The Electron Theory of Simple Metals and Alloys*

Morrel H. Cohen

Institute for the Study of Metals and Department of Physics
University of Chicago
Chicago 37, Illinois

A brief review of the successes and of the difficulties of the Bloch model of a metal is given. The Landau quasi-particle theory and the pseudopotential theory are briefly exposed. These two notions are made the basis for a revised version of the Bloch model which avoids the conceptual difficulties of the original model while retaining, and indeed explaining, all its appealing features.

Well known empirical generalizations concerning the relative importance of conduction-electron density, composition, and structure for the electronic properties are presented and their implications examined anew. A theory of the form required by the empirical generalizations is developed. Recent approximate calculations are discussed and a variety of conclusions drawn from these relating to the possibility of explanation of the empirical generalizations by the theory.

A talk presented to the Special Session on Concentrated Solid Solutions of the Cleveland meeting of the ASM on October 21, 1963.

*Facsimile Price $4.60
Microfilm Price $1.46

Available from the Office of Technical Services
Department of Commerce
Washington 25, D.C.
DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.
DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.
I. INTRODUCTION

In this paper, we are concerned with two separate but related aspects of the electron theory of metals and alloys.

The first is the problem of the proper conceptual framework for describing the motion of individual electrons in metals and alloys together with the problem of devising a reasonable approximation scheme for calculations relating thereto. The Bloch model until recently has provided the basis for all discussion of the motion of electrons in metals and alloys. In section II, we give a brief historical review of the status of the Bloch model, emphasizing particularly the dilemma posed by the contrast between its great empirical success and the apparent weakness of its underpinnings. The resolution of this paradox has occurred only recently with the advent of the Landau quasi-particle theory\(^1-4\) and the pseudopotential of Phillips and Kleinman.\(^5-13\) In section III we briefly describe the Landau quasi-particle theory, and in section IV do the same for the pseudopotential. Finally, in section V, on the basis of these notions we develop an explicit, simple form for the Bloch model which includes electron correlation as well as crystal field effects.

The second aspect of the electron theory of metals relates to the understanding of certain broad empirical generalizations stated in section VI. These are generalizations about the relative importance of conduction-electron density \(n\), atomic composition \(c\), and structure \(S\) for the properties of
metals and alloys. The theory is put into a form which exhibits explicitly the separate dependence of properties on \( n, c, \) and \( S \) in section VII. In section VIII, we finish by reviewing the results of recent approximate calculations which permit one to infer that the relative importance of \( n, c, \) and \( S \) demanded by the empirical generalizations can indeed be derived from the theory of section VII.

Before proceeding to the main body of the paper, we must state that all of the advances described therein are valid primarily for what we term the simple metals. These are best defined by elimination. We shall not consider metals of the transition, rare-earth, or actinide series. Our considerations have borderline validity for the noble metals and for the metals just to the left of the transition series in the periodic table. All others, and there are still some left, we term simple metals. It will be seen that the pseudo-potential provides a quantitative basis for this classification.

II. THE BLOCH MODEL, PARAGON OR PARADOX?

In 1928, Bloch\textsuperscript{14} introduced his famous model as a basis for resolving the difficulties with the electron mean-free path inherent in the simpler Pauli\textsuperscript{15}-Sommerfeld\textsuperscript{16} free-electron model of a metal. We may characterize the Bloch model as follows:

1. The conduction electrons of a metal are identified with the valence electrons of the atoms constituting the metal.

2. The conduction electrons are supposed to move
independently of one another within the metal.

(3) The independent motion of the electrons is generated by a one-particle Hamiltonian, \( H \),

\[
H = \frac{p^2}{2m} + V(r)
\]  

(2.1)

containing the kinetic energy \( \frac{p^2}{2m} \) and a potential \( V(r) \) giving the influence of the environment on each electron.

(4) When the model is applied to crystalline metals, \( V(r) \) is supposed to have the periodicity and symmetry of the crystal. From this alone, most of the consequences of the Bloch model flow.

(5) \( V(r) \) is supposed to be vaguely atomic in character, but it is otherwise unknown.

Used as a conceptual framework within which to fit experimental data, the Bloch model has been enormously successful. We may point to two areas of the physics of metals in which the model has been used to great advantage. The first, the Brillouin-zone theory of metals and alloys, had its inception in the early thirties, grew rapidly to maturity, and continues to grow today, employing the Bloch model most profitably. It is concerned with understanding lattice constant relationships, limits of solid solubility, etc., in terms of the relation between the zone structure and the Fermi surface and has not changed much from its description in the classic book of Mott and Jones. The second,
Fermiology, the study and determination of the Fermi surfaces of metals, is of much more recent vintage. It had its inception in 1957 with Pippard's determination of the Fermi surface of copper and its coming-out party in 1960 at the Fermi Surface Conference in Cooperstown, New York.

One can fairly say that the Bloch Model is a paragon of models with its simplicity and striking successes, such as the two mentioned above. Nevertheless, two serious, fundamental objections to the Bloch model have been repeatedly raised against it over the years, in effect denying the model a proper theoretical sanction.

