1/3}(\underline{r})$. In the initial stage of the calculation, it appears to be necessary to obtain the explicit function $\rho(\underline{r})$. However, assuming that the changes in Fourier coefficients are small in subsequent steps, we may proceed as follows: Let $\rho^{1/3}(\underline{r})$ be the cube root of the charge density at the present stage of iteration, and let $\rho_0^{1/3}(\underline{r})$ be the same quantity at the preceding stage. We write

$$\rho^{1/3}(\underline{r}) = \rho_0^{1/3}(\underline{r}) + \Delta(\underline{r}). \quad (12)$$

$\Delta(\underline{r})$ is expanded in a Fourier series, since it is the Fourier coefficients which will actually be used:

$$\Delta(\underline{r}) = \sum_s \Delta(\underline{k}_s) e^{i \underline{k}_s \cdot \underline{r}}. \quad (13)$$

To determine $\Delta(\underline{k})$, we expand $\rho(\underline{r})$ and $\rho_0(\underline{r})$ in a Fourier series, and extract the cube root by expansion. The result is

$$\begin{aligned} \Delta(\underline{k}) = & \frac{1}{3} \left(\frac{\Omega}{n_e} \right)^{2/3} \{ \rho(\underline{k}) - \rho_0(\underline{k}) - \frac{2}{3} \left(\frac{\Omega}{n_e} \right)^{1/3} \sum_{t \neq 0} [\rho(\underline{k}_t) \rho(\underline{k}-\underline{k}_t) \\ & - \rho_0(\underline{k}_t) \rho_0(\underline{k}-\underline{k}_t)] \} \end{aligned} \quad (14)$$

where we have retained only second order terms. In this equation n_e is the number of electrons in the cell. In this case, the $K=0$ limit is simple. It is a consequence of the use of normalized Bloch functions that

$$\rho(0) = \frac{n_e}{\Omega} \quad (15)$$

at each stage of the iteration. Consequently

$$\Delta(0) = -\frac{2}{9} \left(\frac{\Omega}{n_e}\right)^{5/3} \sum_t [\rho(k_t) \rho(-k_t) - \rho_o(k_t) \rho_o(-k_t)]. \quad (16)$$

The utility of these approximate formulas for $\Delta(k)$ is dependent on the assumption that the changes in the Fourier coefficients of ρ are small.

REFERENCES

*Supported in part by the U.S. Air Force Office of Scientific Research under Grant AFOSR-68-1565.

1. F. Stern, Ph.D. Thesis, Princeton University, (1956) (unpublished).
2. E. E. Lafon and C. C. Lin, Phys. Rev. 152, 579 (1966).
3. J. Callaway, "Energy Band Theory", Academic Press, New York, 1964, p. 64.
4. J. C. Slater, T. M. Wilson, and J. H. Wood, Phys. Rev. 179, 28 (1969).

TOWARD HARTREE-FOCK CALCULATIONS FOR SIMPLE CRYSTALS

Frank E. Harris and Hendrik J. Monkhorst

Department of Physics, University of Utah

Salt Lake City, Utah 84112

ABSTRACT

Hartree-Fock equations are formulated for LCAO wavefunctions whose atomic-orbital composition may depend upon location in the energy band. Computationally practical procedures are described for all quantities entering the formulation, and preliminary results for a simple-cubic atomic-hydrogen lattice are quoted to illustrate the convergence properties of the method. The procedure depends crucially upon reciprocal-lattice transformations and the point-group symmetry about a lattice point; exploitation of these concepts helps provide a natural understanding of the success of independent-particle schemes for solids, and of the structure of the exchange energy contributions.

INTRODUCTION

For many reasons it would be desirable to have accurate Hartree-Fock electronic wavefunctions for at least a few representative crystalline solids. Such wavefunctions would be unlikely to be entirely satisfactory in describing the experimental situation due to the neglect of electron correlation and to the assumption of identical band occupancies for electrons of both spin orientations, but nevertheless Hartree-Fock functions could yield much insight into the methods now used in energy band calculations. In particular, one might hope to learn more about optimum atomic orbitals for use in solids, and to compare Hartree-Fock functions for various systems against the limiting cases exemplified by free-electron and tight-binding methods. One could compare Hartree-Fock and other approximations to the Fermi

surface, and could study the effects of various models for the exchange energy.

Most of the studies aiming toward self-consistent descriptions of electronic states of solids have introduced local operators to approximate the exchange contribution to the energy. This permits the electron-electron interaction to be described in terms of the charge density. A few investigators have considered the explicit inclusion of coulomb and exchange-type electron-repulsion contributions, but that approach leads to large numbers of multicenter integrals. The approach to be described here, involving a systematic exploitation of the properties of transformations to reciprocal space, has been less used, perhaps because of uncertainties as to the rapidity with which the thereby-produced expansions converge.

This paper will describe in detail a method for setting up and solving the Hartree-Fock equations for a simple crystal. As indicated above, the method is based on the systematic use of Fourier-transform formulations. We are not yet reporting a full Hartree-Fock calculation, but we do give data which indicate the satisfactory convergence rate for the reciprocal-space expansions. The transformations to reciprocal space remove entirely the need to evaluate difficult multi-center integrals, as it is then possible to perform the lattice sums before considering the integrations. The result is that lattice orthogonality relations introduce delta functions and the integrals reduce to reciprocal-lattice sums. The Fourier-transform formulation also yields a natural and precise understanding of the cancellation of opposing long-range contributions to the electrostatic energy.

Our numerical calculations, and for simplicity the remainder of this paper, are specialized to an atomic hydrogen crystal in a simple cubic lattice at 0°K, i.e. to a simple cubic system with a proton at each lattice point and an equal number of electrons optimally distributed in the crystal. We assume the electrons to be assigned pairwise to the Bloch-wave states built from atomic functions, but we allow the atomic functions to be optimized individually for each Bloch-wave state.

We have found it convenient to cast the reciprocal-space expansions in terms of stars of lattice points, a star consisting of a set of points interchanged by the operations of the point group about the lattice origin. We describe the Fermi surface, the coefficients defining the atomic functions, and the kernel of the exchange energy expression by expansions in cubic harmonics. The coordinated use of these symmetry concepts simplifies the organization of the calculations and makes explicit the favorable convergence properties. In particular, we are able to identify the exchange energy as the sum of an essentially free-

electron term, a term exactly proportional to the Coulombic electron-electron repulsion, and a small "correction" term.

PROBLEM DEFINITION

Consider a simple cubic lattice with lattice constant a . The lattice is assumed to contain a very large number N of lattice points \vec{R}_μ , at each of which is fixed a proton. A total of N electrons are assigned in opposite-spin pairs to crystal orbitals $|\vec{k}\rangle$ of the Bloch form, with \vec{k} designating the Bloch-wave vector. When necessary to be explicit, we let \vec{r} designate the spatial coordinates upon which $|\vec{k}\rangle$ depends. The crystal orbitals are constructed from sets of atomic orbitals ϕ_i ($i=1, 2, \dots, M$) centered at each lattice point \vec{R}_μ . The coefficients of the ϕ_i in $|\vec{k}\rangle$ are assumed to depend not only on a Bloch-wave phase factor but also upon \vec{k} through coefficients $c_i(\vec{k})$. The specific form to be used for $|\vec{k}\rangle$ is

$$|\vec{k}\rangle = \sum_{i=1}^M c_i(\vec{k}) |\vec{k}_i\rangle \quad (1)$$

with

$$|\vec{k}_i\rangle = e^{2\pi i a^{-1} \vec{k} \cdot \vec{r}} \sum_\mu \phi_i(\vec{r} - a\vec{\mu}) \quad (2)$$

As shown in Eq. (2), the symbol $|\vec{k}_i\rangle$ refers to a Bloch function of wave vector \vec{k} (in reciprocal-lattice units) built from atomic orbital ϕ_i , while the unsubscripted $|\vec{k}\rangle$ designates the entire crystal orbital. Throughout this paper, summations with vector indices are to extend over N points of a simple cubic lattice of unit lattice constant, so that $a\vec{\mu}$ is synonymous with lattice point \vec{R}_μ . When the point $\vec{\mu} = 0$ is to be omitted, a prime is placed on the summation.

The form given in Eqs. (1) and (2) differs slightly from that most frequently encountered, and from that used in previous work of the present authors¹, through the presence of the factor $\exp(2\pi i a^{-1} \vec{k} \cdot \vec{r})$ rather than factors $\exp(2\pi i a^{-1} \vec{k} \cdot \vec{R}_\mu)$ for the individual terms of the sum. The present formulation is appreciably easier to apply, and there appear to be no compelling physical reasons why the more customary form should be preferred. We do note that Eq. (2) is not periodic in \vec{k} , by which we mean that \vec{k} values outside the first Brillouin zone do not describe wavefunctions identical with those associated with \vec{k} values inside that zone. For \vec{k} values within the first Brillouin zone, Eq. (2) describes functions whose \vec{r} dependence cycles smoothly through complex values rather than exhibiting partial or complete interference to produce nodes or minima between lattice points. However, this obvious difference is of little direct significance as a more nodal set of orbitals may be obtained by linear combinations

of $|\vec{k}_i\rangle$, e.g. $|\vec{k}_i\rangle \pm |(-\vec{k})_i\rangle$, and transformations of this sort do not alter the N-electron wavefunction. We also note that the flexibility associated with the use of \vec{k} -dependent coefficients in Eq. (1) can temper the effect of particular assumptions for Eq.(2). The form given in Eq. (2) has also been used by Abrikosov².

The Hamiltonian operator H to be considered here is in a fixed-nucleus and nonrelativistic limit, and consists of electronic kinetic energies and electrostatic interactions among all charges. We write H (in atomic units) as a sum of one-electron and two-electron terms in the following way:

$$H = \sum_{i=1}^N (-\frac{1}{2}\nabla_i^2) + \sum_{1 \leq i < j \leq N} h(\vec{r}_i, \vec{r}_j), \quad (3)$$

where

$$\begin{aligned} h(\vec{r}_1, \vec{r}_2) = & r_{12}^{-1} - \frac{1}{N} \sum_{\mu} |\vec{r}_1 - \vec{R}_{\mu}|^{-1} - \frac{1}{N} \sum_{\mu'} |\vec{r}_2 - \vec{R}_{\mu'}|^{-1} \\ & + \frac{1}{N^2} \sum_{\substack{\mu, \mu' \\ (\mu \neq \mu')}} |\vec{R}_{\mu} - \vec{R}_{\mu'}|^{-1} \end{aligned} \quad (4)$$

Equations (3) and (4) are exact except for terms whose contributions to the total energy are of order less than N . The expectation value E of the total energy then assumes the form

$$\begin{aligned} E = & 2N \int d\vec{k} \frac{\langle \vec{k} | -\frac{1}{2}\nabla^2 | \vec{k} \rangle}{\langle \vec{k} | \vec{k} \rangle} \\ & + 2N^2 \int d\vec{k} d\vec{k}' \frac{(\langle \vec{k} \vec{k}' | h | \vec{k} \vec{k}' \rangle - \frac{1}{2} \langle \vec{k} \vec{k}' | r_{12}^{-1} | \vec{k}' \vec{k} \rangle)}{\langle \vec{k} | \vec{k} \rangle \langle \vec{k}' | \vec{k}' \rangle} \end{aligned} \quad (5)$$

Equation (5) results from the assumption of doubly-occupied crystal orbitals and the requirement of antisymmetry for the crystal wavefunction. The last term is the exchange energy; only the r_{12}^{-1} part of h leads therein to a nonvanishing result. The integrals of \vec{k} and \vec{k}' are over the region enclosed by the Fermi surface; in reciprocal-lattice units this region has volume $\frac{1}{2}$. Except when noted otherwise, all reciprocal-space integrals in this paper will be over this region.

It is now convenient to rewrite Eq. (5) to obtain E in terms of the Bloch waves $|\vec{k}_i\rangle$ and the coefficients $c_i(\vec{k})$. Substituting from Eq. (1), and introducing the restriction that the $c_i(\vec{k})$ have been scaled so that $\langle \vec{k} | \vec{k} \rangle = N$ for all \vec{k} ,

$$\begin{aligned}
 E = & 2 \int d\vec{k} \sum_{ij} c_i^*(\vec{k}) c_j(\vec{k}) \langle \vec{k}_i | -\frac{1}{2} \nabla^2 | \vec{k}_j \rangle \\
 & + 2 \int d\vec{k} d\vec{k}' \sum_{ijmn} c_i^*(\vec{k}) c_j(\vec{k}) c_m^*(\vec{k}') c_n(\vec{k}') \cdot \\
 & \cdot (\langle \vec{k}_i \vec{k}'_m | h | \vec{k}_j \vec{k}'_n \rangle - \frac{1}{2} \langle \vec{k}_i \vec{k}'_m | r^{-1}_{12} | \vec{k}'_n \vec{k}_j \rangle)
 \end{aligned} \quad (6)$$

The scaling restriction may be written

$$S(\vec{k}) = \sum_{ij} c_i^*(\vec{k}) c_j(\vec{k}) \langle \vec{k}_i | \vec{k}_j \rangle = N \quad (7)$$

To form the Hartree-Fock equations, we seek to minimize the expression for E by variation of the $c_i(\vec{k})$ and of the Fermi surface, subject to the constraint given in Eq. (7) and to the requirement that the Fermi surface enclose volume $\frac{1}{2}$. Using the method of Lagrangian multipliers, we consider the variation $\delta(E - 2 \int \epsilon(\vec{k}) S(\vec{k}) d\vec{k})$, where $\epsilon(\vec{k})$ plays the role of a multiplier. The result may be written in the standard form

$$\sum_j F_{ij}(\vec{k}) c_j(\vec{k}) = \epsilon(\vec{k}) \sum_j S_{ij} c_j(\vec{k}), \quad (8)$$

where

$$\begin{aligned}
 F_{ij}(\vec{k}) = & N^{-1} \left[\langle \vec{k}_i | -\frac{1}{2} \nabla^2 | \vec{k}_j \rangle + 2 \sum_{mn} \int d\vec{k}' c_m^*(\vec{k}') c_n(\vec{k}') \cdot \right. \\
 & \cdot \left. (\langle \vec{k}_i \vec{k}'_m | h | \vec{k}_j \vec{k}'_n \rangle - \frac{1}{2} \langle \vec{k}_i \vec{k}'_m | r^{-1}_{12} | \vec{k}'_n \vec{k}_j \rangle) \right], \quad (9)
 \end{aligned}$$

$$S_{ij} = N^{-1} \langle \vec{k}_i | \vec{k}_j \rangle \quad (10)$$

The variation with respect to the Fermi surface leads to the requirement that $\epsilon(\vec{k})$ be constant on the surface; this corresponds to Koopman's theorem but here is exact because of the possibility of continuous variation in \vec{k} .

The Hartree-Fock equations may be solved iteratively, at each iteration determining approximations to the $c_i(\vec{k})$ and $\epsilon(\vec{k})$, and therefrom an approximation to the Fermi surface. The entire process becomes more manageable if expansions are introduced for these quantities, and we have chosen to write them in terms of cubic harmonics³. Accordingly, we introduce $T_{\mu}, \mu = 1, 2, \dots$ to stand for normalized harmonics of the totally symmetric representation of the cubic point group, as described in more detail in Appendix I.

We then write

$$c_j(\vec{k}) = \sum_{p\mu} c_{p\mu}^j k^{p_T}_\mu(\Omega_k) \quad (11)$$

$$\epsilon(\vec{k}) = \sum_{p\mu} \epsilon_{p\mu} k^{p_T}_\mu(\Omega_k) \quad (12)$$

$$f(\Omega_k) = \sum_\mu f_\mu T_\mu(\Omega_k) \quad (13)$$

where k and Ω_k stand for the magnitude and angular coordinates of \vec{k} and $f(\Omega_k)$ is the k coordinate of the Fermi surface in the direction Ω_k .

In the preliminary studies now under way, we are producing $\epsilon(\vec{k})$ and $c_j(\vec{k})$ at specific points \vec{k} , are fitting these values to the expansions given in Eqs. (11) and (12), and are then determining the f_μ of Eq. (13) from the $\epsilon(\vec{k})$ expansion. Thus the description of our current procedures will be essentially completed by a discussion of the means for evaluating the quantities entering Eqs. (9) and (10). These topics comprise the remainder of this paper.

MATRIX ELEMENTS

Our first main step in the evaluation of $F_{ij}(\vec{k})$ and S_{ij} is to use Fourier representation theory to express the Bloch-wave matrix elements as reciprocal-lattice sums. For all but the kinetic-energy integrals, the orbitals enter only as Fourier transforms of lattice sums of orbital products:

$$\phi_{ij}(\vec{q}) = \sum_{\mu} \langle \phi_i(\vec{r}) | e^{2\pi i a^{-1} \vec{q} \cdot \vec{r}} | \phi_j(\vec{r} - a\vec{\mu}) \rangle \quad (14)$$

Here \vec{q} is the transform variable, expressed in reciprocal lattice units.

The overlap integrals reduce as shown below.

$$\begin{aligned} \langle \vec{k}_i | \vec{k}_j \rangle &= \sum_{\lambda\mu} \langle e^{2\pi i a^{-1} \vec{k} \cdot \vec{r}} \phi_i(\vec{r} - a\vec{\lambda}) | e^{2\pi i a^{-1} \vec{k} \cdot \vec{r}} \phi_j(\vec{r} - a\vec{\mu}) \rangle \\ &= N \sum_{\mu} \langle \phi_i(\vec{r}) | \phi_j(\vec{r} - a\vec{\mu}) \rangle = N \phi_{ij}(0). \end{aligned} \quad (15)$$

The factor N arises when the lattice translational symmetry is used to remove the $\vec{\lambda}$ summation. We see that $\langle \vec{k}_i | \vec{k}_j \rangle$ is related to ϕ_{ij} evaluated at $\vec{q} = 0$. Noting the definition of S_{ij} , we also see that

$$S_{ij} = \phi_{ij}(0) \quad (16)$$

Next consider the kinetic-energy integrals. Writing

$$\begin{aligned} -\frac{1}{2}\nabla^2 |\vec{k}_j\rangle &= e^{2\pi i a^{-1}\vec{k}\cdot\vec{r}} \sum_{\vec{\mu}} (-\frac{1}{2}\nabla^2) \phi_j(\vec{r} - a\vec{\mu}) + 2\pi^2 a^{-2} k^2 |\vec{k}_j\rangle \\ &\quad - 2\pi i a^{-1} e^{2\pi i a^{-1}\vec{k}\cdot\vec{r}} \sum_{\vec{\mu}} \vec{k}\cdot\nabla \phi_j(\vec{r} - a\vec{\mu}) \end{aligned} \quad (17)$$

we find

$$\langle \vec{k}_i | -\frac{1}{2}\nabla^2 | \vec{k}_j \rangle = N(T_{ij} + \frac{2\pi^2 k^2}{a^2} S_{ij}), \quad (18)$$

where

$$T_{ij} = \sum_{\vec{\mu}} \langle \phi_i(\vec{r}) | -\frac{1}{2}\nabla^2 | \phi_j(\vec{r} - a\vec{\mu}) \rangle \quad (19)$$

The last term of Eq. (17) gives no contribution because its summation is antisymmetric in \vec{r} .

The two-electron matrix element of h is reduced with the aid of the Fourier convolution and inversion theorems. Proceeding tentatively term by term, the Fourier representation needed for the r_{12}^{-1} integral is⁴

$$\begin{aligned} &\langle \phi_i(\vec{r}_1 - a\vec{\lambda}) \phi_m(\vec{r}_2 - a\vec{\lambda}') | r_{12}^{-1} | \phi_j(\vec{r}_1 - a\vec{\mu}) \phi_n(\vec{r}_2 - a\vec{\mu}') \rangle \\ &= \frac{1}{\pi a} \int \frac{d\vec{q}}{q^2} e^{-2\pi i \vec{q} \cdot (\vec{\lambda}' - \vec{\lambda})} [\phi_i^*(\vec{r}_1) \phi_j(\vec{r}_1 - a\vec{\mu} + a\vec{\lambda})]^T(\vec{q}) \cdot \\ &\quad \cdot [\phi_m^*(\vec{r}_2) \phi_n(\vec{r}_2 - a\vec{\mu}' + a\vec{\lambda}')]^T(-\vec{q}) \end{aligned} \quad (20)$$

where $[\cdot]^T$ indicates Fourier transform (in reciprocal-lattice units). A main virtue of Eq. (20) is that the dependence on the spatial locations of the orbitals is partially transferred to the complex exponential. In $\langle \vec{k}_i | \vec{k}'_m | r_{12}^{-1} | \vec{k}_j | \vec{k}'_n \rangle$, the left side of Eq. (20) occurs summed over $\vec{\lambda}$, $\vec{\lambda}'$, $\vec{\mu}$ and $\vec{\mu}'$. Independently of $\vec{\lambda}$, we note that the $\vec{\mu}$ summation produces the factor $\phi_{ij}(\vec{q})$ introduced in Eq. (14), while the $\vec{\mu}'$ summation yields $\phi_{mn}(-\vec{q})$. We thus have

$$\langle \vec{k}_i \vec{k}'_m | r_{12}^{-1} | \vec{k}_j \vec{k}'_n \rangle = \frac{1}{\pi a} \int \frac{d\vec{q}}{q^2} \sum_{\vec{\lambda} \vec{\lambda}'} e^{-2\pi i \vec{q} \cdot (\vec{\lambda} - \vec{\lambda}')} \phi_{ij}(\vec{q}) \phi_{mn}(-\vec{q}) \quad (21)$$

Now that we see that the right side of Eq. (21) depends on $\vec{\lambda}$ and $\vec{\lambda}'$ only through the complex exponential, we reduce it further by invoking the lattice orthogonality relation

$$\sum_{\vec{\lambda} \vec{\lambda}'} e^{-2\pi i \vec{q} \cdot (\vec{\lambda} - \vec{\lambda}')} = N \sum_{\vec{v}} \delta(\vec{q} - \vec{v}) \quad (22)$$

The result is

$$\langle \vec{k}_i \vec{k}'_m | r_{12}^{-1} | \vec{k}_j \vec{k}'_n \rangle = \frac{N}{\pi a} \sum_{\vec{v}} v^{-2} \phi_{ij}(\vec{v}) \phi_{mn}(-\vec{v}) \quad (23)$$

We defer temporarily the question of the singularity at $\vec{v} = 0$.

Next consider the electron-nuclear interaction terms of h. The result analogous to Eq. (21) is

$$\langle \vec{k}_i \vec{k}'_m | \frac{1}{N} \sum_{\vec{\mu}} |\vec{r}_1 - a\vec{\mu}|^{-1} | \vec{k}_j \vec{k}'_n \rangle = \frac{1}{\pi a} \int \frac{d\vec{q}}{q^2} \sum_{\vec{\lambda} \vec{\mu}} e^{-2\pi i \vec{q} \cdot (\vec{\lambda} - \vec{\mu})} \phi_{ij}(\vec{q}) \phi_{mn}(0) \quad (24)$$

where $\langle \vec{k}'_m | \vec{k}'_n \rangle$ has been replaced by $N \phi_{mn}(0)$. Again applying Eq. (22), we reach

$$\langle \vec{k}_i \vec{k}'_m | \frac{1}{N} \sum_{\vec{\mu}} |\vec{r}_1 - a\vec{\mu}|^{-1} | \vec{k}_j \vec{k}'_n \rangle = \frac{N}{\pi a} \sum_{\vec{v}} v^{-2} \phi_{ij}(\vec{v}) \phi_{mn}(0) \quad (25)$$

Similarly,

$$\langle \vec{k}_i \vec{k}'_m | \frac{1}{N} \sum_{\vec{\mu}'} |\vec{r}_2 - a\vec{\mu}'|^{-1} | \vec{k}_j \vec{k}'_n \rangle = \frac{N}{\pi a} \sum_{\vec{v}} v^{-2} \phi_{ij}(0) \phi_{mn}(-\vec{v}) \quad (26)$$

We have replaced \vec{v} by $-\vec{v}$ in Eq. (26) to make the symmetry more explicit. Again we defer discussion of $\vec{v} = 0$.

To handle the last term of h, we introduce the Fourier representation of $|\vec{R}_{\mu} - \vec{R}_{\mu'}|^{-1}$, obtaining

$$\begin{aligned}
 & \langle \vec{k}_i \vec{k}'_m | \frac{1}{N^2} \sum_{\substack{\mu\mu' \\ (\mu \neq \mu')}} |\vec{R}_\mu - \vec{R}_{\mu'}|^{-1} | \vec{k}_j \vec{k}'_n \rangle \\
 &= \frac{1}{\pi a} \int \frac{d\vec{q}}{q^2} \sum_{\substack{\mu\mu' \\ (\mu \neq \mu')}} e^{-2\pi i \vec{q} \cdot (\vec{\mu}' - \vec{\mu})} \phi_{ij}(0) \phi_{mn}(0). \tag{27}
 \end{aligned}$$

To apply the lattice orthogonality relation to Eq. (27), we add and subtract on its right hand side the terms with $\mu = \mu'$, after which we have

$$\begin{aligned}
 & \langle \vec{k}_i \vec{k}'_m | \frac{1}{N^2} \sum_{\substack{\mu\mu' \\ (\mu \neq \mu')}} |\vec{R}_\mu - \vec{R}_{\mu'}|^{-1} | \vec{k}_j \vec{k}'_n \rangle \\
 &= \frac{N}{\pi a} \left(\sum_{\vec{v}} v^{-2} - \int \frac{d\vec{q}}{q^2} \right) \phi_{ij}(0) \phi_{mn}(0) \tag{28}
 \end{aligned}$$

It is now time to consider the behavior of the $\vec{v} = 0$ terms in Eqs. (23), (25), (26), and (28). A superficial check indicates that in the combination needed for h , these terms would cancel. However, the integration involving the δ function from Eq. (22) is at best questionable when the remainder of the integrand is singular, and there is in addition a problem in the interchange of summation and integration which was carried out without comment to reach Eqs. (21) and (24). Near $\vec{q} = 0$ this interchange is improper due to a lack of necessary uniform convergence. In spite of these difficulties, a more careful analysis verifies that the $\vec{v} = 0$ terms do indeed cancel.

After dropping the $\vec{v} = 0$ terms, the right side of Eq. (28) still contains the difference of a divergent sum and a divergent integral. The difference is nevertheless convergent if (as is appropriate here) the sum and integral are extended over a carefully defined region which is then allowed to become infinite. The result, which depends upon the lattice symmetry, is $-8.913633\dots$ for the simple cubic unit lattice⁵.

Combining the results of the foregoing analysis, we have

$$\begin{aligned} \langle \vec{k}_i \vec{k}'_m | h | \vec{k}_j \vec{k}'_n \rangle = \frac{N}{\pi a} \sum_{\vec{v}}' v^{-2} & \left(\phi_{ij}(\vec{v}) \phi_{mn}(-\vec{v}) - \phi_{ij}(\vec{v}) \phi_{mn}(0) \right. \\ & \left. - \phi_{ij}(0) \phi_{mn}(-\vec{v}) \right) - \frac{8.913633N}{\pi a} \phi_{ij}(0) \phi_{mn}(0). \end{aligned} \quad (29)$$

An analysis of the exchange matrix element also proceeds by the argument leading toward Eq. (23), but the complex exponentials in \vec{k} and \vec{k}' do not cancel. In place of Eq. (20), we have

$$\begin{aligned} & \langle \phi_i(\vec{r}_1 - a\vec{\lambda}) \phi_m(\vec{r}_2 - a\vec{\lambda}') | r_{12}^{-1} e^{2\pi i a^{-1}(\vec{k}-\vec{k}') \cdot (\vec{r}_2 - \vec{r}_1)} | \phi_n(\vec{r}_1 - a\vec{\mu}) \phi_j(\vec{r}_2 - a\vec{\mu}') \rangle \\ &= \frac{1}{\pi a} \int \frac{d\vec{q}}{q^2} e^{-2\pi i (\vec{q} - \vec{k} + \vec{k}') \cdot (\vec{\lambda}' - \vec{\lambda})} [\phi_i^*(\vec{r}_1) \phi_n(\vec{r}_1 - a\vec{\mu} + a\vec{\lambda})]^T (\vec{q} - \vec{k} + \vec{k}') \cdot \\ & \quad \cdot [\phi_m^*(\vec{r}_2) \phi_j(\vec{r}_2 - a\vec{\mu}' + a\vec{\lambda}')]^T (-\vec{q} + \vec{k} - \vec{k}'). \end{aligned} \quad (30)$$

Applying the summations and the lattice orthogonality relation, we reach

$$\langle \vec{k}_i \vec{k}'_m | r_{12}^{-1} | \vec{k}'_n \vec{k}_j \rangle = \frac{N}{\pi a} \sum_{\vec{v}}' \frac{1}{|\vec{v} + \vec{k} - \vec{k}'|^2} \phi_{in}(\vec{v}) \phi_{mj}(-\vec{v}). \quad (31)$$

Notice that the point $\vec{v} = 0$ remains in the exchange sum.

Equation (29), for the Coulomb matrix elements, may be given a simple physical interpretation if we note that its last term describes the entire Coulomb energy in the limit that the electronic distribution is uniformly distributed (in which case $\phi_{ij}(\vec{v})$ and $\phi_{mn}(\vec{v})$ vanish for all nonzero \vec{v}). The other terms of Eq. (29) therefore describe energies associated with the non-uniformity of the electronic charge, the first term giving the modification of the electron repulsion energy while the second and third terms give the nonuniformity contribution to the electron-nuclear attraction energy. In crystals possessing a relatively diffuse charge distribution, we would expect the electron-nuclear terms to be small relative to the final term (or to the kinetic energy), and the electron-electron terms would be smaller yet. This observation indicates that the bulk of the electronic energy will often be explainable in terms of one-electron models, thereby providing a rationale for the success of pseudopotential and related methods.

Equation (31) provides some understanding of the nature of the exchange energy. To start, the $\vec{v} = 0$ term describes a free-electron-like contribution to the exchange energy arising from

overlap charges (remember that $\phi_{in}(0)$ and $\phi_{mj}(0)$ are proportional to overlap matrix elements). The remaining terms of Eq. (31) approach electron-repulsion terms of the same type as enter the Coulomb matrix elements, because $|\vec{v} + \vec{k} - \vec{k}'|^{-2}$ rapidly approaches v^{-2} as v increases. The $\vec{v} \neq 0$ terms of Eq. (31) may therefore be characterized as a Coulomb-like contribution plus a small remainder. This point of view may be emphasized by writing

$$\begin{aligned} & \langle \vec{k}_i \vec{k}'_m | r_{12}^{-1} | \vec{k}'_n \vec{k}_j \rangle = \frac{N}{\pi a} \frac{1}{|\vec{k} - \vec{k}'|^2} \phi_{in}(0) \phi_{mj}(0) \\ & + \frac{N}{\pi a} \sum_{\vec{v}}' v^{-2} \phi_{in}(\vec{v}) \phi_{mj}(-\vec{v}) + \frac{N}{\pi a} \sum_{\vec{v}}' \left(\frac{1}{|\vec{v} + \vec{k} - \vec{k}'|^2} - \frac{1}{v^2} \right) \phi_{in}(\vec{v}) \phi_{mj}(-\vec{v}) \end{aligned} \quad (32)$$

Equation (32) shows that the total exchange energy consists of a term of a free-electron-like magnitude plus a contribution comparable in size to the nonuniformity contribution to the Coulomb energy. This fact will be useful in determining the precision needed in computations of exchange matrix elements.

WORKING FORMULAS

We now collect the results of the preceding section into working expressions for the Fock matrix and the energy. In doing so it is convenient to rewrite the lattice sums to consolidate terms which are identical for symmetry reasons. Accordingly we introduce the notion of the star of a vector, consisting of the set of all vectors made equivalent to the original vector by the operations of the point group of the lattice⁶. There will be 48 members in the star of a vector having an asymmetric relation to the symmetry directions of the lattice, but more symmetrically oriented vectors will have stars with fewer members.

We proceed by grouping the vectors describing reciprocal-lattice points into stars; the 26 such stars of shortest length are listed in Table I, which also gives each star's number of members (denoted by g). Where possible we then replace lattice sums denoted by a vector index, say \vec{v} , by sums over stars v ($v = 1, 2, \dots$) with weights g_v . A representative vector from Star v is denoted \vec{q}_v .

Making use of the fact that $\phi_{ij}(\vec{q})$ is identical for all vectors \vec{q} of the same star, we write the Fock matrix elements in the form

Table I. Vectors defining stars of lattice points in reciprocal space. The components q_x , q_y , q_z and the square length q^2 are in reciprocal-lattice units; g is the number of equivalent vectors in the star. The symbols assigned to certain stars indicate directions in the notation of Bouckaert, Smoluchowski, and Wigner⁷.