The first and less severe objection is that in fitting the model to the experimental data on simple metals and alloys, the required matrix elements of the crystal potential turn out to be small compared to the free-electron Fermi energy, $E_F$, which is a measure of the average kinetic energy of the conduction electrons. This result apparently contradicts the notion that the potential in the Bloch model is atomic in character. One would have expected the potential and kinetic energy to be of comparable magnitudes.

The second and more fundamental objection is that the electrons are treated as independent despite the existence of the Coulomb interaction among them, which surely must be comparable in magnitude to the kinetic energy of the electrons.

One is thus faced with a paradox: the Bloch model
appears to work quantitatively, but it also appears to demand for its very existance weak electron-atom interactions and negligible electron-electron interactions. The resolution of this paradox has been effected by a twofold reinterpretation of the Bloch model. The first element of reinterpretation is provided by the Landau quasi-particle theory\textsuperscript{1-4} introduced in 1956,\textsuperscript{1} which eliminates the difficulty of the strong electron-electron interaction and accordingly is discussed in section III. The second element of reinterpretation is provided by the pseudopotential theory\textsuperscript{5-13} introduced by Phillips and Kleinman in 1959,\textsuperscript{5} which eliminates the difficulty of the strong electron-atom interaction and accordingly is discussed in section IV.

III. THE LANDAU QUASI-PARTICLE THEORY

The difficulty posed by the strong electron-electron interaction is overcome by recognizing that the Bloch model does not in fact describe the motion of the individual electrons, which are of course not independent. Rather, the Bloch model describes the motion of more complex entities, called Landau quasi-particles, which stand in precise one-to-one correspondence with the independent electrons in the original interpretation of the Bloch model.

Landau pointed out that one should regard the electrons in a metal as forming a Fermi liquid rather than a Fermi gas because of their strong interactions and high density. He then set up a general theory of a Fermi liquid
in which the basic entities are the quasi-particles rather than the actual particles comprising the system. The quasi-particle corresponding to an electron in a metal may be pictured as in Fig. 1 as having three parts: first, an electron at its center moving uniformly through the motion; second, a surrounding region of electron deficiency, or correlation hole, of order an average electron separation in diameter and containing a net deficiency of one electron; and third, a region of backflow associated with the motion of the correlation hole, i.e., electrons pushed by the Coulomb repulsion away from the central electron flow around it much as water around a moving obstacle. The theory of a Fermi liquid was set up by Landau in terms of these quasi-particles in such a way that for a metal reinterpretation of the Bloch model in terms of quasi-particles requires only relatively minor modifications of the form of the model. The important thing to remember is that in the past when we talked about independent electrons forming an electron gas in a metal, we were really talking about the nearly independent quasi-particles which are the elementary excitations of the Fermi liquid actually formed by the electrons in the metal.

The quasi-particle structure is not infinitely stable except for those quasi-particles corresponding to electrons at the Fermi energy. However, as long as we deal with energies not too far removed from the Fermi energy, we can suppose the lifetime of the quasi-particles to be
infinite. In that case it becomes possible to write down a single-particle-like Hamiltonian which generates the motion of the individual quasi-particles. This quasi-particle Hamiltonian replaces that of Eq. (1) in the Bloch model; we set it up in section V after discussion of the pseudopotential in section IV.

IV. THE PSEUDOPOTENTIAL

Atomic electrons may be separated into two groups, the core and the valence electrons. The core electrons are the inner, tightly bound electrons of the atom. The corresponding set of one-electron wave functions, \( \psi_c \), are essentially unaffected by the state of aggregation of the atoms. Consequently, the \( \psi_c \) are often known already from studies of isolated atoms or ions.

The valence electrons, on the other hand, are the outer, relatively loosely bound electrons. They are highly sensitive to the state of aggregation of the atoms; indeed they are responsible for the interatomic forces leading to aggregation. The set of one-electron, valence-electron wave functions, \( \psi_v \), must thus be determined anew for each different state of atomic aggregation from some suitable form of the Schrödinger equation, e.g.,

\[
\mathcal{H} \psi_v = E_v \psi_v , \tag{4.1}
\]

where \( E_v \) represents the set of valence electron energies. However, they must be determined from (4.1) in such a way
that the Pauli exclusion principle is satisfied, the mathematical expression of which is the orthogonality requirement

\[ (\psi_v, \psi_c) = 0. \]  

(4.2)

It is clear that in the usual application of the Bloch model to metals the core electrons are completely ignored, although this is certainly not the case for attempts to calculate the band structure \textit{a priori}. We are thus faced with a dilemma: it is most convenient to be able to ignore the existence of core electrons in developing and applying a model like that of Bloch, but we must still satisfy the exclusion principle through observation of Eq. (4.2).