Symbol	Star No.	q_x	q_y	q_z	q^2	g
Γ	1	0	0	0	0	1
Δ	2	1	0	0	1	6
Σ	3	1	1	0	2	12
Λ	4	1	1	1	3	8
Δ	5	2	0	0	4	6
	6	2	1	0	5	24
	7	2	1	1	6	24
Σ	8	2	2	0	8	12
	9	2	2	1	9	24
Δ	10	3	0	0	9	6
	11	3	1	0	10	24
	12	3	1	1	11	24
Λ	13	2	2	2	12	8
	14	3	2	1	13	24
	15	3	2	1	14	48
Δ	16	4	0	0	16	6
	17	3	2	2	17	24
	18	4	1	0	17	24
Σ	19	3	3	0	18	12
	20	4	1	1	18	24
	21	3	3	1	19	24
	22	4	2	0	20	24
	23	4	2	1	21	48
	24	3	3	2	22	24
	25	4	2	2	24	24
	26	4	3	0	25	24

$$F_{ij}(\vec{k}) = T_{ij} + \frac{2\pi^2 k^2}{a^2} S_{ij} + V_{ij} + \frac{1}{\pi a} \sum_{v \neq 1} g_v q_v^{-2} \sum_{mn} P_{mn} \phi_{ij}(\vec{q}_v) \phi_{mn}(-\vec{q}_v) - \frac{1}{2\pi a} \sum_v g_v \sum_{mn} Q_{mn}^v(\vec{k}) \phi_{in}(\vec{q}_v) \phi_{mj}(-\vec{q}_v), \quad (33)$$

where ϕ_{ij} , S_{ij} and T_{ij} are given in Eqs. (14), (16), and (19), and

$$V_{ij} = -\frac{1}{\pi a} \sum_{v \neq 1} g_v q_v^{-2} \phi_{ij}(\vec{q}_v) \quad (34)$$

$$P_{mn} = 2 \int c_m^*(\vec{k}') c_n(\vec{k}') d\vec{k}' \quad (35)$$

$$Q_{mn}^v(\vec{k}) = \left\langle 2 \int \frac{c_m^*(\vec{k}') c_n(\vec{k}') d\vec{k}'}{|\vec{q}_v + \vec{k} - \vec{k}'|^2} \right\rangle_v \quad (36)$$

The notation $\langle \rangle_v$ in Eq. (36) denotes the average over the members of Star v . We have used the scaling condition of Eq. (7), in the form

$$\sum_{mn} c_m^*(\vec{k}') c_n(\vec{k}') \phi_{mn}(0) = 1 \quad (37)$$

and the fact that the \vec{k}' integration region has volume $\frac{1}{2}$, to eliminate $\phi_{mn}(0)$ and obtain the V_{ij} term of Eq. (33). We have also dropped from $F_{ij}(\vec{k})$ the terms proportional to S_{ij} and independent of \vec{k} , as these can be accounted for by an additive change in the multiplier $\epsilon(\vec{k})$.

An energy expression consistent with Eq. (33) is

$$E = N \left[\sum_{ij} d\vec{k} c_i^*(\vec{k}) c_j(\vec{k}) (F_{ij}(\vec{k}) + T_{ij} + \frac{2\pi^2 k^2}{a^2} S_{ij} + V_{ij}) - \frac{8.913633}{2\pi a} \right] \quad (38)$$

If the Hartree-Fock equations are satisfied, Eq. (38) can also be written

$$E = N \left[\int \epsilon(\vec{k}) d\vec{k} + \sum_{ij} \int d\vec{k} c_i^*(\vec{k}) c_j(\vec{k}) (T_{ij} + \frac{2\pi^2 k^2}{a^2} S_{ij} + V_{ij}) - \frac{8.913633}{2\pi a} \right] \quad (39)$$

COMPUTATIONAL METHODS

The first main computational problem encountered in evaluating the quantities of the preceding section is for $\phi_{ij}(\vec{q})$. Fourier transforms for the individual terms in the summation of Eq. (14) are cumbersome but can be used. We hope to report on such an approach elsewhere. However, the lattice sum permits rearrangement based on lattice orthogonality relations. Using the Fourier convolution theorem, we write

$$\begin{aligned}\phi_{ij}(\vec{q}) &= \sum_{\vec{\mu}} \left[\phi_i(\vec{r}) \phi_j(\vec{r} - a\vec{\mu}) \right]^T(\vec{q}) \\ &= \sum_{\vec{\mu}} a^{-3} \int d\vec{p} \phi_i^*(\vec{p}) \phi_j^T(\vec{q} - \vec{p}) e^{2\pi i (\vec{q} - \vec{p}) \cdot \vec{\mu}}\end{aligned}\quad (40)$$

where the complex exponential allows for the fact that the argument of ϕ_j is $\vec{r} - a\vec{\mu}$. Invoking Eq. (22), we find

$$\phi_{ij}(\vec{q}) = a^{-3} \sum_{\vec{v}} \phi_i^*(\vec{v}) \phi_j^T(\vec{q} - \vec{v}) \quad (41)$$

If we use 1s Slater-type orbitals with screening parameters ζ_i and ζ_j ,

$$\phi_{ij}(\vec{q}) = \frac{8}{\pi^2} \sum_{\vec{v}} \frac{(\delta_i \delta_j)^{5/2}}{(q_v^2 + \delta_i^2)^2 (|q - q_v|^2 + \delta_j^2)^2} \quad (42)$$

where $\zeta_i = a\zeta_i/2\pi$.

Proceeding to the other one-electron quantities, the kinetic-energy matrix element T_{ij} may be handled analogously, leading to

$$T_{ij} = \frac{16}{a^2} \sum_{\vec{v}} g_v q_v^2 \frac{(\delta_i \delta_j)^{5/2}}{(q_v^2 + \delta_i^2)^2 (q_v^2 + \delta_j^2)^2} \quad (43)$$

In deriving Eq. (43), we note that in the reciprocal-lattice units we use for \vec{q} , $[-\frac{1}{2}\nabla^2 \phi]^T = (2\pi^2 q^2/a^2) \phi^T(\vec{q})$. There is no difficulty in carrying out the summation of Eq. (34) to find V_{ij} .

Before continuing to a discussion of the quantities P_{mn} and $Q_{mn}^V(\vec{k})$ which are needed for the two-electron contributions to

Table II. Fourier transforms of lattice sums of 1s orbital products (as defined in Eq. (14) of the text), evaluated at the reciprocal-lattice points listed in Table I. Orbital screening parameters: $\zeta_1 = 2$, $\zeta_2 = 1$. Data are for lattice spacing $a = 2$.

Star No.	ϕ_{11}	ϕ_{12}	ϕ_{22}
1	3.3136	8.9326	25.1456
2	0.6063	0.8350	0.4306
3	0.2646	0.2954	0.1228
4	0.1525	0.1511	0.0575
5	0.1076	0.0943	0.0342
6	0.0742	0.0627	0.0220
7	0.0549	0.0449	0.0154
8	0.0343	0.0265	0.0088
9	0.0279	0.0212	0.0070
10	0.0283	0.0213	0.0071
11	0.0234	0.0174	0.0057
12	0.0198	0.0145	0.0047
13	0.0169	0.0123	0.0040
14	0.0147	0.0106	0.0034
15	0.0128	0.0092	0.0030
16	0.0101	0.0071	0.0023
17	0.0090	0.0063	0.0020
18	0.0090	0.0063	0.0020
19	0.0081	0.0057	0.0018
20	0.0081	0.0057	0.0018
21	0.0073	0.0051	0.0016
22	0.0067	0.0046	0.0015
23	0.0061	0.0042	0.0013
24	0.0056	0.0038	0.0012
25	0.0047	0.0032	0.0010
26	0.0044	0.0030	0.0009

Let us examine some numerical values of $\phi_{ij}(\vec{q})$, T_{ij} , and V_{ij} . In Table II we present some $\phi_{ij}(\vec{q})$ values for the first 26 inequivalent reciprocal-lattice points. These data are significant in that they indicate the rate at which the ϕ_{ij} approach zero as q increases. The calculations are reported for two sets of orbitals, the more localized of which has $\zeta a = 4$, meaning its relative amplitude has dropped to e^{-4} at the nearest neighboring lattice points. This set of orbitals is more localized than would be expected in an insulator of appreciable cohesive energy. Even for this set of orbitals, which exhibits the poorer convergence, we observe a highly gratifying decrease within increasing q . This behavior appears in part to be attributable to the effect of the lattice sum, i.e. to the combination of orbital products in all orientations. As we shall see more fully later, the convergence in the $\phi_{ij}(\vec{q})$ helps ensure the convergence of the reciprocal-space

Table III. One-electron contributions to kinetic energy T and potential energy V , in hartrees, for normalized 1s-orbital Bloch functions of various k (in reciprocal-lattice units). These quantities are independent of the direction of the Bloch wave vector. Subscripts indicate orbitals; orbital screening parameters are $\zeta_1 = 2$, $\zeta_2 = 1$. Data are for lattice spacing $a = 2$.

k	T_{11}	V_{11}	T_{12}	V_{12}	T_{22}	V_{22}
0	0.367	-0.330	0.034	-0.142	0.003	-0.024
0.1	0.417	-0.330	0.083	-0.142	0.053	-0.024
0.2	0.564	-0.330	0.227	-0.142	0.200	-0.024
0.3	0.802	-0.330	0.460	-0.142	0.438	-0.024
0.4	1.157	-0.330	0.807	-0.142	0.793	-0.024
0.5	1.601	-0.330	1.242	-0.142	1.237	-0.024

summations of the two-electron energy contributions.

In Table III are listed some one-electron kinetic and potential energy contributions. We see that the kinetic energy dominates, particularly for the larger k values. This is to be expected, being due to the fact that the potential energy is only that associated with the distortion from a uniform electron distribution.

Consider now the calculation of P_{mn} and $Q_{mn}(\vec{k})$, both of which involve reciprocal-space integrations to a Fermi surface given by the expansion, Eq. (13), of integrands including coefficients expanded according to Eq. (11). For P_{mn} the integration is simply of the product $c_m^* c_n$. The expansions of c_m^* and c_n can be multiplied together, using the cubic-harmonic equivalent of the Clebsch-Gordan expansion to reduce products of cubic harmonics. The result is

$$c_m^*(\vec{k}) c_n(\vec{k}) = \sum_{pp'} \sum_{\mu\mu'\lambda} c_m^{m*} c_n^{n*} D_{\mu\mu'\lambda} k^{p+p'} T_\lambda(\Omega_k) \quad (44)$$

$D_{\mu\mu'\lambda}$ is a cubic Clebsch-Gordan coefficient (see Appendix I) defined by

$$T_\mu(\Omega) T_{\mu'}(\Omega) = \sum_\lambda D_{\mu\mu'\lambda} T_\lambda(\Omega) \quad (45)$$

We next integrate Eq. (44) to the Fermi surface, performing the radial integration first. At that stage, writing Ω for Ω_k ,

$$\int d\Omega \int_0^{f(\Omega)} dk k^{p+p'+2} T_\lambda(\Omega) = \frac{1}{p+p'+3} \int d\Omega T_\lambda(\Omega) [f(\Omega)]^{p+p'+3} \quad (46)$$

Introduction of Eq. (13), repeated use of Eq. (45), and recognition that $T_1 = (4\pi)^{-\frac{1}{2}}$ is the only cubic harmonic whose angular integral does not vanish, all lead to

$$\int d\vec{k} k^{p+p'} T_\lambda(\Omega) = \frac{(4\pi)^{\frac{1}{2}}}{p+p'+3} (\underline{E}^{p+p'+3})_{\lambda 1}, \quad (47)$$

where \underline{E} is a matrix of elements

$$E_{\lambda\sigma} = \sum_{\mu} f_{\mu} D_{\mu\lambda\sigma} \quad (48)$$

Thus,

$$P_{mn} = \sum_{pp'} \sum_{\mu\mu'\lambda} \frac{(4\pi)^{\frac{1}{2}}}{p+p'+3} c_{p\mu}^{m*} c_{p'\mu'}^n D_{\mu\mu'\lambda} (\underline{E}^{p+p'+3})_{\lambda 1} \quad (49)$$

Evaluation of $Q_{mn}^v(\vec{k})$ uses the ideas illustrated above, but has the additional complication of the factor $|\vec{q}_v + \vec{k} - \vec{k}'|^{-2}$ and the average over the orientations of \vec{q}_v . For $q > k + k'$, an effective procedure is to expand $|\vec{q}_v + \vec{k} - \vec{k}'|^{-2}$ into cubic harmonics, keeping only the totally symmetric terms (the others would not survive the \vec{k}' integration and the orientation averaging of \vec{q}_v). The expansion is

$$\frac{1}{|\vec{q}_v + \vec{k} - \vec{k}'|^2} = \frac{16\pi^2}{q_v^2} \sum_{\rho\sigma\sigma'} \sum_{s=0}^{\infty} \sum_{s'=0}^{\infty} G_{\rho\sigma\sigma'}^{ss'} \left(\frac{\vec{k}}{q_v}\right)^{\ell+2s} \left(\frac{\vec{k}'}{q_v}\right)^{\ell'+2s'}.$$

• $T_\rho(\Omega_v) T_\sigma(\Omega_k) T_{\sigma'}(\Omega_{k'}) + \text{less symmetric terms,} \quad (50)$

where

$$G_{\rho\sigma\sigma'}^{ss'} = (-1)^\ell D_{\rho\sigma\sigma'} \frac{(2s+2s'+\ell+\ell'-L-1)!! (2s+2s'+\ell+\ell'+L)!!}{(2s+2\ell+1)!! (2s)!! (2s'+2\ell'+1)!! (2s')!!} \quad (51)$$

and ℓ , ℓ' , and L are the angular momentum quantum numbers of T_σ , $T_{\sigma'}$, and T_ρ , respectively.

For the present problem, Eq. (50) is certainly valid for all \vec{q}_v except $\vec{q}_v = 0$ and possibly the points equivalent to $\vec{q}_v = (1, 0, 0)$, i.e. stars 1 and 2. We have chosen to use Eq. (50) for all \vec{q}_v except those just mentioned. Because the terms retained on the right side of Eq. (50) are symmetric with respect to the possible orientations of \vec{q}_v , the averaging indicated in the definition of $Q_{mn}^v(\vec{k})$ may be dropped and any member of Star v may be used.

Straightforward analysis gives the following expression:

$$Q_{mn}^v(\vec{k}) = \frac{16\pi^2}{q_v^2} \sum_{pp'} \sum_{\mu\mu'\lambda} \sum_{ss'} \sum_{\rho\sigma\sigma'} \frac{c_m^{m*} c_n^{n*} D_{\mu\mu'\lambda} G_{\rho\sigma\sigma'}^{ss'}}{(p+p'+\ell'+2s'+3)q_v^{\ell+\ell'+2s+2s'}} \cdot T_p(\Omega_v) k^{\ell+2s} T_{\sigma}(\Omega_k) (E^{p+p'+\ell'+2s'+3})_{\lambda\sigma'}, \quad v>2 \quad (52)$$

The expansion indicated in Eq. (52) seems formidable, as it contains five infinite summations in addition to those needed for a good representation of $c_m^* c_n$. However, these summations converge rapidly, and for larger v we are helped by the convergence of the quantities multiplying $Q_{mn}^v(\vec{k})$ in the equation for the Fock matrix. We have determined which terms need be retained in Eq. (52) to give the exchange energy to four significant figures. For orbital screening parameters $\zeta \leq 2$, we find that for all stars $v > 2$ we need a total of 60 points, as listed in Table IV. That table shows that for Star 8 and beyond, only the leading term (which is simply q_v^{-2}) is needed, and that no significant contribution at all lies beyond Star 26. We see that only through Star 5 is there a significant angular dependence in $|\vec{q}_v + \vec{k} - \vec{k}'|^{-2}$. The convergence becomes poorer as the maximum orbital screening increases, but as we previously pointed out, $\zeta = 2$ is already quite large for the present problem.

To gain greater confidence in the computational method just described, we calculated $Q_{mn}^v(\vec{k})$ for a number of \vec{k} values. Samples of these results are given in Table V, and the contributions of various v values to the Fock matrix are presented in Table VI. In making the calculations for Tables V and VI we assumed spherically symmetric $c_s(\vec{k})$, so that the departures from spherical symmetry in the Fock matrix are entirely due to the cubic crystal structure.