The pseudopotential introduced by Phillips and Kleinman\textsuperscript{5} provides us with the escape from this dilemma. We can ignore to some extent the existence of the core functions if we deal explicitly only with a \textit{pseudo} wave function \( \phi_v \), which stands in one-to-one correspondence with the \textit{actual} valence-electron wave function. The pseudo wave function is not required to satisfy the orthogonality condition (4.2), which is an expression of the statistical exclusion of the valence electrons from the region of space occupied by the core electrons. Instead, an electron described by the pseudo wave function \( \phi_v \) is kept out of the core dynamically by adding a repulsive potential\textsuperscript{5,7,8,10} \( V_R \) to the usual Hamiltonian to get the Schrödinger equation satisfied by \( \phi_v \),
\[ H \phi_V = (\mathcal{H} + V_R) \phi_V = E_V \phi_V \quad (4.3) \]

where the valence-electron energy \( E_V \) is the same as in (4.1).

There is no unique way to introduce \( \phi_V \) and \( V_R \) by relaxing condition (4.2). We may take advantage of this arbitrariness by requiring that \( V_R \) cancel the core contribution to the potential acting on a valence electron, \( V_c \), to as great a degree as possible. This cancellation is excellent in the core region as is illustrated in Fig. 2 for the Si\(^{+++}\) core potential.

The core potential \( V_c \) and repulsive potential \( V_R \) may be combined into a single pseudopotential,

\[ V_I = V_c + V_R = \sum_l v_l^{\phi}(p,\mathbf{r}) \quad (4.4) \]

which is a superposition of individual ion-core pseudopotentials \( v_l^{\phi}(p,\mathbf{r}) \) which are spherically symmetric but nonlocal, i.e., momentum dependent, in character. Thus far we have considered only the influence of the cores on the valence electrons. In a metal the valence, i.e., conduction, electrons screen the ion-core pseudopotential so that the total potential in the pseudo wave equation \( (4.3) \), may be considered quite weak for many applications.

Thus, recognizing the one-to-one correspondence between the pseudo wave function and the real wave function, we see that (1) the original Bloch model in fact referred to the pseudo wave function and not the actual wave function.
(2) The crystal potential of the Bloch model is in fact the pseudopotential of Kleinman and Phillips. (3) The weakness of the potential required to fit the Bloch model to experiment can be understood quite simply as arising from the cancellation of the strong potential inside the core by the repulsive potential and from the screening of the resultant ion-core pseudopotential outside the core by the conduction electrons themselves. Finally (4), the classification of metals into simple and otherwise is based on the completeness of this cancellation: Simple metals have weak pseudopotentials; other metals do not.

V. THE BLOCH MODEL, REVISED

We have obtained the following simple Hamiltonian for the Landau quasi-particles in a metal,

\[ H = \frac{p^2}{2m} + V_I + V_H - V_{EX} \]  

(5.1)

The derivation of (5.1) involves current many-body theory in all of its technical complexity; discussion of the derivation and the attendant approximations would thus be out of place here. Suffice it to say (1) that the Green's function formalism of, e.g., Martin and Schwinger\textsuperscript{22} and some of the remarks of Luttinger and Nozières\textsuperscript{3,4} concerning the relations between the Green's function and the quasi-particles form a convenient starting point for obtaining (5.1) in a straightforward manner, (2) that the most serious deficiency of (5.1) arises from inadequacy in the
treatment of short-range correlations between pairs of electrons, and (3) that the quasi-particles are supposed to brave infinite lifetimes.

In (5.1), the contributions to the quasi-particle Hamiltonian are first the kinetic energy $\frac{1}{2} \mathbf{p}^2$; next the total ion-core pseudopotential $V_I$ discussed in IV; third the Hartree field, or average Coulomb potential $V_H$ produced by all the conduction electrons; and finally the dynamically screened exchange potential $V_{EX}$ of all of the conduction electrons. Thus the quasi-particle Hamiltonian $H$ is much like the usual Hartree-Fock Hamiltonian in that a Coulomb and exchange term are present which must both be determined selfconsistently. $H$ differs from a Hartree-Fock Hamiltonian, however, in two essential respects: first because it contains the pseudopotential and thus relates to the pseudo wave function, and secondly because the correlation hole around each quasi-particle screens the electrostatic potential in $V_{EX}$. The result of the screening process is to divide the matrix element of the Coulomb potential for an exchange collision of given energy and momentum transfer by the dynamical dielectric function of the conduction electron system for the corresponding frequency and wave number. The dynamical dielectric function itself must be determined selfconsistently.

We now consider the application of (5.1) explicitly to simple metals, where $V_I$ may be treated as weak. We consider $V_I$ to be of first-order and calculate $V_H$ and $V_{EX}$.
selfconsistently to first order in \( V_I \):

\[
V_I = V_I^{(1)}
\]

\[
V_H = V_H^{(0)} + V_H^{(1)} + \ldots
\]

\[
V_{EX} = V_{EX}^{(0)} + V_{EX}^{(1)} + \ldots
\]

(5.2)

The superscripts in (5.2) indicate the order of each term. Because of electrical neutrality, \( V_H^{(0)} \) vanishes; further, \( V_{EX}^{(0)} \) depends only on momentum. Together, the zero order term in (5.1) give the energy \( E_0(p) \) of a quasi-particle in a system of electrons neutralized by a positive, uniform, background charge,

\[
E_0(p) = \frac{p^2}{2m} - V_{EX}^{(0)}.
\]

(5.3)

\( E_0(p) \) replaces the kinetic energy \( \frac{p^2}{2m} \) in the original Bloch model.