There remains the problem of the evaluation of $Q_{mn}^v(\vec{k})$ for stars $v = 1$ and $v = 2$. The best method we have found for these cases is entirely general, but is somewhat slower than that already described and is therefore only used when necessary. We simply expand $|\vec{q}_v + \vec{k} - \vec{k}'|^{-2}$ as

$$|\vec{q}_v + \vec{k} - \vec{k}'|^{-2} = \frac{2\pi}{k' |\vec{q}_v + \vec{k}|} \sum_p T_p(\Omega') T_p(\Omega_k) Q_L(w), \quad (53)$$

where Q_L is a Legendre function of the second kind, Ω' is the direction of $\vec{q}_v + \vec{k}$, and $w = (|\vec{q}_v + \vec{k}|^2 + k'^2)/2k' |\vec{q}_v + \vec{k}|$, and L

Table IV. Terms needed to obtain relative accuracy of 10^{-4} in expansion of exchange energy, Eq. (52), for 1s orbitals with $\zeta \leq 2$. Data are for lattice spacing $a = 2$.

v	Star No.	σ	σ'	s	s'	v	Star No.	σ	σ'	s	s'
3	1	1	0	0		5	1	1	0	0	
	1	1	1	0			1	1	1	0	
	1	1	0	1			1	1	0	1	
	1	1	2	0			2	1	0	0	
	1	1	1	1			1	2	0	0	
	1	1	0	2			1	1	0	0	
	1	1	3	0			1	1	1	0	
	1	1	2	1			1	1	0	1	
	1	1	1	2			1	1	0	0	
	1	1	0	3			1	1	1	0	
	1	1	3	1			1	1	0	1	
	1	1	2	2			1	1	0	0	
	1	1	1	3			1	1	0	0	
	2	1	0	0			10	1	1	0	
	2	1	0	1			11	1	1	0	
	1	2	0	0			12	1	1	0	
	1	2	1	0			13	1	1	0	
	3	1	0	0			14	1	1	0	
	3	1	0	1			15	1	1	0	
	2	2	0	0			16	1	1	0	
	1	3	0	0			17	1	1	0	
	1	3	1	0			18	1	1	0	
4	1	1	0	0			19	1	1	0	
	1	1	1	0			20	1	1	0	
	1	1	0	1			21	1	1	0	
	1	1	2	0			22	1	1	0	
	1	1	1	1			23	1	1	0	
	1	1	0	2			24	1	1	0	
	2	1	0	0			25	1	1	0	
	1	2	0	0			26	1	1	0	

Table V. Dependence of $Q_{11}^v(\vec{k})$, Eq. (36), on direction and magnitude of Bloch vector \vec{k} , for v at reciprocal-lattice points listed in Table I. Bloch vector \vec{k} is specified (in reciprocal-lattice units) by its magnitude k and the direction of a star from Table I. Calculations are for spherical Fermi surface and only orbital 1 occupied. The Q_{11} values are compared with their large- v asymptote P_{11}/v^2 , where P_{11} is given by Eq. (35). Subscript indicates Orbital 1, a 1s orbital with screening parameter $\zeta_1 = 2$. Data are for lattice spacing $a = 2$.

v Star no.	3	4	5	6	7
P_{11}/v^2	0.1509	0.1006	0.0754	0.0604	0.0503
$Q_{11}^v(k = 0.2236)$					
\vec{k} stars no. 2,6,11,12	0.1562	0.1029	0.0767	0.0611	0.0508
no. 3,4,7,9,14,15	0.1563	0.1029	0.0767	0.0611	0.0508
$Q_{11}^v(k = 0.3873)$					
\vec{k} star no. 2	0.1588	0.1038	0.0775	0.0615	0.0511
3	0.1597	0.1043	0.0772	0.0615	0.0511
4	0.1596	0.1044	0.0772	0.0615	0.0511
6	0.1594	0.1041	0.0774	0.0615	0.0511
7	0.1595	0.1043	0.0773	0.0615	0.0511
9	0.1596	0.1043	0.0772	0.0615	0.0511
11	0.1591	0.1040	0.0774	0.0615	0.0511
12	0.1593	0.1041	0.0774	0.0615	0.0511
14	0.1595	0.1042	0.0773	0.0615	0.0511
15	0.1596	0.1043	0.0772	0.0615	0.0511
$Q_{11}^v(k = 0.5)$					
\vec{k} star no. 2	0.0611	0.1045	0.0786	0.0619	0.0514
3	0.1637	0.1057	0.0778	0.0619	0.0514
4	0.1632	0.1061	0.0775	0.0619	0.0514
6	0.1628	0.1053	0.0781	0.0619	0.0514
7	0.1630	0.1057	0.0778	0.0619	0.0514
9	0.1634	0.1060	0.0776	0.0619	0.0514
11	0.1621	0.1049	0.0783	0.0619	0.0514
12	0.1625	0.1053	0.0781	0.0619	0.0514
14	0.1633	0.1055	0.0779	0.0619	0.0514
15	0.1632	0.1057	0.0778	0.0619	0.0514

Table VI. Contributions (in hartrees) of stars of reciprocal-lattice points \vec{v} listed in Table I to Fock matrix elements $F_{11}(\vec{k})$, Eq. (33), for Bloch vectors \vec{k} specified (in reciprocal-lattice units) by magnitude k and the direction of a star from Table I. Calculations are for spherical Fermi surface and only Orbital 1 occupied. Subscript indicates Orbital 1, a 1s orbital with screening parameter $\zeta_1 = 2$. Data are for lattice spacing $a = 2$.

\vec{v} star no.	$\vec{k} = 0.2236$	$\vec{k} = 0.3873$	$\vec{k} = 0.5000$	$\vec{k} = 0.5000$
	\vec{k} star no.2	\vec{k} star no.2	\vec{k} star no.2	\vec{k} star no.3
3	0.009733	0.009555	0.009403	0.009231
4	0.001455	0.001441	0.001431	0.001413
5	0.000410	0.000405	0.000400	0.000404
6	0.000627	0.000623	0.000619	0.000619
7	0.000287	0.000285	0.000283	0.000283
8	0.000042	0.000042	0.000042	0.000042
9	0.000050	0.000050	0.000050	0.000050
10	0.000013	0.000013	0.000013	0.000013
20	0.000002	0.000002	0.000002	0.000002
26	0.000000	0.000000	0.000000	0.000000

is the angular-momentum quantum number of T . Use of Eq. (53) leads to

$$Q_{mn}^v(\vec{k}) = \left\langle \frac{2\pi}{|\vec{q}_v + \vec{k}|} \sum_{\rho\sigma} T_\rho(\Omega') \sum_{pp'} \sum_{\mu\mu'\lambda} c_{p\mu}^{m*} c_{p'\mu'}^n D_{\mu\mu'} D_{\lambda\rho\sigma} \int T_\sigma(\Omega_{k'}) k'^{p+p'-1} Q_L(w) dk' \right\rangle_v \quad (54)$$

The radial part of the \vec{k}' integration is done analytically, followed by a numerical angular integration. The orientational average for Star $v = 2$ is obtained as a direct average of the six contributing orientations.

The radial integrations can be reduced to the general form

$$I_{pL}(\bar{u}) = \int_0^{\bar{u}} u^{p+1} Q_L\left(\frac{1+u^2}{2u}\right) du \quad (55)$$

where $\bar{u} = f(\Omega_{k'})/|\vec{q}_v + \vec{k}|$. The I_{pL} may be generated recursively, starting from

$$I_{00}(\bar{u}) = \frac{2}{3} [Q_2(\bar{u}) - Q_0(\bar{u})] + 2\bar{u} \quad (56)$$

and

$$I_{-1,1}(\bar{u}) = \frac{1}{3} [Q_2(\bar{u}) - Q_0(\bar{u})] + \Lambda(\bar{u}), \quad (57)$$

where

$$\Lambda(\bar{u}) = \sum_{j=0}^{\infty} \frac{\bar{u}^{2j+1}}{(2j+1)^2}, \quad \bar{u} \leq 1 \quad (58a)$$

$$= \frac{\pi^2}{4} - \sum_{j=0}^{\infty} \frac{\bar{u}^{-2j-1}}{(2j+1)^2} \quad \bar{u} \geq 1. \quad (58b)$$

At $\bar{u} = 1$, the right sides of Eqs. (56) and (57) approach continuously their limiting values: $I_{00}(1) = 1$, $I_{-1,1}(1) = (\pi^2 - 4)/8$.

Further details of the evaluation of $Q_{mn}^v(\vec{k})$ for $v = 1$ and 2 will be presented elsewhere, as the work is still in progress.

ACKNOWLEDGEMENTS

This work has been supported in part by National Science Foundation Grant GP-11170. The authors also gratefully acknowledge stimulating discussions with many colleagues, particularly Drs. B. Gale Dick and Ryo Togei. Finally, they thank Miss Karen Bryan for skillfully and cheerfully converting crude scribbling into finished copy.

APPENDIX I. CUBIC HARMONICS

The set of eigenfunctions of ℓ_{op}^2 of a given quantum number ℓ , $\ell = 0, 1, 2, \dots$, where ℓ_{op} is the angular momentum operator, may be thought of as the basis functions for various irreducible representations of the three-dimensional rotation group. As the cubic point group is a subgroup of the full rotation group, the eigenfunctions of given ℓ are also bases for cubic-group representations which may, however, be reducible. This implies that the eigenfunctions of given ℓ can be transformed among themselves so as to have definite symmetry properties under the cubic-group operations. The resulting functions are called cubic harmonics. Our interest here is in the ℓ_{op}^2 eigenfunctions which are invariant under all operations of the cubic group, i.e. the totally symmetric cubic harmonics which form bases for the A_{1g} representation of the cubic group. Let $T_\mu(\Omega)$, $\mu = 1, 2, \dots$ be the set of such cubic harmonics, assigning μ in order of ℓ (when several T_μ have the

same ℓ their ordering is arbitrary). The T_μ are to be orthonormal, so that

$$\int d\Omega T_\mu(\Omega) T_{\mu'}(\Omega) = \delta_{\mu\mu'}, \quad (59)$$

Our method for obtaining the $T_\mu(\Omega)$ stems from the fact that the spherical harmonics of degree ℓ can be represented in cartesian form by homogeneous polynomials of degree ℓ in the coordinates x, y, z . It therefore follows that the completely symmetric cubic harmonics can be represented by completely symmetric homogeneous polynomials. A construction of the T_μ of given ℓ can therefore be accomplished by finding all the completely symmetric polynomials of degree ℓ which are eigenfunctions of ℓ_{op}^2 with eigenvalue $\ell(\ell+1)$ in units of \hbar^2 .

Completely symmetric polynomials of degree ℓ exist only if ℓ is even, in which case they are linear combinations of basic functions (uvw) with $u+v+w = \frac{1}{2}\ell$, defined as

$$(uvw) = \frac{1}{3!} \left[x^{2u} y^{2v} z^{2w} + x^{2u} y^{2w} z^{2v} + x^{2v} y^{2w} z^{2u} + x^{2w} y^{2u} z^{2v} + x^{2w} y^{2v} z^{2u} \right] \quad (60)$$

There will be as many different (uvw) as there are partitions of $\frac{1}{2}\ell$ with $u \geq v \geq w$. Our approach is simply to construct the matrices of ℓ_{op}^2 and unity among the (uvw) of a given ℓ , and to diagonalize, thereby obtaining ℓ_{op}^2 eigenfunctions. The eigenfunctions of eigenvalue $\ell(\ell+1)$ are the desired cubic harmonics.

The algebraic information needed to carry out the above-described process is not extensive and is listed below without comment.

$$(uvw)(u'v'w') = \frac{1}{3!} \left[(u+u', v+v', w+w') + (u+u', v+w', w+v') + (u+v', v+u', w+w') + (u+v', v+w', w+u') + (u+w', v+u', w+v') + (u+w', v+v', w+u') \right] \quad (61)$$

$$\int (uvw) d\Omega = \left(\frac{2v - 1}{2u + 1} \right) (u+1, v-1, w) \quad (62)$$

$$\int (uoo) d\Omega = \frac{4\pi}{2u + 1} \quad (63)$$

$$\begin{aligned} {}_0^2(uvw) &= (4u + 4v + 4w + 8uv + 8uw + 8vw)(uvw) \\ &\quad - 2u(2u - 1) [(u-1, v+1, w) + (u-1, v, w+1)] \\ &\quad - 2v(2v - 1) [(u+1, v-1, w) + (u, v-1, w+1)] \\ &\quad - 2w(2w - 1) [(u+1, v, w-1) + (u, v+1, w-1)] \end{aligned} \quad (64)$$

Once the cubic harmonics have been obtained, the Clebsch-Gordan coefficients can be obtained directly from the relation

$$D_{\mu\mu'\lambda} = \int T_\mu(\Omega) T_{\mu'}(\Omega) T_\lambda(\Omega) d\Omega \quad (65)$$

by multiplying together the necessary (uvw) according to Eq. (61) and evaluating the integral with Eqs. (62) and (63).

REFERENCES

1. F. E. Harris and H. J. Monkhorst, "Complete Calculations of the Electronic Energies of Solids", Phys. Rev. Letters 23, 1026 (1969).
2. A. A. Abrikosov, "The Equations of State of Hydrogen at High Pressures", U. S. S. R. Astronomical J. 31, 112 (1954).
3. F. C. Von der Lage and H. A. Bethe, "A Method for Obtaining Electronic Eigenfunctions and Eigenvalues in Solids with an Application to Sodium", Phys. Rev. 71, 612 (1947); see also D. D. Betts, "Solid Harmonics as Basis Functions for Cubic Crystals", Can. J. Phys. 37, 350 (1959).
4. See for example R. A. Bonham, J. L. Peacher, and H. L. Cox, Jr., "On the Calculation of Multicenter Two-Electron Repulsion Integrals Involving Slater Functions", J. Chem. Phys. 40, 3083 (1964).
5. F. E. Harris and H. J. Monkhorst, "Lattice Sums and Madelung Constants", Chem. Phys. Letters 4, 181 (1969).

6. J. C. Slater, "Symmetry and Energy Bonds in Crystals",
Quantum Theory of Molecules and Solids, Vol. 2 (McGraw-Hill,
New York, 1965), p. 130.
7. L. P. Bouckaert, R. Smoluchowski, and E. Wigner, "Theory of
Brillouin Zones and Symmetry Properties of Wave Functions
in Crystals", Phys. Rev. 50, 58 (1936).

THE USE OF THE GI METHOD IN BAND CALCULATIONS ON SOLIDS

William A. Goddard III and Patricia M. O'Keefe

A. A. Noyes Laboratory of Chemical Physics, California
Institute of Technology, Pasadena, California 91109

The application of the GI method to band calculations is discussed with primary emphasis on the alkali metals. The GI method goes beyond Hartree-Fock by allowing all orbitals to split, however the total wavefunction is modified so as to retain the correct spin symmetry. We find that the resulting band structure leads to a natural explanation of such puzzling properties as, the magneto-resistance, the prepeaking of the soft X-ray emission spectrum of Li, and the Mott Paradox.

I. INTRODUCTION

The basis of nearly all band calculations on solids has been the Hartree-Fock method. Of course in the application of this method to solids, it has been necessary to make a number of approximations and simplifications so that many results have been at best crude approximations to Hartree-Fock. Far more accurate Hartree-Fock calculations have been carried out on molecules in recent years. From this work it has become clear that although Hartree-Fock is adequate for treating some aspects and properties, it is totally inadequate in other ways.¹ One outcome of this work has been the GI method which goes beyond Hartree-Fock to remove some deficiencies and yet retains the conceptually useful orbital interpretation.¹ However the new GI orbitals differ significantly from the HF orbitals and lead to a much clearer interpretation of excited states of molecules² and of chemical reactions between molecules.^{3,4} For solids the GI method leads to bands of orbitals; however the resulting band structure differs qualitatively in some respects from the usual Hartree-Fock band model of the solid.

In this paper we will discuss the GI band structures for the alkali metals, a case in which quantitative results have already been obtained. We will find that the GI band structure leads to rather simple explanations of several properties which were difficult to understand on the basis of Hartree-Fock theory. These properties include the X-ray emission spectra, the magnetoresistance, the Hall effect, and the Mott paradox.

Before discussing the direct application to solids we will compare the Hartree-Fock and GI wavefunctions for some simple systems in order to illustrate the differences in the corresponding wavefunctions for solids.

II. THE WAVEFUNCTION

A. Introduction

First we will consider the ground state of the H_2 molecule. In this case the Hartree-Fock wavefunction is

$$\mathcal{A}[\phi(1)\phi(2)\alpha(1)\beta(2)] = \phi(1)\phi(2)[\alpha\beta - \beta\alpha] , \quad (1)$$

where \mathcal{A} is the antisymmetrizer (the determinant operator) and α and β refer to spin up and spin down. The orbital ϕ is double occupied (once with each spin) and is functionally optimized so as to minimize the total energy of (1). In Fig. 1 we show the HF orbital for two internuclear distances, $R = 1.4 a_0$ (the equilibrium value) and $R = 12 a_0$, at which point we should have essentially two free H atoms. We see that each orbital has σ_g symmetry, even at large internuclear distances. We can write this HF orbital as

$$\phi = \chi_l + \chi_r , \quad (2)$$

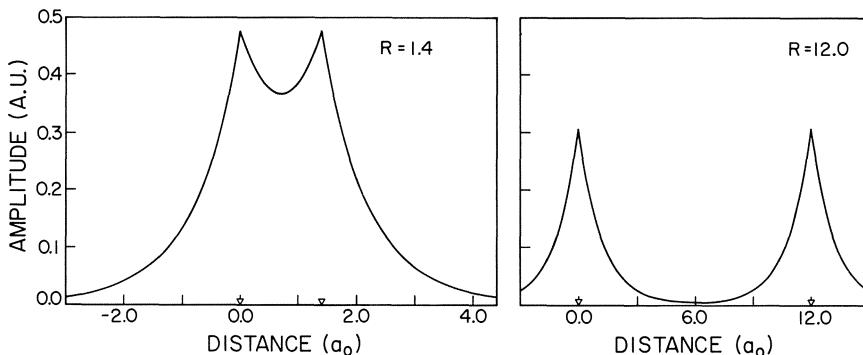


Fig. 1. The Hartree-Fock orbitals of H_2 .

where χ_ℓ and χ_r are localized on the left and right protons. (For large internuclear distances R (say a foot) χ_ℓ and χ_r have essentially no overlap.) Thus we can expand the spatial part of (1) as

$$\begin{aligned}\varphi(1)\varphi(2) = & [\chi_\ell(1)\chi_\ell(2) + \chi_r(1)\chi_r(2)] \\ & + [\chi_\ell(1)\chi_r(2) + \chi_r(1)\chi_\ell(2)].\end{aligned}\quad (3)$$

But if the nuclei are far apart, the spatial wavefunction should be just a product of the separated atomic functions $\chi_\ell\chi_r$, properly symmetrized,

$$\chi_\ell(1)\chi_r(2) + \chi_r(1)\chi_\ell(2). \quad (4)$$

Because of the spurious ionic terms in (3), the HF wavefunction behaves very poorly at large R as shown in Fig. 2.^{7,8} The problem here is just that the orbital φ is doubly occupied, whereas the correct wavefunction for large R , (4), requires singly occupied orbitals. The solution then is to allow each electron in (1) to be in a different orbital,

$$\alpha[\varphi_a(1)\varphi_b(2)\alpha(1)\beta(2)] = \varphi_a\varphi_b\alpha\beta - \varphi_b\varphi_a\beta\alpha \quad (5)$$