The first-order terms may be collected together into \( v^{(1)} \),

\[
v^{(1)} = V_I + V_H^{(1)} - V_{EX}^{(1)},
\]

(5.4)

which replaces the crystal potential \( V \) in the original Bloch model. \( V_H^{(1)} \) arises from the first-order change of the electron density introduced by \( V_I \) and acts so as to screen the potential of each ion-core completely at large distances. \( V_{EX}^{(1)} \) arises both from first-order changes in the momentum distribution of the electrons and from first-order changes
in the dielectric function; it modifies the details of the screening process.

We have thus arrived at a revised form for the Bloch model,

\[ H = E_0(p) + V^{(1)}, \quad (5.5) \]

suitable both for the approximate calculation of band structures and Fermi surfaces of simple metals from first principles and for the interpretation of experimental data. In the latter case, \( E_0(p) \) can be regarded as the exact quasi-particle energy for the electron gas (apart from lifetime and retardation effects) and \( V^{(1)} \) replaced by \( V \), which stands for the remainder of the exact quasi-particle Hamiltonian obtained by ignoring lifetime effects. For calculations, it is best to retain the approximations described above.

Present theoretical estimates\(^{23}\) and interpretation of experimental evidence\(^{24}\) indicates that \( E_0(p) \) differs by order 10-25% from \( \Phi_{\lambda m} \) for the simple metals. The approximations leading to \( V^{(1)} \) give it the form of a sum of spherically symmetric, momentum-dependent, single ion-core potentials

\[ V^{(1)} = \sum_{\lambda} v_{\lambda}(p,r). \quad (5.6) \]

Recent approximate calculations of Harrison\(^{11,25}\) make it clear that the \( v_{\lambda}(p,r) \) may usually be considered individually weak so that for most processes
\[ \langle k | V^{(1)} | k' \rangle \ll E_F \] (5.7)

where \( \langle k | V^{(1)} | k' \rangle \) is a matrix element of \( V^{(1)} \) between the plane-wave states \(|k\rangle\) and \(|k'\rangle\).

We have thus arrived at a simple, useful form for the Bloch model in which the paradoxes raised by the original model are resolved and the consonance of theory and experiment made manifest. With this model, one can understand the electronic structures of simple metals and alloys. The theoretical problems remaining are to make sufficiently accurate calculations of \( E_0(p) \) and of \( V^{(1)}(p, \mathbf{p}) \), but they are within our grasp.

VI. THE PROPERTIES OF METALS AND ALLOYS; SOLID OR LIQUID: SOME EMPIRICAL GENERALIZATIONS FOR SIMPLE METALS

In the previous sections of this paper we have focused our attention on theoretical understanding of only one empirical generalization concerning metallic properties, that these seem paradoxically to imply the validity of an independent-particle, weak-potential model for electrons in metals. In the present section we wish to call attention to other empirical generalizations concerning metallic properties, all very well known but not perhaps viewed with proper perspective.
A. CONSIDERABLE STRUCTURAL INSENSITIVITY

The properties of simple metals usually show a remarkable structural insensitivity, both cohesive or collective properties and "one-electron" properties.

Consider the cohesive energy, for example. Allotropic transformations, order-disorder transitions, and melting occur, say, at some transformation temperature \( T_t \) with an energy of transformation \( E_t \). The ratio of the boiling temperature, \( T_b \), to \( T_t \) or the cohesive energy \( E_{COH} \) to \( E_t \) is ordinarily much greater for simple metals for all three categories of transformations than is the case for other kinds of materials, e.g., molecular crystals. The allotropic transformation from the bcc structure to the fcc occurring in Na at 35°K furnishes an extreme example. The ratio \( E_t/E_{COH} \) is only of order \( 10^{-3} \).

To illustrate the structural insensitivity of one-electron properties one might consider the resistivity change on melting. The resistivity is ordinarily treated theoretically in such a manner as to imply that a periodic structure is necessary for high conductivity, yet the resistivity only increases by a factor of two on melting for many simple metals. Still more striking are the Knight-shift measurements of Knight, Berger, and Heine\(^25\) who report that the Knight shifts of a variety of metals are essentially unchanged upon melting.
B. RELATIVE COMPOSITION INSENSITIVITY

The properties of simple metals usually show a relative insensitivity to composition.

In support and clarification of this second generalization we can break it down further in two statements:

(1) The properties of pure metals depend primarily on the column of the periodic table rather than on the row, i.e., on valence rather than on atomic number.

(2) Within a row, alloys having the same average valence have similar properties, irrespective of constitutive details, and, as implied to some extent by (1), the average valence remains primary even when rows are mixed.