[this is called unrestricted Hartree-Fock (UHF)]. In this case we find that at large R , φ_a becomes a Hls function localized on the left and φ_b becomes a Hls function localized on the right. Thus the wavefunction dissociates to the correct state at $R = \infty$. However the ground state of H_2 should be a singlet state. To see if (5) has the correct spin symmetry we apply the spin raising operator

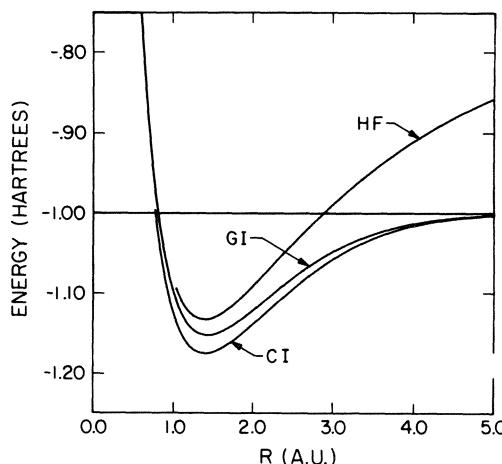


Fig. 2. The energy curves for H_2 .

to (5) obtaining,

$$\hat{S}^{\dagger} \mathcal{A}[\varphi_a \varphi_b \alpha\beta] = \mathcal{A} \varphi_a \varphi_b \alpha\alpha = (\varphi_a \varphi_b - \varphi_b \varphi_a) \alpha\alpha . \quad (6)$$

In order for (5) to correctly be a singlet state, (6) must be zero, which can only occur if φ_a is proportional to φ_b . Thus (5) treats dissociation correctly but leads to the wrong spin symmetry. In order to get the correct spin symmetry with (5) we must take

$$\varphi_a = \varphi_b ,$$

and hence we obtain the wrong dissociation. We really want both the correct spin symmetry and the correct dissociation, and we see now that (5) cannot do both. In order to accomplish these objectives, we will replace the antisymmetrizer in (5) by the Group Operator,⁹ G_i^{γ} , to obtain

$$G_i^{\gamma}[\varphi_a \varphi_b \alpha\beta] . \quad (7)$$

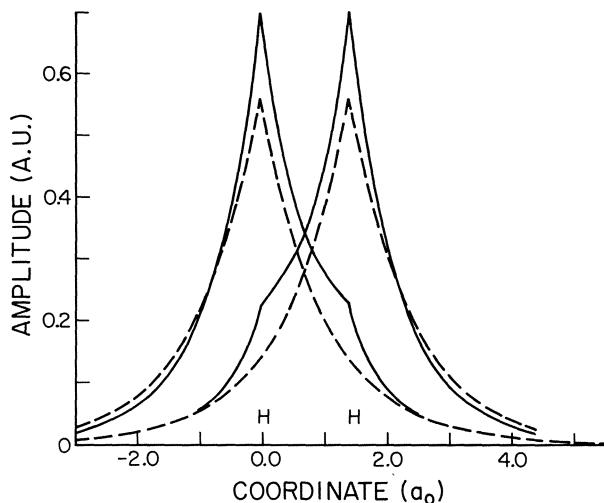
The group Operator, G_i^{γ} , has the property that regardless of the choice of φ_a and φ_b the total wavefunction (7) is an eigenfunction \hat{S}^2 and satisfies Pauli's principle.⁹ For a two electron singlet state this wavefunction has just the form

$$(\varphi_a \varphi_b + \varphi_b \varphi_a) (\alpha\beta - \beta\alpha) . \quad (8)$$

We now require that the orbitals of (8) be functionally optimized. The result is two coupled integro-differential equations¹⁰ to solve

$$\begin{aligned} H_a \varphi_a &= \epsilon_a \varphi_a \\ H_b \varphi_b &= \epsilon_b \varphi_b \end{aligned} \quad (9)$$

for the optimum orbitals of (8). The optimum orbitals¹ for H_2 (at $R = 1.4 a_0$) are shown in Fig. 3. As we pull H_2 apart each GI orbital changes continuously into a H_{1s} function on its center and hence the wave correctly dissociates into the ground state H atoms. But for all R the wavefunction (8) is correctly a singlet state.

Fig. 3. The GI orbitals of H_2 .

Solid lines are for $R = 1.4a_0$; dashed line is for $R = \infty$.

B. The Group Operator

For more electrons, the GI wavefunction^{1,9} can no longer be factored generally into spatial and spin parts as in (8). In addition, for more than two electrons there is generally more than one way to couple the spins to obtain a specific resulting spin S .

These different spin couplings⁷ are denoted by the i in G_i^γ . (In general the spin coupling must be optimized along with the orbitals.³)

The G_i^γ operator is expressed as

$$G_i^\gamma = \sum_r \xi_{ri} \sigma_{ri}^{\gamma} \omega_{ri}^{\bar{\gamma}} \quad (10)$$

in terms of Wigner projection operators, σ_{ri}^{γ} and $\omega_{ri}^{\bar{\gamma}}$ for the symmetric group (the permutation group).^{9,11} Here σ_{ri}^{γ} and $\omega_{ri}^{\bar{\gamma}}$ involve only permutation operators, where σ_{ri}^{γ} operates on spatial coordinates and $\omega_{ri}^{\bar{\gamma}}$ operates on spin coordinates. The $\omega_{ri}^{\bar{\gamma}}$ has the property that for any function, χ , of the spin coordinates of N electrons, $\omega_{ri}^{\bar{\gamma}} \chi$, is an eigenstate of \hat{S}^2 (the specific value of the spin being determined by the irreducible representation, $\bar{\gamma}$). The combination of σ_{ri}^{γ} and $\omega_{ri}^{\bar{\gamma}}$ in (10) is such that for any function of the spatial and spin coordinates, Θ , of N electrons, $G_i^\gamma \Theta$ satisfies Pauli's principle. We will generally consider wavefunctions of the form

$$G_i^\gamma \Phi \chi , \quad (11)$$

where Φ is a function of the spatial coordinates of N electrons and χ is a function of the spin coordinates. This wavefunction (11) will then always satisfy Pauli's principle and be an eigenfunction of \hat{S}^2 , regardless of the form of Φ and χ . We really do not have to deal with all these operators. The advantage of the group operator (10) is that for an operator which is independent of spin, such as the Hamiltonian H , the expectation value for (11) can be written in terms of spatial coordinates only, e.g., the energy of (11) is given by¹

$$E = \langle \Phi | \hat{H}_{ii} | \Phi \rangle / \langle \Phi | O_{ii} | \Phi \rangle \quad (12)$$

Thus we can avoid all unnecessary algebra relating to the spin coordinates and need deal with only a single spatial operator, O_{ii} . This operator is expanded as⁹

$$O_{ii}^\gamma = \frac{1}{\theta^\gamma} \sum_\tau U_{iit\tau}^\gamma \hat{\tau} , \quad (13)$$

where the sum is over all $N!$ spatial permutations τ and $U_{iit\tau}^\gamma$ is the representative matrix for τ in the γ irreducible representation of the symmetric group (θ^γ is a numerical constant). Further details on $U_{iit\tau}^\gamma$ are presented elsewhere.⁹ Here we will illustrate the forms of the resulting wavefunctions with several examples.

For comparing (11) to other types of wavefunctions it is convenient to transform to the following forms⁹

$$G_i^\gamma \Phi \chi = c \mathcal{A}[O_{ii} \Phi] \chi = c' \mathcal{A}[\Phi(\omega_{ii} \chi)] \quad (14)$$

(where c is a constant and unimportant here).

C. Optimization of Spin Coupling

For a three electron doublet, as for Li atom or H₃, we obtain¹² (substituting products for Φ and χ)

$$G_1^\gamma \Phi \chi = c \mathcal{A}[\varphi_{1a} \varphi_{1b} \varphi_{2a} (\alpha \beta \alpha + \beta \alpha \alpha)] \quad (15a)$$

$$G_f^\gamma \Phi \chi = c' \mathcal{A}[\varphi_{1a} \varphi_{2a} \varphi_{1b} (2 \alpha \alpha \beta - \beta \alpha \alpha - \alpha \beta \alpha)] \quad (15b)$$

[the last coupling operator for a given spin (γ) is denoted as $i = f$]. Thus each wavefunction can be expanded as a linear combination of determinants. The HF wave function is

$$\alpha[\varphi_1 \varphi_1 \varphi_2 \alpha\beta\alpha] \quad (16)$$

and the UHF wavefunction is

$$\alpha[\varphi_{1a} \varphi_{1b} \varphi_{2a} \alpha\beta\alpha] . \quad (17)$$

[Of course the optimum orbitals φ_{1a} , φ_{1b} , and φ_{2a} would be different for (15a), (15b) and (17).] We see that setting

$$\varphi_{1a} = \varphi_{1b} = \varphi_1$$

$$\varphi_{2a} = \varphi_2$$

in either (15a) or (15b) yields (16) and thus the usual HF wavefunction is a special case of both (15a) and (15b). Note that in (16) it is no restriction to take

$$\langle \varphi_1 | \varphi_2 \rangle = 0$$

and in (15b) and (17) we can take

$$\langle \varphi_{1a} | \varphi_{2a} \rangle = 0$$

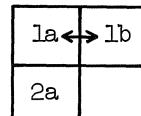
without changing the energy. However in (15a) we can make no such orthogonality restrictions without raising the energy. The optimum orbitals for the wavefunction (15a) are called the G1 orbitals and the optimum orbitals for (15b) are called the GF orbitals.¹ These two sets of orbitals are generally different and generally lead to different energies (both better than HF).

To obtain the spin coupling of (15a), we first couple two orbitals (φ_{1a} and φ_{1b}) into a singlet pair and then couple in the third orbital (φ_{2a}) to obtain a doublet. In (15b) we first couple two orbitals (φ_{1a} and φ_{2a}) into a triplet pair and then couple in the third orbital (φ_{1b}) to obtain a doublet. We will denote the coupling in (15a) as (18a),

which just indicates that orbitals φ_{1a} and φ_{1b} are first

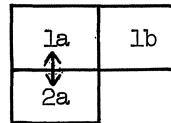
symmetrically coupled (corresponding to singlet pairing)

and then φ_{2a} is coupled so as to yield a doublet state. The coupling



(18a)

in (15b) is denoted as (18b), which indicates that φ_{1a} and φ_{2a} are first antisymmetrically coupled (corresponding to triplet pairing) and then φ_{1b} is coupled so as to yield a doublet state.



(18b)

That these wavefunctions (15) should correspond to such spin couplings is determined by the representation matrices $\underline{U}_{\tau}^{\gamma}$ used in (13). In order to avoid any bias inherent in the specific choice of the representation matrices, we must simultaneously optimize both the orbitals and the representation matrices.³ The resulting function can be written as³

$$G_1^{\gamma L \Phi \chi}, \quad (19)$$

where now the G_{11} is based on a new set of $\underline{U}_{\tau}^{\gamma}$. This wavefunction (19) can be expanded as

$$G_1^{\gamma L \Phi \chi} = \sum_i c_i G_i^{\gamma \Phi \chi}$$

in terms of the standard group operators and can be written as

$$G_1^{\gamma L \Phi \chi} = c \mathcal{Q} [\Phi (\sum_i c_i^* W_{ii}^{\gamma} \chi)] \quad (20)$$

Optimizing the representation matrices in (19) is equivalent to optimizing the spin-coupling terms in (20) and we refer to the optimum wavefunction of form (19) as the Spin-Coupling Optimized GI (or SOGI) wavefunction.³

D. The Li Atom

For the Li atom the resulting SOGI wavefunction is nearly identical with the GI wavefunction, the energy being only .000005 h lower.³ The energies for the ground state of Li for the various methods are^{3,13}

Hartree-Fock	-7.432725
UHF	-7.432749
GF	-7.432813
GI	-7.447560
SOGI	-7.447565
Experimental	-7.478

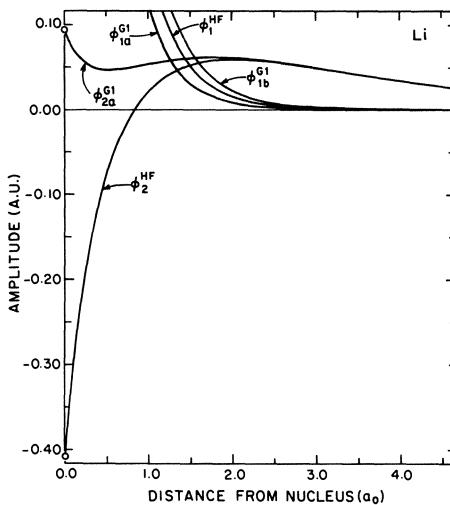


Fig. 4. The G1 and HF orbitals of Li atom.

Thus the SOGI wavefunction for Li accounts for about 1/3 of the correlation or many-body effects left out of the Hartree-Fock wavefunction. The resulting G1 orbitals¹² are plotted in Fig. 4 (the G1 and SOGI orbitals are nearly identical) where they are compared to the Hartree-Fock orbitals. We see that the core orbitals [ϕ_{1a} and ϕ_{1b} of (15a)] are similar to the Hartree-Fock orbital, although split somewhat. The G1 valence orbital (ϕ_{2a}) is also similar to the Hartree-Fock orbital in the valence region but quite different in the core region. In particular, the G1 valence orbital is smooth and nodeless. This property is used in obtaining the effective potentials due to the core orbitals.¹⁴

E. Four Electrons

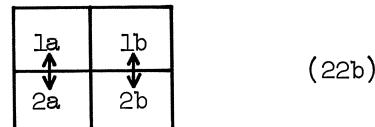
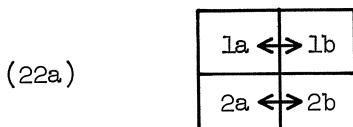
For a four-electron singlet state the two G_i^γ operators can be expanded as¹⁵

$$G_1^\gamma \Phi \chi = c \mathcal{A} [\phi_{1a} \phi_{1b} \phi_{2a} \phi_{2b} (\alpha\beta - \beta\alpha)(\alpha\beta - \beta\alpha)] \quad (21a)$$

and

$$G_f^\gamma \Phi \chi = c \mathcal{A} [\phi_{1a} \phi_{2a} \phi_{1b} \phi_{2b} (2\alpha\beta\beta - \alpha\beta\alpha\beta - \alpha\beta\beta\alpha - \beta\alpha\alpha\beta - \beta\alpha\beta\alpha + 2\beta\beta\alpha\alpha)] \quad (21b)$$

Corresponding to (18) the orbital coupling in (21a) is denoted as (22a) and the orbital coupling in (21b) is denoted as (22b). Just

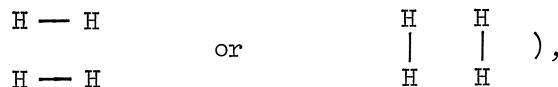


as for (15b) we can take

$$\langle \phi_{1a} | \phi_{2a} \rangle = 0 \quad \text{and} \quad \langle \phi_{1b} | \phi_{2b} \rangle = 0 \quad (23)$$

for (21b), but no orthogonality restrictions can be made on the orbitals in (21a).

Comparing (22) and (21) with (18) and (15), we would expect the SOGI wavefunction for LiH to be nearly identical with the G1 wavefunction and it is. The energy drops by only .000001 h in going from G1 to SOGI for LiH (at $R = 3.015 a_0$).^{22,3} On the other hand for a system such as square H_4 in which all orbitals have roughly the same energy and in which there is more than one way to consider the bonding (e.g., for square H_4 we could have



we would expect the coupling to be nearly GF. In fact, for the lowest singlet state of H_4 ($^1B_{1g}$) near the optimum bond distance ($2.5 a_0$), the optimum SOGI wavefunction is nearly the GF wavefunction, the energy being only .002 h lower.¹⁶ The resulting GF orbitals are shown in Fig. 5. Note here that ϕ_{1a} and ϕ_{2a} are eigenfunctions of the one-electron Hamiltonian \bar{H}_a ,

$$\bar{H}_a \phi_{ka} = \epsilon_{ka} \phi_{ka} \quad k = 1, 2 \quad (24a)$$

and are localized mainly along one diagonal of H_4 , while ϕ_{1b} and ϕ_{2b} are eigenfunctions of \bar{H}_b

$$\bar{H}_b \phi_{kb} = \epsilon_{kb} \phi_{kb} \quad k = 1, 2 \quad (24b)$$

and localized along the other diagonal. In addition, ϕ_{1a} and ϕ_{2a} are equivalent to ϕ_{1b} and ϕ_{2b} except for being rotated by 90° . For the UHF wavefunction

$$\alpha[\phi_{1a} \phi_{2a} \phi_{1b} \phi_{2b} \alpha\beta\beta] \quad (25)$$

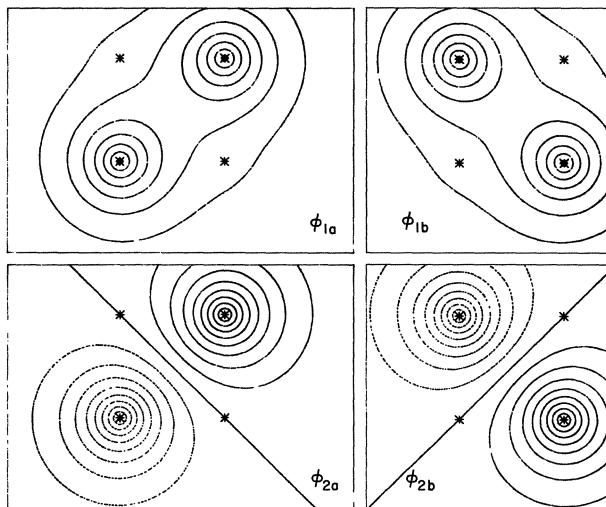


Fig. 5. The GF orbitals for square H_4 (side = $2.5a_0$).

equation (23) would also be satisfied and the orbitals would be solutions of equations like (24) except that the H_a and H_b would be different. However for (25), orbitals ϕ_{1a} and ϕ_{2a} are always associated with α spin and ϕ_{1b} and ϕ_{2b} with β spin. As a result if the UHF orbitals were similar to the GF orbitals of Fig. 5, we would find a net up spin near centers 1 and 3 and a net down spin near centers 2 and 4. That is, the UHF wavefunction would lead to a sort of antiferromagnetic structure for square H_4 . This problem occurs because (25) is a mixture of spin-states rather than a pure singlet as is (21). Thus although the GF orbitals are formally similar to the UHF orbitals, we must not associate any definite spin with the GF orbitals. For H_4 the GF wavefunction (21b) is a singlet and there is no net spin density at any point in the molecule.

F. Six Electrons

For six electron singlet states there are five possible G_i^γ operators. Of these the two most important are G_1^γ and G_f^γ , which correspond to orbital couplings of

(25a)

$1a \leftrightarrow 1b$	
$2a \leftrightarrow 2b$	
$3a \leftrightarrow 3b$	

and

(25b)

$1a$	$1b$
$2a$	$2b$
$3a$	$3b$

respectively. For Li_2 we find that¹⁷ φ_{1a} and φ_{1b} are Li core functions on the left center, φ_{2a} and φ_{2b} and Li core functions on the right center, and φ_{3a} and φ_{3b} are the valence or bonding orbitals. Comparing (24) with (18a) we would expect the Li_2 SOGI wavefunction to be nearly exactly G1 and it is (the energy difference being less than .000001 h).¹⁷ The core functions are nearly identical with the atomic core functions and the valence orbitals are as shown in Fig. 6 (this shows one bonding orbital, the other is equivalent but centered on the right nucleus).¹⁷ Thus the G1 orbitals are σ functions but do not have inversion symmetry (the Hartree-Fock bonding orbital of $\overline{\text{Li}}_2$ is a σ_g function).

Comparing (25) with (22b) we would expect a system such as hexagonal H_6 to be nearly GF. In fact, we find that it is,¹⁸ the GF energy is -3.2902 compared to a SOGI energy of -3.2943. Thus again the SOGI wavefunction is nearly the same as the GF wavefunction. Just as for H_4 , the GF wavefunction involves two sets of orbitals ($\varphi_{1a}, \varphi_{2a}, \varphi_{3a}$) and ($\varphi_{1b}, \varphi_{2b}, \varphi_{3b}$) which are solutions of equations like (23). We find that the a orbitals of H_6 are concentrated on centers 1, 3, and 5 (numbering in sequence around the ring) and the b orbitals are concentrated on centers 2, 4, and 6. Thus the GF orbitals are symmetry functions of D_{3h} rather than of the full molecular symmetry group D_{6h} .

G. Alkali Metals

On the basis of the results on Li, Li_2 , H_4 , and H_6 we would expect that in crystalline Li the core electrons should be essentially G1 coupled and the valence electrons should be GF coupled as in (26)

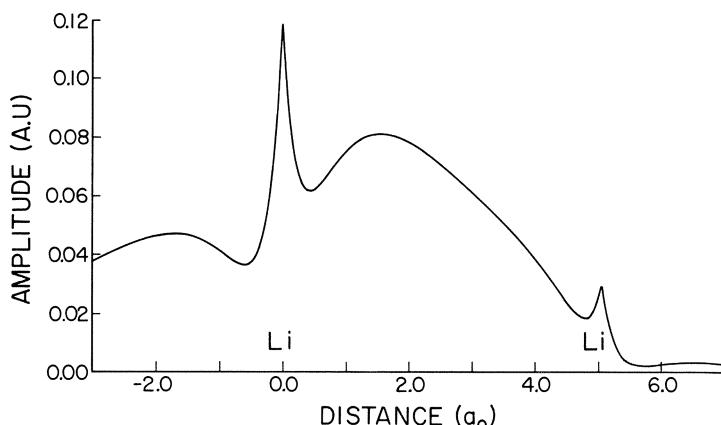


Fig. 6. A G1 bonding orbital of Li_2 .

(here N is the number of Li atoms in the crystal and $m = N/2$). Since the core orbitals in the metal should be essentially the same as the Li atom core orbitals, we will ignore these orbitals in the following discussion (except for including their potential as seen by the conduction electrons).¹⁴

G1	1a	\leftrightarrow	1b	(26)
	2a	\leftrightarrow	2b	
	⋮		⋮	
	N_a	\leftrightarrow	N_b	
	$(N+1)_a$	\downarrow	$(N+1)_b$	
	$(N+2)_a$	\downarrow	$(N+2)_b$	
	$(N+m)_a$	\downarrow	$(N+m)_b$	

Just as for H_4 and H_6 , the a and b conduction orbitals of Li metal are solutions of

$$\begin{aligned} H_a \varphi_{ka} &= \epsilon_{ka} \varphi_{ka} \\ H_b \varphi_{kb} &= \epsilon_{kb} \varphi_{kb} \end{aligned} \quad (27)$$

Just as for Li and Li_2 these conduction orbitals need not be orthogonal to the Li core orbitals, but the a conduction orbitals are mutually orthonormal as are the b orbitals,¹⁹

$$\langle \varphi_{ia} | \varphi_{ja} \rangle = \delta_{ij}$$

$$\langle \varphi_{ib} | \varphi_{jb} \rangle = \delta_{ij}$$

Generally for Hartree-Fock wavefunctions to lead to the correct spatial symmetry for the total wavefunction, each orbital must transform as a symmetry function for the total spatial symmetry group.²⁰ As we have seen above, this is no longer true for the GI wavefunctions. For example, the H_2 and Li_2 orbitals were symmetry functions only of D_{co} not D_{coh} . In addition the GF orbitals of

square H_4 and hexagonal H_6 were symmetry functions of D_{2h} and D_{3h} , respectively, rather than of the full molecular symmetry groups, D_{4h} and D_{6h} , respectively. In all of these cases the total energy is lowered by the orbitals concentrating on different centers. In this way one obtains a sort of static correlation between the electrons.

Despite this lower symmetry in the orbitals, the many-electron wavefunction has the correct spatial symmetry. For example, for H_2

$$\begin{aligned} i\varphi_a &= \varphi_b \\ i\varphi_b &= \varphi_a, \end{aligned} \tag{28}$$

where i is the inversion operator. Thus

$$i(\varphi_a \varphi_b + \varphi_b \varphi_a) = \varphi_a \varphi_b + \varphi_b \varphi_a$$

and the total wavefunction is correctly a $^1\Sigma_g^+$ state of $D_{\infty h}$ symmetry. Similarly for the H_4 wavefunction we have

$$\begin{aligned} C_4 \varphi_{1a} &= \varphi_{1b}; & C_4 \varphi_{1b} &= \varphi_{1a} \\ C_4 \varphi_{2a} &= \varphi_{2b}; & C_4 \varphi_{2b} &= -\varphi_{2a} \end{aligned} \tag{29}$$

and hence C_4 operating on (10) leads to minus (10). Thus (10) correctly transforms as a $^1B_{1g}$ state of the D_{4h} symmetry group, despite the lower symmetry of the orbitals.

In the same way the $\{\varphi_{ka}\}$ orbitals for bcc Li would be expected to concentrate on alternate sites and thus to have sc symmetry. That is, we expect these functions to transform as

$$\underline{R} \varphi_{ka}(r) = e^{ik \cdot \underline{R}} \varphi_{ka}(r) \tag{30a}$$

for all \underline{R} belonging to the simple cubic (sc) subgroup of the bcc space group. [In (30) we now take k to be the wavevector for the orbital rather than a simple numerical index as in (27).] Similarly

$$\underline{R} \varphi_{kb}(r) = e^{ik \cdot \underline{R}} \varphi_{kb}(r) \tag{30b}$$

for simple cubic \underline{R} . We will denote the transformations which are in the bcc space group but not in the sc space group as \underline{R}_{ab} (these transformations take corner sites into body-centered sites). In order that the total wavefunction have the correct symmetry, we have²¹

$$\begin{aligned} \underline{R}_{ab} \varphi_{ka}(r) &= e^{ik \cdot \underline{R}_{ab}} \varphi_{kb}(r) \\ \underline{R}_{ab} \varphi_{kb}(r) &= e^{ik \cdot \underline{R}_{ab}} \varphi_{ka}(r) \end{aligned} \tag{31}$$

which is analogous to (28) and (29).

If we let R in (30) be a primitive translation in the x direction

$$R = a \hat{e}_x , \quad (32)$$

where a is the cubic lattice constant, we see that two orbitals ϕ_{ka} and $\phi_{k'a}$ transform the same way under (32) if

$$e^{i(k_x - k'_x)} = 1 .$$

Since (32) and the corresponding translations in the x and y directions form primitive translations for the sc space group, the set of inequivalent k -vectors can be taken as the set

$$-\frac{\pi}{a} < (k_x, k_y, k_z) \leq \frac{\pi}{a} . \quad (33)$$

That is the first Brillouin zone (BZ) for describing the GI orbitals is just the simple cube defined by (33). This cube is half the volume of the usual bcc BZ; in fact, the boundaries of the usual bcc BZ are the boundaries of the second sc BZ.

If there are N different Li atoms (in a microcrystal), then there are N conduction electrons and hence $N/2$ a orbitals and $N/2$ b orbitals [$m = N/2$ in (27)]. But there are $N/2$ states in the first sc BZ. In constructing the total wavefunction we separately consider the a states and the b states. We must have $N/2$ a orbitals and hence we could just exactly fill the first BZ. We will find below that for large lattice parameters this is just what happens and we obtain an insulator. However at equilibrium the first GI band overlaps the second and hence each is partially occupied (in fact, the calculated Fermi surface is very nearly a sphere). The b orbitals lead to an equivalent energy band, and hence exactly the same considerations apply to them; consequently we need only consider the a orbital band scheme.

There are now two band equations to solve, (27), rather than one as in HF. However because of symmetry we need only solve one of them (for bcc Li). The operator in (27) can be written as

$$H_a = -\frac{1}{2} \nabla^2 + U_a ,$$

where U_a is an integral operator representing the potential due to the other $N-1$ conduction electrons and to the N core states. This potential is more complicated than for bcc HF (since we now have

essentially two atoms per cell). Instead of empirically determined pseudopotentials, we use the effective potentials¹⁴ obtained from ab initio GI calculations on the Li atom.^{3,12} Note that since the GI valence orbital of Li is smooth in the core region,^{3,12} it is not necessary or correct to use orthogonalized plane waves (OPW) for the band calculations. For Li metal, simple plane wave expansions converge quite rapidly.²²

III. RESULTS AND DISCUSSION

The calculated band structure of Li metal is shown in Fig. 7 (for the equilibrium lattice spacing of $a = 6.575 a_0$). Here the special symmetry points are $\Gamma = 2\pi/a (0,0,0)$; $X = 2\pi/a (\frac{1}{2},0,0)$; $M = 2\pi/a (\frac{1}{2},\frac{1}{2},0)$; $R = 2\pi/a (\frac{1}{2},\frac{1}{2},\frac{1}{2})$. The calculated Fermi level is indicated; this was calculated on the basis of band calculations at points throughout the BZ based on increments of $0.1 2\pi/a$ in each direction.

A. The Mott Paradox

Calculations were also carried out as a function of lattice constant. We found that as a increases the lowest R_1 , M_1 , and X_1 points drop (and Γ_1 increases) so that for $a > 11 a_0$ the lowest band no longer overlaps the second band, as shown in Fig. 8 (for $a = 18 a_0$). Thus for large a , bcc Li is correctly predicted to be an insulator, whereas for small a , bcc Li is correctly predicted to be a metal. For the HF wavefunction the first band is half full at all a and hence the system would be expected to be a metal even for $a = \infty$. This paradox is called the Mott paradox²³ and really arises from the double occupation restriction in the HF wavefunction (for H₂ we saw that the same restriction also leads to an incorrect behavior for large R).

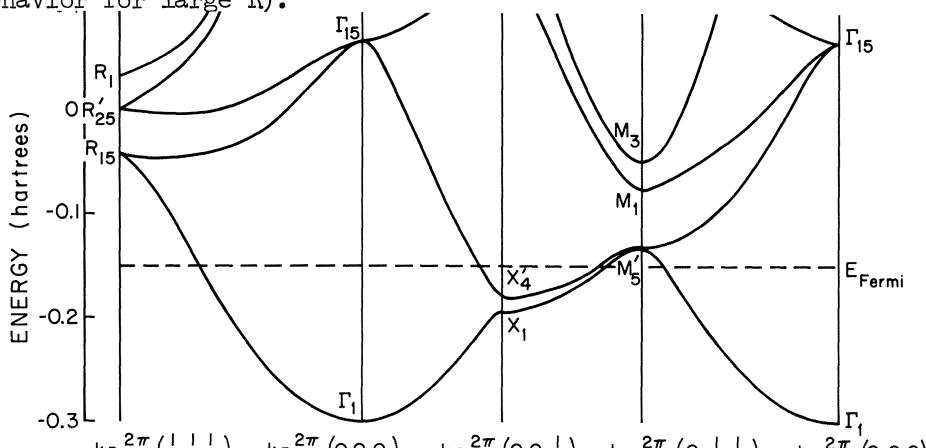


Fig. 7. The GI band structure of bcc Li ($a = 6.575 a_0$).

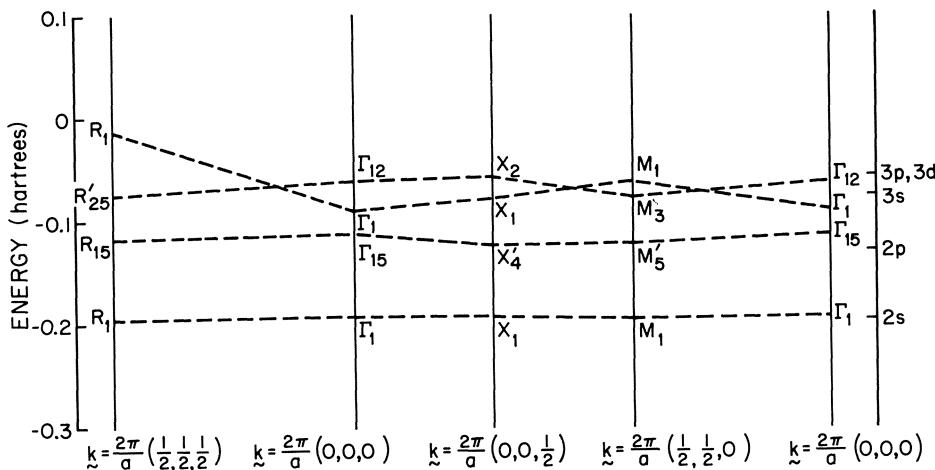


Fig. 8. The GI band structure for Li ($a = 18 a_0$).
(The Li atomic energies are shown at the right).

For the GI wavefunction the bands behave correctly as $a \rightarrow \infty$ and hence no Mott paradox occurs.

B. The X-ray Emission Spectra

For lithium the X-ray emission spectrum involves transitions from the conduction band down to empty Li core orbitals, which are s-like. Thus the emission spectrum is a measure of the density of p-states in the conduction band (assuming the transition matrix elements to vary smoothly). The p-density of states should start at zero for the bottom of the conduction band and increase monotonically until the Fermi surface touches the BZ boundary. If the state at the BZ boundary has p-character, then one may obtain a peak in the p-density of states and hence a peak in the X-ray emission spectrum corresponding to this energy.

In fact for Li the X-ray emission spectrum²⁴ shows just such a peak about 0.9 eV below the band edge²⁵ as shown²⁶ in Fig. 9. However experiment and reliable Hartree-Fock calculations²⁷ both show that for the alkali metals the Fermi surface is not distorted enough to touch the usual bcc BZ. Thus on the basis of the HF band scheme we would expect a monotonic increase in X-ray emission intensity, as indicated in Fig. 9, in dramatic disagreement with the experimental results.

On the other hand we note that for the GI energy an X_{14}' state

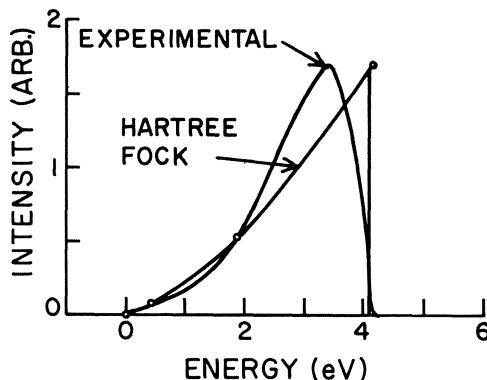


Fig. 9. The X-ray emission spectrum of Li (after parameter, Ref. 26).

is about 0.8 eV below the Fermi level. Since X_4' is of p-symmetry the GI energy band should lead to a peak in the p-density of states about 0.8 eV below the Fermi level and hence a prepeaking in the X-ray emission about 0.8 eV below the edge, in good agreement with the experimental value.