C. STRONG DEPENDENCE ON CONDUCTION-ELECTRON DENSITY

It is well known that when neighboring B-subgroup elements form binary alloys with the noble metals, no special regularities emerge in the dependence of an electronic property $P$ on atomic concentration $C$, as is illustrated in Fig. 3. It is equally well known that striking regularities emerge when the dependence of the same $P$ on electron-to-atom ratio, or $e/a$, is considered in the same alloys as is illustrated in Fig. 4. We are thus led to our final empirical generalization:

The properties of simple metals depend strongly on $e/a$, or more accurately, conduction-electron density $n$. 
D. IMPLICATIONS FOR THE THEORY

Generalizations A, B, and C taken together suggest that an electronic property $P$ can be decomposed into three parts

$$P = P_e(n) + P_c(n,c) + P_S(n,c,S) .$$  \hspace{1cm} (6.1)

The first part $P_e$ depends only on electron density $n$,

$$P_e = P_e(n) \hspace{1cm} (6.1a)$$

and is large. The second part $P_c$ depends in detail on composition, symbolized by $c$, as well,

$$P_c = P_c(n,c) \hspace{1cm} (6.1b)$$

and is smaller than $P_e$. The third and last part $P_S$ depends in addition on structure, symbolized by $S$,

$$P_S = P_S(n,c,S) \hspace{1cm} (6.1c)$$

and is very much smaller than $P_e$ or $P_c$. The theoretical task facing us is to effect the decomposition (6.1) formally and then to demonstrate through approximate calculations that the relative magnitudes of $P_e$, $P_c$, and $P_S$ demanded by experience follow also from the theory.

Before doing so in the next two sections, we wish to illustrate both the limitations and the validity of such a decomposition as (6.1) with three simple examples:

1. The Pauli spin paramagnetism $\chi$ evidently could be a property for which the decomposition (6.1) holds. It
exists for an electron gas, and, if one carried out a perturbation treatment for $X$ based on the weakness of the pseudo-potential just as for the quasi-particle Hamiltonian $H_{\text{in}}$, one could presumably find,

$$X = X_e + X_c + X_S$$

(6.2)

$$X_e > X_c >> X_S$$

(6.2a)

(2) The Knight Shift $K$ is closely related to the Pauli spin paramagnetism in that it is a measure of the magnetic field produced at a nucleus by the net conduction-electron spin induced by an external magnetic field through the Pauli paramagnetism. Unlike $X$, which is essentially a collective property of the conduction electrons, $K$ is essentially an atomic property. Were we to carry out a corresponding perturbation calculation based on the weakness of the pseudo-potential, we should find

$$K = K_c + K_S$$

(6.3)

$$K_e = 0, \quad K_c >> K_S,$$

(6.3a)

as is demanded by the results of reference 26.

(3) The mechanical strength, or, more precisely, a yield stress or a Peierls-Nabarro stress, depends in the most detailed way on the structure and primarily on the structure. Most deformation phenomena owe their existence to a structural defect, the dislocation, and are most sensitive to
the presence and arrangement of other structural imperfections. For such properties, no decomposition of the type 6 is possible. When one considers the ideal mechanical strength of a perfect crystal on the other hand, one might expect theory to yield

\[ I = I_e + I_c \neq I_s \]  

(6.4)

\[ I_e \sim I_c \sim I_s \]  

(6.4a)

so that the decomposition (6.1) though possible is not useful.

To summarize, we can classify the electronic properties of metals into three categories:

(i) collective properties for which a decomposition like (6.2) exists.

(ii) atomic properties for which a decomposition like (6.3) exists.

(iii) structure-sensitive properties for which either no decomposition is possible or, if possible, is not useful.

VII. A THEORY OF THE FORM REQUIRED BY THE EMPIRICAL GENERALIZATIONS

A. SYSTEM

The system we shall consider is comprised of N atoms of arbitrary composition. We shall make the Born-Oppenheimer approximation in the sense that we separate out the nuclear motion to be treated at a stage of the theory later than is included in our present studies. The N atoms are thus to be
regarded as fixed in an arbitrary, but locally uniform, arrangement of positions within a volume $\Omega$. We suppose the $l$th atom to contribute $z_l$ conduction electrons, the total number of conduction electrons in the system being $zN$ and the average valence, or $e/a$, being $z$. A pseudo-potential form of the theory will be used, the core electrons not being treated explicitly. The $l$th ion core thus interacts with the conduction electrons through a pseudopotential $V_l^{-1}(p,r)$. 

B. ILLUSTRATION: SINGLE PARTICLE ENERGIES IN THE BLOCH MODEL.

To illustrate simply that a decomposition of the type (6.1) is theoretically possible, we now work out the eigenvalues of the Bloch model, Eq.(5.5), by treating the crystal potential $V(1)$ as a perturbation on the "free" quasi-particle energy $E_0(p)$. We shall carry the expansion to second order in $V(1)$, but, as a further simplification, ignore the corresponding second-order contribution to the crystal potential. That is, we shall not carry out a fully selfconsistent expansion.