For Na the dominant X-ray transition is down to the Na 2p levels and hence the intensity is a measure of the s+d density of states in the conduction band. Since the second band starts with the X_1 states and since this state is of even parity (s- and d-like), we would expect the s+d density of states and the X-ray emission to be something like that shown in Fig. 10. In fact the X-ray emission spectrum of Na shows just such an increase in intensity about 0.4 eV below the band edge.²⁸

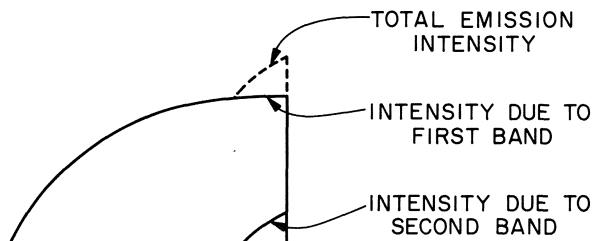


Fig. 10. Schematic indication of the expected X-ray emission spectrum for Na.

Thus for both Li and Na the GI band scheme allows a natural explanation for some anomalous properties of the X-ray emission spectrum.

Some other explanations have been offered for the prepeaking in Li, for example, involving the distortions and modifications in the electronic states due to the localized hole in the core-band.²⁷ It is not clear whether this type of explanation can account for the differences between Li and Na.

C. Magnetoresistance and Hall Effect

As discussed elsewhere³⁰ the presence of small energy gaps at the Fermi surface (for Li the calculated gap in Fig. 7 is $E_g = 0.14$ eV) can lead to dramatic effects upon the magnetoresistance. In the GI band scheme we have two partially filled bands, but at the BZ boundaries these bands are separated only by an amount comparable to E_g . Thus even at relatively small magnetic fields we can have band-to-band transitions (magnetic breakdown³¹) at these points. At intermediate magnetic fields this leads to extended orbits and results in large increases in the resistance.³⁰ Eventually for large fields band-to-band transitions dominate, and we obtain essentially circular orbitals. This results in a saturation of the magnetoresistance at these high fields.³⁰

Experimental measurements^{32,33} of the magnetoresistance have been carried out on K, but we have not yet done calculations on this system. However based on the Li and Na calculations we would expect the energy gap of K to be in the range of 0.02 eV to 0.06 eV. In Fig. 11 we show the resulting magnetoresistance in the [001] direction for K calculated for two values of E_g . For $E_g = 0.04$ the magnetoresistance is essentially linear from 12 to 55 kG and saturates at 110 kG. This linear region is in agreement with the linear range found by Bowers *et al.*³² on single crystals of K. For $E_g = 0.02$ we find a short linear region and saturation by 15 kG. This is the sort of behavior found by Babiskin *et al.*³³ Basically the experimental problem is the great sensitivity of the magnetoresistance of alkali metals to strain. Babiskin *et al.*³³ assume that all of the high field linear region is due to strain and subtract it out, whereas Bowers *et al.*³⁴ find that going to purer and presumably more strain free samples leads to larger slopes in the high field region. We are beginning preparations for the GI band calculations on K. Perhaps these calculations could help resolve some of the experimental uncertainties.

The GI band structure (plus magnetic breakdown leads to a Hall coefficient that decreases with field, eventually becoming the free electron value for large fields.³⁰ Such a decrease has been

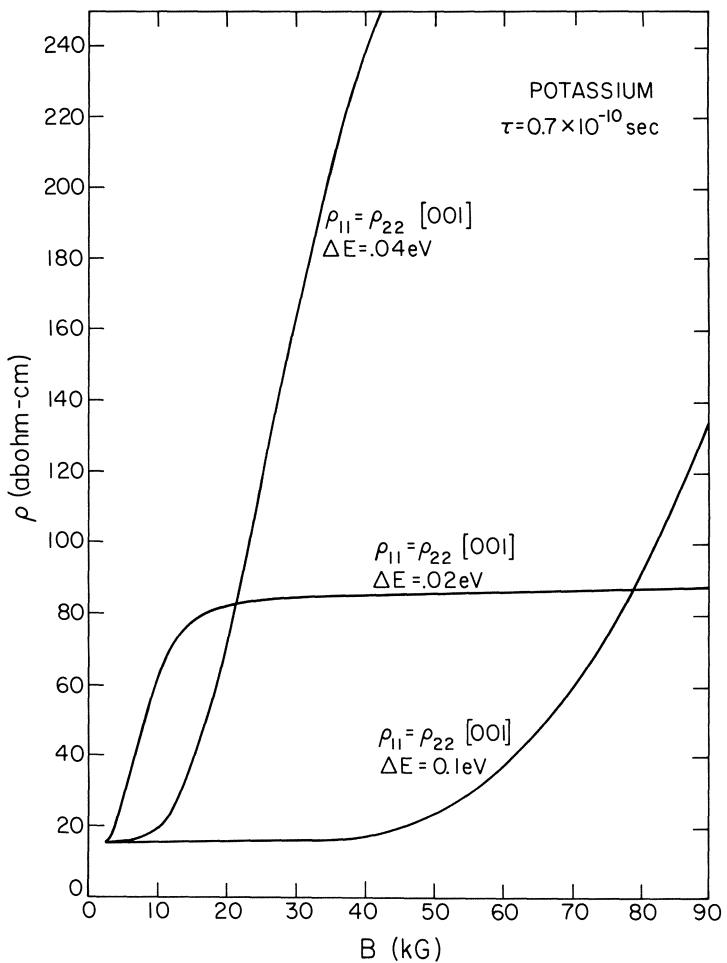


Fig. 11. The calculated magnetoresistance of K.

observed by Penz for K and a value larger than the free electron value has been observed for Na by Chambers and Jones.³⁶

D. Other Properties

Position annihilation experiments³⁷ have been carried out on Li and interpreted to yield the wave vector to the Fermi surface, k_F , in the [001], [011], and [111] directions. Within experimental error the GI Fermi surface and the HF Fermi surface agree with these results.

De Haas-van Alphen experiments³⁸ have been done on several alkali metals, showing a rather spherical Fermi surface. However these experiments are carried out at high fields (e.g., 50 kG for K) which should be within the saturation region for magnetic breakdown. Thus one would expect de Haas-van Alphen behavior as if the GI surface were nearly spherical.

Cyclotron resonance results³⁹ on the alkali metals have been interpreted in terms of a simple nearly spherical Fermi surface. However these experiments are also carried out for $\omega\tau > 1$ and hence well into the magnetic breakdown region.

Because of the smaller BZ, optical transitions become allowed for GI which were forbidden for the Hartree-Fock band structure. As a result we would expect an increase in absorption intensity below the usual band-to-band threshold (1.2 eV for K). Early experiments by Mayer and El Naby⁴⁰ showed a peak in this region (centered at 0.62 eV for K). This was interpreted by Cohen and Phillips⁴¹ as due to collective resonance states and by Overhauser⁴² as due to spin density waves. In fact, Overhauser was able to account for the shape of the Mayer-El Naby band through use of the spin density wave hypothesis.⁴² Unfortunately more recent experiments by Smith⁴³ give no indication of such low energy bands for Na and K; although an extra increase is observed in the Drude absorption starting at 0.5 eV for K. Although new band-band transitions are allowed by the GI band structure in this region of energy, it is not clear yet just how large the matrix elements will be and hence how important these transitions will be compared to the Drude absorption.

E. Summary

The discussion here was not meant to provide an exhaustive quantitative explanation of each property but rather to indicate some of the motivation in going beyond Hartree-Fock to develop the GI method for band structure calculations. We see that the GI band

structure leads to a coherent natural explanation for several anomalous properties for the alkali metals.

IV. OTHER METHODS

There have been other approaches suggested for going beyond Hartree-Fock.

A. Unrestricted Hartree-Fock

The most common has been the UHF method in which the Slater determinant form of the total wavefunction is retained but the double occupation restriction is relaxed. As we noted for H₂, the spins of the split orbitals no longer cancel and we obtain spurious spin densities which must entirely disappear as we improve the wavefunction toward the exact one.

The use of UHF for solids dates back to Slater⁴⁴ who suggested that for bcc Na the up-spin orbitals would localize on one sc sublattice and the down-spin orbitals on the other. Overhauser has suggested that the resulting spin density waves one gets for a solid actually have physical significance and has discussed the possibility of experimentally detecting these spin density waves.⁴⁵ In any case, Reitz and Overhauser⁴⁵ suggested that the spin density waves could account for the magnetoresistance behavior of K and as mentioned in III.D Overhauser⁴² through use of the same hypothesis was able to account for the Mayer-El Naby low-energy band of K. Overhauser has also suggested that rather than the doubly-occupied orbitals splitting to yield spin density waves, the orbitals may remain doubly occupied but take on lower spatial symmetry (charge density waves).⁴⁵ For UHF, such a decrease in symmetry of the orbitals (say from bcc to sc) would lead to a decrease in symmetry of the many-electron wavefunction (say from bcc to sc). This is in contrast to the case for the GI wavefunction where sc GI orbitals lead to a many-electron wavefunction with bcc symmetry. Some of Overhauser's discussions can be carried over in discussing the GI wavefunctions. However the decreases in the symmetry of the total wavefunction implied in the spin density waves and charge density waves probably have no physical significance.

In addition to Overhauser's work, there have been several UHF band calculations. This approach is sometimes referred to as different bands for different spins. The most extensive work has been by Wilson⁴⁷ who has considered such magnetic systems as NiO and MnO.

In addition Calais⁴⁸ has used the UHF approach to calculate the cohesive energy of Li.

B. Alternate Molecular Orbitals

For a cyclic system of N atoms with one valence electron per atom (e.g., H_N), there are N HF orbitals of which half (assuming N to be even) are doubly occupied

$$\{\varphi_k; \quad k = 1, \dots, m = \frac{N}{2}\}$$

and half are empty

$$\{\varphi_{m+k}; \quad k = 1, \dots, m\} .$$

The total HF wavefunction is

$$\mathcal{A}[\varphi_1 \varphi_1 \varphi_2 \varphi_2 \cdots \alpha \beta \alpha \beta \cdots] ;$$

we can improve the energy by (1) allowing each pair to split

$$\begin{aligned} \theta_k &= \varphi_k + c_k \varphi_{2m-k} \\ \bar{\theta}_k &= \varphi_k - c_k \varphi_{2m-k} \end{aligned} \quad (34)$$

and (2) spin projecting the total wavefunction to get the correct total spin symmetry⁴⁹

$$\sigma^S \mathcal{A}[\theta_1 \bar{\theta}_1 \theta_2 \bar{\theta}_2 \cdots \alpha \beta \alpha \beta \cdots] . \quad (35)$$

In this method we assume that each φ_i is a linear combination of atomic orbitals, with one atomic orbital on each center. Then for a system with spatial symmetry group, C_{Nv} , the orbitals $\{\varphi_i\}$ are completely determined by symmetry and the occupied and empty orbitals to be associated in (34) are also determined by symmetry.⁴⁹ All that is left to determine is the set of coefficients $\{c_k\}$.

This method is called the alternant molecular orbital (AMO) method⁴⁹ and has been applied to the π orbitals of benzene⁵⁰ and to the H₆ system⁵¹ and with further simplifications to other systems.⁴⁹ One problem is that because of the pairing (34), there is only one allowed state for each spatial symmetry. Thus for benzene the singlet state must be⁵⁰ $1A_{1g}$. Thus this method is not expected to be useful for band calculations in which one wants the whole spectrum of states. However it could be useful for cohesive energy.^{48,49}

An additional problem is that the method is really based upon using one basis function per atom. This is not so much of a restriction for the ground state of solid hydrogen, but it would be for solid Li since here we should include at least the 2s and 2p atomic functions on each center.

V. CONCLUSIONS

We find that the band structure of alkali metals as obtained from the GI method explains several puzzling properties (e.g., the prepeaking in the soft X-ray emission, the Mott paradox, and the magnetoresistance) not accounted for by the HF band structure. However the GI method leads to an interpretation of the electronic wavefunction in terms of the conceptually useful band orbitals. Thus it would appear that the GI method might be of general use in treating the wavefunctions and properties of solids.

VI. FOOTNOTES

1. W. A. Goddard III, "Improved Quantum Theory of Many-Electron Systems. II. The Basic Method", Phys. Rev. 157, 81 (1967).
2. (a) C. F. Melius and W. A. Goddard III, "The Excited States of LiH According to the GI Spin Generalized SCF Method", to be published;
(b) D. L. Huestis, W. A. Goddard III, and T. H. Dunning, Jr., "The Projected GI Method and the Excited States of H₂", to be published.
3. R. C. Ladner and W. A. Goddard III, "An Improved Quantum Theory of Many-Electron Systems. V. The Spin-Coupling Optimized GI Method", J. Chem. Phys. 51, 1073 (1969).
4. (a) W. A. Goddard III and R. C. Ladner, "The Optimum Orbitals for the H₂ + D ⇌ H + HD Exchange Reaction", Int. J. Quantum Chem. 3S, 63 (1969);
(b) R. C. Ladner and W. A. Goddard III, "Potential Energy Surfaces for the Reactions of Small Systems. I. Ab Initio SOGI Energies and Orbitals for Linear H₃ and LiH₂", to be published.
(c) W. A. Goddard III and R. C. Ladner, "The Orbital Description of the Reactions of Small Molecules", to be published.
(d) W. A. Goddard III, "The Orbital Phase Continuity Principle and Selection Rules for Chemical Reactions", to be published.
5. Figure 1 is based on the HF calculations of reference 6.

6. S. Fraga and B. J. Ransil, "Potential Curve for the Interaction of Two Hydrogen Atoms in the LCAO MO SCF Approximation", *J. Chem. Phys.* 35, 1967 (1961).
7. In Fig. 2, the HF energy curve is from ref. 6, the GI energy curve is from ref. 4b, and the exact energy curve is from W. Kolos and L. Wolniewicz, "Accurate Adiabatic Treatment of the Ground State of the Hydrogen Molecule", *J. Chem. Phys.* 41, 3663 (1964).
8. (a) We use Hartree atomic units in which $\hbar = 1$, $|e| = 1$, and $m_e = 1$ ($|e|$ is the magnitude of the charge of an electron and m_e is the mass of the electron). In these units the unit of length is the Bohr, $1 a_0 = .529178 \text{\AA}$, and the unit of energy is the Hartree, $1 h = 27.2117 \text{ eV} = 2 R_\infty$. These units are used almost exclusively for calculations on atoms and molecules. In solid state work another common set of units are the Rydberg atomic units in which the unit of energy is the Rydberg, ($R_\infty = 13.6058 \text{ eV}$). In this case, we could take $|e| = \sqrt{2}$, $\hbar = \sqrt{2}$, and $m_e = 1$;
 (b) B. N. Taylor, W. H. Parker, and D. N. Langenberg, The Fundamental Constants and Quantum Electrodynamics (Academic Press, New York, 1969).
9. W. A. Goddard III, "Improved Quantum Theory of Many-Electron Systems. I. Construction of Eigenfunctions of \hat{S}^2 which Satisfy Pauli's Principle", *Phys. Rev.* 157, 73 (1967).
10. W. A. Goddard III, "Wavefunctions and Correlation Energies for Two-, Three-, and Four-Electron Atoms", *J. Chem. Phys.* 48, 1008 (1968).
11. W. A. Goddard III, "The Symmetric Group and the Spin Generalized SCF Method", *Int. J. Quantum Chem.* 3S, 593 (1969).
12. W. A. Goddard III, "New Type of Wave Function for Li, Be^+ , and B^{++} ", *Phys. Rev.* 169, 120 (1968).
13. W. A. Goddard III, "Magnetic Hyperfine Structure and Core Polarization in the Excited States of Lithium", *Phys. Rev.* 176, 106 (1968).
14. (a) W. A. Goddard III, "A New Foundation for the Use of Pseudopotentials in Metals", *Phys. Rev.* 174, 659 (1968);
 (b) L. R. Kahn and W. A. Goddard III, "A Direct Test of the Use of Pseudopotentials in Molecules", *Chem. Phys. Lett.* 2, 667 (1968).