We are required to solve

$$ H \psi = E \psi, \quad H = E_0(p) + V, \quad V = \sum_l V_l $$

(7.1)

with $\sum_l V_l$ treated as a weak perturbation. The unperturbed solutions are

$$ \psi_k^0 = \mathcal{N}^{-\frac{1}{2}} e^{i k \cdot r}, \quad E_0(p) \psi_k^0 = \epsilon_k \psi_k^0, $$

(7.2)
which we use as a basis set for the expansion. The matrix elements of $v_L$ in this basis set are

$$\langle k | v_L | k' \rangle = \langle k | v_{L0} | k' \rangle e^{i(k'-k) \cdot R_L}, \quad (7.3)$$

where $\langle k | v_{L0} | k' \rangle$ is the matrix element for the $l$th ion core at the origin and $R_L$ the actual position of the $l$th ion core. To second order, we have for the eigenvalue in (7.1)

$$E_k = E_k + \langle k | v_k \rangle + \sum_{k'} \frac{\langle k | v_{k'0} | k' \rangle \langle k' | v_k | k' \rangle}{E_k - E_{k'}} \quad (7.4)$$

This may be readily rearranged into

$$E_k = E_k + \sum_l \left[ \frac{\langle k | v_{L0} | k \rangle + \sum' \frac{\langle k | v_{L0} | k' \rangle \langle k' | v_k | k' \rangle}{E_k - E_{k'}}}{E_k - E_{k'}} \right]^2$$

$$+ \sum_{l \neq k} \left[ \sum_{k'} \langle k | v_{L0} | k+q \rangle \langle k+q | v_{L0} | k' \rangle e^{iq \cdot (R_k - R_{k'})} \right] \quad (7.5)$$

by using (7.3).

The first term in (7.5) is $E_k$, the "free" quasi-particle energy, and is thus of collective character, giving rise only to contributions to electronic properties depending on conduction-electron density only. The second term in $E_k$ depends explicitly on the kinds of atoms present but only on the kinds, so that it is a composition-dependent term. Closer examination reveals, however, that a considerable part of this term depends only
on the average ion-core charge and should be grouped with the first term as collective in character. Finally, the remaining term depends explicitly on separations between pairs of ion cores and hence is of the structure-dependent type. Again, a note of caution must be struck. The \( q = 0 \) term is omitted from the sum over intermediate states in this last term, leading to a value of the \( \mathbf{L}^2 \) contribution which becomes constant and does not vanish as the separation \( |\mathbf{P}_i - \mathbf{P}_f| \) becomes large. This would lead to a structure independent part which must be subtracted out and apportioned appropriately between the collective term and the composition dependent term. In practice, this is easy to do.

Thus terms of the form required by the empirical generalizations emerge and further will have the required size relationships if, as is the case for simple metals, the pseudopotential is weak so that perturbation theory converges rapidly. Actually the long-wavelength components of the pseudopotential are not weak, so that a more general method of proof must be used. This objection affects primarily the electron-density and composition-dependent terms. The structure dependent terms remain small relative to the others because of the relative weakness of those parts of the pseudopotential which contribute to the \( P_g \) terms.

C. EXACT ANALYSIS: FREE ENERGY

We turn now to an exact analysis of the free energy of our system both to show less trivially that a property of
the system can be decomposed as in (6.1) and to furnish a prototype for similar theoretical decompositions of other properties.

The total Hamiltonian $\mathcal{H}$ of the system is

$$\mathcal{H} = K + V_{ee} + \sum_{\ell} V_{\ell} + V_{II} \tag{7.6}$$

In (7.6), $K$ is the total kinetic energy of the conduction electrons, $V_{ee}$ is the total Coulomb interaction among all the electrons,

$$V_{\ell} = \sum_{j} V_{\ell}^{1}(p_{j}, p_{j}) \tag{7.7}$$

is the total interaction of all conduction electrons with the $\ell^{th}$ ion core via the ion-core pseudopotential $V_{\ell}^{1}(p, r)$, and $V_{II}$ is the total ion-core interaction energy. The latter is a sum of pairwise central interactions, Coulombic at nuclear separations such that the cores do not overlap, and approximately representable by a Born-Mayer repulsion otherwise. The van der Waals interaction between the ion cores may be of some small importance for the heavier simple metals.

We now propose to treat the ion core-conduction electron interaction, $\sum V_{\ell}$, as a perturbation of arbitrary strength on

$$\mathcal{H}_0 = K + V_{ee}, \tag{7.8}$$

the electron-gas Hamiltonian. $V_{II}$ is, of course, a constant.
It is of later importance that we do not neutralize $V_{ee}$ by adding to the unperturbed system a uniform positive background. Instead, we shall just keep track of the structure and composition independent divergence in $V_{ee}$, which will later be cancelled by corresponding divergences in $V$ and $V_{II}$.