15. W. E. Palke and W. A. Goddard III, "The Electronic Structure of LiH According to a Generalization of the Valence Bond Method", *J. Chem. Phys.* 50, 4524 (1969).
16. C. W. Wilson, Jr., and W. A. Goddard III, "Ab Initio Calculations on the $H_2 + D_2 = 2HD$ Four-Center Exchange Reactions. II. Orbitals, Contragradience, and the Reaction Surface", *J. Chem. Phys.*, to be published.
17. R. J. Blint, W. E. Palke, and W. A. Goddard III, unpublished; see Fig. 3 of ref. 14b.
18. P. M. O'Keefe, R. C. Ladner, and W. A. Goddard III, "The Orbital Description of the Symmetrical H_6 and H_4 Systems", to be published.
19. W. A. Goddard III, "Improved Quantum Theory of Many-Electron Systems. III. The GF Method", *J. Chem. Phys.* 48, 450 (1968).
20. In some cases we can take linear combinations of HF orbitals to yield lower symmetry functions, but canonical HF orbitals which diagonalize the one electrons Hartree-Fock Hamiltonian, $H^{HF} \Phi_k = \epsilon_k \Phi_k$, must be symmetry functions.
21. Note that $R_{ab}R_{ab}^T$ is an element, say R , of the sc space group; thus with (31) we obtain (30).
22. P. M. O'Keefe and W. A. Goddard III, "Lithium Energy Band Structure from Ab Initio Pseudopotentials", *Phys. Rev.* 180, 747 (1969).
23. N. F. Mott, "The Basis of the Electron Theory of Metals with Special Reference to the Transition Metals", *Proc. Roy. Soc. (London)* A62, 416 (1949); J. M. Ziman, Principles of the Theory of Solids (Cambridge University Press, 1964).
24. H. W. B. Skinner, "The Soft X-Ray Spectroscopy of the Solid State", *Rept. Progr. Phys.* 5, 257 (1938); D. E. Bedo and D. H. Tomboulian, "K-Emission Spectrum of Metallic Lithium", *Phys. Rev.* 109, 35 (1958).
25. There is not a sharp edge for Li, but rather a steep linear decrease in intensity ending in a tail of width kT . We have taken the edge as the energy extrapolated from the linear region.
26. Figure 9 is based on R. H. Parmenter, "Electronic Energy Bands in Crystals", *Phys. Rev.* 86, 552 (1952).

27. F. S. Ham, "Energy Bands of Alkali Metals. I. Calculated Bands", Phys. Rev. 128, 82 (1962).
28. R. S. Crisp and S. E. Williams, "The Soft X-Ray Emission Spectra of Sodium, Beryllium, Boron, Silicon, and Lithium", Phil. Mag. 6, 365 (1961).
29. D. A. Goodings, "Interpretation of the Soft X-Ray Emission Spectrum of Lithium Metal", Proc. Phys. Soc. 86, 75 (1965).
30. P. M. O'Keefe and W. A. Goddard III, "A New Approach to Energy Band Calculations with Results for Lithium Metal", Phys. Rev. Lett. 23, 300 (1969).
31. M. H. Cohen and L. M. Falicov, "Magnetic Breakdown in Crystals", Phys. Rev. Lett. 7, 231 (1961).
32. P. A. Penz and R. Bowers, "Strain Dependent Magnetoresistance of Potassium", Phys. Rev. 172, 991 (1968).
33. J. Babiskin and P. G. Siebenmann, "Saturating and Linear Magnetoresistive Effects in Sodium and Potassium", Phys. Kondens Materie 2, 113 (1969).
34. H. Taub, R. Schmidt, B. W. Maxfield, and R. Bowers, unpublished; private communication from H. Taub.
35. P. A. Penz, "Field-Dependent Hall Coefficient in Potassium", Phys. Rev. Lett. 20, 725 (1968).
36. R. G. Chambers and B. K. Jones, "Measurement of High-Field Hall Effect by an Inductive Method", Proc. Roy Soc. A270, 417 (1962).
37. J. J. Donaghy and A. T. Stewart, "Fermi Surface of Lithium by Positron Annihilation", Phys. Rev. 164, 391 (1967).
38. D. Shoenberg and P. J. Stiles, "The De Haas-van Alphen Effect in Alkali Metals", Proc. Roy. Soc. A281, 62 (1964).
39. C. C. Grimes and A. F. Kip, "Cyclotron Resonance in Sodium and Potassium", Phys. Rev. 132, 1991 (1963).
40. H. Mayer and M. H. El Naby, "Zum Inneren Lightelektrischen Effekt (Quantensprungabsorption) im Alkali Metall Kalium", Z. Physik. 174, 289 (1963).
41. M. H. Cohen and J. C. Phillips, "Evidence for a New Collective Resonance in a Free Electron Metal", Phys. Rev. Lett. 12, 662 (1964).

42. A. W. Overhauser, "Spin-Density Wave Antiferromagnetism in Potassium", Phys. Rev. Lett. 13, 190 (1964).
43. N. V. Smith, "Optical Constants of Sodium and Potassium from 0.5 to 4.0 eV by Split-Beam Ellipsometry", Phys. Rev. 183, 634 (1969).
44. J. C. Slater, "Cohesion in Monovalent Metals", Phys. Rev. 35, 509 (1930).
45. A. W. Overhauser, "Spin Density Waves in an Electron Gas", Phys. Rev. 128, 1437 (1962).
46. (a) J. R. Reitz and A. W. Overhauser, "Magnetoresistance of Potassium", Phys. Rev. 171, 749 (1968).
(b) A. W. Overhauser, "Exchange and Correlation Instabilities of Simple Metals", Phys. Rev. 167, 691 (1968).
47. T. M. Wilson, "A Study of the Electronic Structure of the First Row Transition-Metal Compounds", Int. J. Quantum Chem. 3S, 757 (1970).
48. J. L. Calais, "Different Bands for Different Spins. I. The Cohesive Energy of an Alkali Metal", Ark. Fysik. 28, 479 (1965).
49. R. Paunz, J. De Heer, and P. O. Lowdin, "Studies on the Alternant Molecular Orbital Method, I, II", J. Chem. Phys. 36, 2247, 2257 (1962).
50. J. De Heer, "The Method of Different Orbitals for Different Spins and Its Application to Alternant Hydrocarbons", Rev. Mod. Phys. 35, 631 (1963).
51. (a) E. G. Larson and W. R. Thorson, "Study of Electron Correlation in the H₆ Ring, Using a Novel Approximation", J. Chem. Phys. 43, 3832 (1965); "Model of Electron Correlation in Solids", J. Chem. Phys. 45, 1539 (1966).
(b) J. W. Moskowitz, "Study of the H₆ Ring in the Molecular Orbital and Alternant Molecular Orbital Approximations", J. Chem. Phys. 38, 677 (1963).

LIST OF CONTRIBUTORS

Anderson, A. G., IBM Research, Yorktown Heights	xiii
Andersen, O. K., University of Pennsylvania	178
Arbman, G., Northwestern University	33
Arlinghaus, F. J., General Motors Research Laboratories . .	400
Ashcroft, N. W., Cornell University	368
Bohn, G., University of Munich	44
Boring, A. M., Los Alamos Scientific Laboratory	495
Bross, H., University of Munich	44
Callaway, J., Louisiana State University	512
Chaney, R. C., University of Wisconsin	284
Collins, T. C., Aerospace Research Laboratories, WPAFB .	82, 124
Connolly, J. W. D., Pratt and Whitney Aircraft	3
Cooper, B. R., GE Research and Development Center	207
Dalton, N. W., IBM Research, San Jose	225
Davis, H. L., Oak Ridge National Laboratory	183
DeCicco, P. D., Massachusetts Institute of Technology . .	296
Diamond, J. B., University of Pennsylvania	347
Doniach, S., Stanford University	500
Dresselhaus, G., Lincoln Laboratory, MIT	260
Ellis, D. E., Northwestern University	271, 276
Euwema, R. N., Aerospace Research Laboratories, WPAFB .	82, 124
Faulkner, R. A., Bell Telephone Laboratories, Murray Hill .	16
Fry, J. L., Louisiana State University	512
Goddard, W. A., III, California Institute of Technology .	542
Gray, D. M., Watervliet Arsenal	144

Harris, F. E., University of Utah	517
Herman, F., IBM Research, San Jose	
Hu, S. M., IBM Components Division, Fishkill	157
Jacobs, R. L., Imperial College, London	340
Janak, J. F., IBM Research, Yorktown Heights	323
Jepsen, D. W., IBM Research, Yorktown Heights	157, 416
Johnson, K. H., Massachusetts Institute of Technology . . .	377
Kambe, K., Fritz-Haber Inst. d. Max-Planck-Ges., Berlin . .	409
Karpfen, R. J., Watervliet Arsenal	144
Koelling, D. D., Northwestern University	25, 33
Kreiger, E. L., GE Research and Development Center	207
Lafon, E. E., Oklahoma State University	284
Lee, M. J. G., University of Chicago	63
Li, E. K., Massachusetts Institute of Technology	400
Liberman, D. A., Los Alamos Scientific Laboratory	489
Lin, C. C., University of Wisconsin	284
Lipton, D., Imperial College, London	340
Marcus, P. M., IBM Research, Yorktown Heights	xv, 416
Mattheiss, L. F., Bell Telephone Laboratories, Murray Hill .	355
Meister, G., University of Munich	44
Monkhorst, H. J., University of Utah	517
Mueller, F. M., Argonne National Laboratory	305
O'Keefe, P. M., California Institute of Technology	542
Overhof, H., University of Marburg	218
Ortenburger, I. B., IBM Research, San Jose	469
Painter, G. S., Oak Ridge National Laboratory	271, 276
Pratt, G. W., Jr., Massachusetts Institute of Technology .	400
Schoen, J. M., Bell Telephone Laboratories, Allentown . . .	373
Schubö, W., University of Munich	44
Segall, B., Case Western Reserve University	200, 207
Sham, L. J., University of California, San Diego	458
Shankland, D. G., Air Force Institute of Technology, WPAFB	362
Slater, J. C., University of Florida	447

Smith, F. C., Jr., Massachusetts Institute of Technology . . .	377
Snow, E. C., Los Alamos Scientific Laboratory	495
Stöhr, H., University of Munich	44
Stukel, D. J., Aerospace Research Laboratory, WPAFB . . .	82, 124
Tong, B. Y., University of Western Ontario	476
Young, C-Y, Lincoln Laboratory, MIT	260
Wepfer, G. G., Aerospace Research Laboratory, WPAFB	124
Williams, A. R., IBM Research, Yorktown Heights	157
Wilson, A. R., Lincoln Laboratory, MIT	260
Woo, J. W. F., IBM Research, Yorktown Heights	473
Wood, J. H., Los Alamos Scientific Laboratory	59

SUBJECT INDEX

A

- Ag 207, 213, 215
 - phase shifts 214
- Al, density of states 327, 332
 - MAPW 52
 - pseudopotential for 10
- Alkalies, optical properties 260
- APW 4, 59
 - alternative method 25
 - modified (MAPW) 44
 - relativistic (RAPW) 33

C

- Cellular method 9
- Charge density in SCOPW 102ff
- Clebsch-Gordan coefficients 49, 170
- Clusters of atoms, electronic properties 377
- Comparison of APW, OPW, KKR, LCAO 176
- Compounds
 - LiF 294
 - ReO₃ 357ff
 - SiC 282
 - SO₄⁼ 379, 386
 - CuS₂ 383
- Conservation of effort, band methods 7, 41, 176
- Copper 8, 29, 52, 146, 212, 253, 255, 498
 - phase shifts 210
 - interband oscillator strength 306
- Core shift 10
- Critical points 308
- Cs 77
- Cubic harmonics 538

D

- Density of states
 - basis decomposition 313
 - critical structure 308
 - derivatives of 334ff
 - GR method 323, 349
 - polynomial expansions 308ff
 - sampling considerations 310
 - transition metals 342
 - two OPW's, exact 368
- Determinantal equation (see secular equation)
- Diagonalization
 - Hermitean matrices 16
 - QR method 18
- Diamond 280, 293
- Discrete variational method (DVM) 271, 276

E

- Electron gas, SCHF dielectric function 473
- Error functional for wave functions 273
- Evanescence Bloch waves 417
- Ewald expansion 187
- Ewald parameter 194
- Exchange
 - Cu, Ni 9
 - approximations 94, 272
- Exchange approximations,
 - comparison of for copper 495
 - Xα 447
 - Xα, compared to HF 450ff
 - Hohenberg-Kohn, for atomic systems 477
- Exchange-correlation potential 458

- Kohn-Sham-Gaspar (KSG) 459
 gradient correction 460, 464ff, 469
 fourth-order exchange corrections 470
 local effective potential for Na 481ff
 study of several forms for Cu 495
 Exchange operator, Overhauser divergence 462
 Exchange potential, tight-binding method 515
 Exchange potential, new 489
- G
- Gaunt coefficients 180, 186, 383, 412, 431
 Gaussian type orbitals 287
 Ge 162
 GI method 542
 Gilat-Rabenheimer (GR)
 method 61, 315, 323ff
 photoemission 332
 tests of 331
 Givens-Householder reduction 17
 Gradients of $E(k)$ 59
 Graphite 280
 Green's function 166
 Green's function method (GFM)
 (see KKR method)
- H
- Hartree-Fock (HF), in
 crystalline hydrogen 517
 unrestricted 563
 Hermite functions 317
 Hole recoil, effect on band structure 505
- I
- Impurity cluster 393
 Impurity problem (vacancy) 400
- Incomplete Gamma functions 190, 413
 Integration 305, 323, 340, 347
 Interband transitions, relaxed orbital corrections (ROC) 500
 Interpolation 117, 305, 323, 328, 340, 355, 362
 determinantal method 331
 Fourier series method 362
 GR scheme 311, 323, 337, 340
 $k \cdot \Pi$ 400ff
 LCAO (nonorthogonalized orbitals) 355ff
 LCOAO (orthogonalized orbitals) 355ff
 QUAD scheme 312
 roughness functional 364
 Iron 253ff
- J
- Jacobi method diagonalization 17
- K
- K 76, 262
 magnetoresistance (calculated) 560
 KKR method 4, 157, 178, 183, 200, 207, 218, 225
 reciprocal lattice representation 230
 approximate schemes for transition metals 221, 235
 numerical results 245
 non-muffin tin 158, 200
 non-muffin tin, test using Mathieu potential 205
 predictor-corrector algorithm 171
 relativistic corrections 221
 structure factors 7, 159
 183, 202, 218, 390
 tables of structure factors 195, 196

- two dimensional lattices 409
 wave function 161, 178, 200
 KKR-layer method 416, 429
 scattering matrix 428, 435
 structure constants 412
 KKRZ 6, 260
 k.p method 38
 k. π method, application to vacancy problem 400
- L
- LCAO method 4, 276, 284
 cellular method 296
 Hartree-Fock 517
 interpolation method 355
 nearest neighbor approximation 289
 Taylor series method 298
 variational improvement 290
 LEED 409, 417
 Li 56, 74, 278, 288, 557
 x-ray emission 558ff
 Logarithmic derivatives 46
- M
- Modified APW (MAPW) 44
 MAPW, convergence of (E and ψ) 52
 MOPW 5
 d orbitals 144ff
 Matrix, Hermitean, diagonalization 16
 Matrix elements 36, 67, 306, 401
 Molecular orbitals, self-consistent field approach 377ff
 Monte Carlo integration 264, 305
 Mott paradox resolved 557
 Muffin tin orbital 179
 Multicenter integrals 285
- N
- Na 75, 262, 486
 x-ray emission spectrum 559
 Ni, MAPW 52
- O
- Occupation numbers, spin orbitals in ground state 448
 Optical interband conductivity 264
 Optical properties, core effects 266
 Optical excitations of crystals 452
 OPW 4, 85, 126
 convergence 89
 isolated atom model 98
 modified (MOPW) 5, 144
 relativistic 129
 self-consistent (SCOPW) 82, 101
 symmetrized (SOPW) 97, 124
 Overcompleteness 203
 Overlap integrals 522
- P
- Partial waves 3, 380
 Pd, density of states 330
 Phase shifts, alkalis 262
 transition metals 233
 parameterization 207, 209, 232
 Photoemission 332ff
 Plane waves 3
 symmetrized combination 134
 Potential
 complex 438
 Madelung 108ff
 Mathieu 205
 nonspherical, in APW 47
 non-muffin tin 200, 276
 overlapping spheres 10
 periodic 3
 warped muffin tin 10
 Propagation matrix method 416

- Pseudopotentials 5, 10, 260
 APW, for alkalis 63, 72
 phase shift analysis 65, 234
 Pt, density of states 314
 Q
 QR method 18
- R
 Radial functions 45, 165
 Rb 76
 Relaxation of holes 500
- S
 Secular equation 4, 28, 59, 66, 87, 168, 229, 239, 273, 285, 382, 390
 Choleski decomposition 88
 RAPW 34
 Self consistent field method 447
 Self consistent field scattered wave method (SCF-SW) 377
 Self consistent OPW (SCOPW) 82
 Self-consistency, tight-binding method 512
 Si 162
 Single-particle states in solids, ground state as interband transitions 500
 Slater type orbital (STO) 277
 Stars, FCC 40
 linearly independent members 132ff
 relativistic 132ff
 Structure constants (see KKR)
 Susceptibility 340, 347
 free electron 350
 wave vector dependence 351
 SRAPW 37
 SROPW 124
 Surface matching equations for wave function 424
 Symmetrization procedures 275
- T
 T matrix 384
 Transfer matrix Q 411, 427
 Transformed symmetrized APW's (TSAPW) 403
 Transport coefficients, integration scheme 373
 Transition metals 225
- V
 Vacancy states 400
 Variational principle 47, 51, 158
- W
 Wannier functions 401, 406
 Watson sphere 380
- X
 X α method 448ff
 vs. Hartree-Fock 450