The free energy is defined as

$$F = -kT \ln \text{Tr} \, e^{-\beta \mathcal{H}}$$

which may be rewritten as

$$F = V_{II} - kT \ln \text{tr} \, e^{-\beta (\mathcal{H}_0 + \sum \mathcal{U})},$$

(7.10)

where $k_B$ is Boltzmann's constant, $T$ the absolute temperature, and $\beta = (kT)^{-1}$. The free energy of the unperturbed system is, correspondingly,

$$F_0 = -kT \ln \text{tr} \, e^{-\beta \mathcal{H}_0}$$

We now define the average $\langle \mathcal{O} \rangle$ of any many electron operator $\mathcal{O}$ as

$$\langle \mathcal{O} \rangle = \frac{\text{tr} (e^{-\beta \mathcal{H}_0} \, \mathcal{O})}{\text{tr} \, e^{-\beta \mathcal{H}_0}},$$

(7.11)

This is clearly an average over the canonical ensemble for the unperturbed system. Going back to (7.10), we can put the second term in $F$ in the form of an average like (7.11) by disentangling $e^{-\beta \mathcal{H}_0}$ from $e^{-\beta}$:

$$e^{-\beta (\mathcal{H}_0 + \sum \mathcal{U})} = e^{-\beta \mathcal{H}_0} \, e_0^{- \left[ \int_0^\beta d\lambda \sum \mathcal{U}_2 (\lambda) \right]}.$$

(7.12a)
In (7.12a) and (7.13) the symbol $e_0$ stands for the ordered exponential function in which factors containing larger values of $\lambda$ stand to the left of factors containing smaller values.

It is now convenient to introduce the operator

$$g_i = e_0^{-\int_0^\beta d\lambda \tilde{V}_i(\lambda)}.$$  

(7.14)

The average in (7.13) can then be written as the average of the ordered product of the

$$\langle e_0^{-\int_0^\beta d\lambda \tilde{V}_i(\lambda)} \rangle = \langle \prod_{i=0}^T g_i \rangle.$$  

(7.15)

With the aid of a new theorem containing a generalization of the notion of Thiele semi-invariants, 27 it is possible to effect a convenient rearrangement of (7.15). The theorem has been proved by M. Coopersmith and myself; the proof and a detailed discussion will be given elsewhere. The theorem is similar to one already given by Kubo 28 but more convenient for our present purposes. The rearranged form of (7.15) is

$$\langle \prod_{i=0}^T g_i \rangle = \left[ \prod_{l=0}^T G_l \right] \left[ \prod_{l<m} G_{lm} \right] \left[ \prod_{l<m<n} G_{lmn} \right], \quad \ldots.$$  

(7.16)
with

\[ G_l = \langle \varepsilon_l \rangle \]  
(7.17a)

\[ G_{lm} = \frac{\langle \varepsilon_l \varepsilon_m \rangle_0}{\langle \varepsilon_l \rangle \langle \varepsilon_m \rangle} \]  
(7.17b)

\[ G_{lmn} = \frac{\langle \varepsilon_l \varepsilon_m \varepsilon_n \rangle_0 (\langle \varepsilon_l \rangle \langle \varepsilon_m \rangle \langle \varepsilon_n \rangle)^2}{\pi S \langle \varepsilon_l \rangle \langle \varepsilon_m \rangle_0 \langle \varepsilon_n \rangle} \]  
(7.17c)

In (7.17b) and (7.17c), the \( \langle \rangle_0 \) stands for the average of an ordered product, and the product in the denominator of (7.17c) is symmetrized with regard to \( l, m, \) and \( n \). It is possible to write the general term of (7.16) explicitly, but we shall not do so here.

We now introduce the logarithms of the \( G \), etc.,

\[ W_l = -kT \ln G_l \]  
(7.18a)

\[ W_{lm} = -kT \ln G_{lm} \]  
(7.18b)

\[ W_{lmn} = -kT \ln G_{lmn} \]  
(7.18c)

This permits us to write the free energy in the form

\[ F = F_0 + \sum_l W_l + V_{II} + \sum_{l<m} W_{lm} \]

\[ + \sum_{l<m<n} W_{lmn} + \ldots, \]

(7.19)

taking cognizance of (7.13) - (7.18). It can be shown that

\[ G_{lm} \xrightarrow{\text{as } |R_l - R_m| \to \infty} 1. \]  
(7.20a)
Similarly any of the higher order G's goes to unity if any one of the ion-core positions involved recedes to \( \infty \). From (7.20a) it follows that

\[
W_{lm} \xrightarrow{|R_l - R_m| \to \infty} 0 \quad (7.20b)
\]

and similarly for the higher W's. Thus, the free energy is beginning to appear in the desired form, containing position independent terms like \( W_l \), position dependent terms like \( W_{lm} \) which have the character of interatomic interactions because, in part, of (7.20).

The divergences left in by refusing to use neutralized Coulomb potentials cancel out again in the first four terms of (7.19). Let \( \mathcal{F}' \) be the free energy of the electron gas neutralized by a uniform positive background. Divide \( W_l \) into a part \( \mathcal{W}_l \) which depends only the charge \( Z_l \) of the \( l \)th ion core and a part, \( U_l \) which depends explicitly on atomic number,

\[
W_l = \mathcal{W}_l + U_l \quad (7.21)
\]

The pairwise interactions in \( V_{II} \) may be combined with those in the fourth term of (7.19) to give

\[
V_{II} + \sum_{l < m} W_{lm} = \sum_{l < m} U_{lm} \quad (7.22a)
\]

\[
U_{lm} \xrightarrow{|R_l - R_m| \to \infty} 0 \quad (7.22b)
\]
These definitions permit us to put \( F \) in precisely the form (6.1)

\[
F = F_e + F_c + F_S
\]

(7.23)

\[
F_e = \mathcal{G}_o + \sum_l W_l
\]

(7.23a)

\[
F_c = \sum_l U_l
\]

(7.23b)

\[
F_S = \sum_{l<m} U_{lm} + \sum_{l<m<n} W_{lmn} + \ldots
\]

(7.23c)

The electron-density dependent part of \( F, F_e \), arises from the free energy of the neutralized electron gas \( \mathcal{G}_o \) plus that part of the total free energy of interaction of each ion-core with the screening charge associated particularly with it, \( \sum_l W_l \), which depends only on electron density, \( \sum_l W_l \).

The composition-dependent term \( F_c \) arises from the remainder of the interaction of the electrons with a single ion-core at a time, \( U_l \). Finally, the structure dependent part \( F_S \) arises from a sum of pair-wise central interactions between ion cores \( U_{lm} \) plus triplet interactions, plus \( \ldots \).

One can say nothing about the relative magnitudes of the three parts or even whether the series of terms in 2-, 3-, \ldots \( n \)-body... interactions for \( F_S \) converges on the basis of the present analysis. The latter is a purely formal demonstration of the possibility of arranging the theory of the electronic properties of metals in such a way as to exhibit explicitly and separately the three dependences required by the empirical generalizations of §VII.
A number of studies of the ground state energy of the neutralized electron gas have been carried out. Of the remaining terms in (7.23), recent approximate results have been obtained for \( \sum \nu \xi \), \( \sum \phi \), and \( \sum \mu \phi m \) for the quantities equivalent to these. The present author obtained an expansion like (7.23) by carrying out a perturbation expansion of the SCF dielectric formulation. A momentum-independent pseudopotential was used to second-order in perturbation theory, and no exchange or correlation effects were included in \( F_c \) and \( F_s \). Bellemans and de Heenner carried out similar calculations but based them on the dense-electron gas approximation in many-body perturbation theory and obtained results which represent an approximation to those of reference 33. W. A. Harrison has carried out for Zn the most extensive numerical calculation to date as well as less extensive studies of other metals. He ignores exchange and correlation effects throughout but does take into account the momentum dependence of the pseudopotential; again, second-order perturbation theory was used. Wiser has carried the work of reference 33 to the next step by adding exchange and correlation effects in the lowest order approximation of the dielectric formalism. Finally, Pick and Sarma have been able to effect a simplification of the procedures used by Harrison.
In addition, I have carried out extensive SCF perturbation calculations without exchange and correlation, but to all orders in the pseudopotential. On the basis of this as yet unpublished work\textsuperscript{39} and of the above cited publications I have come to the following conclusions:

1. $F_{e}$ is large and hard to calculate accurately. The principal difficulty here is the simultaneous inclusion of long-range and short-range electron correlation, either being easy to include by itself.

2. $\sum_{l} U_{ll}$ is comparable to or smaller than $F_{e}$; it is also hard to calculate accurately. The principal difficulty here is that small momentum transfers caused by the screened ion-core pseudopotential make an important contribution to $U_{ll}$. For these, the corresponding matrix elements of the pseudopotential are large, in fact equal to $-2/3 E_{F}$ in first order, so that perturbation theory converges poorly.

3. On the other hand, $\sum_{l} U_{mn}$ is relatively small, and $U_{mn}$ is easier to calculate. Here only long-range correlation of electrons is important, and the relevant matrix elements of the pseudopotential are small.

4. The multiatomic force terms, $\sum_{k \leq m \leq n} W_{kmn}$ and higher terms, are still smaller but hard to calculate because of their intrinsic complexity.

5. The convergence parameter giving the ratio of an $n$-fold interaction to an $(n-1)$-fold one is, very roughly, $E_{g}/E_{F}$, the ratio of a typical, low-lying energy gap
in the conduction band to the unperturbed Fermi energy. This ratio is $\lesssim 0.1$ for most simple metals, leading one to conclude that the expansion implicit in (7.22) is fairly rapidly convergent.

6. As a direct consequence of conclusions 1-5, we can expect to make more accurate calculations of $F_S$, the structure dependent term in the free energy, than of $F_e$ and $F_C$. We have within our grasp at the present time the possibility of determining accurately the relative stability of different crystal structures at constant volume, the more difficult terms $F_e$ and $F_C$ depending only on volume for fixed composition. Some order-disorder transformations may also be accessible.

7. As a consequence of conclusion 5., an excellent empirical model for the free energy of a simple metal is

$$F = F_e + F_C + \sum_{l>m} U_{lm}.$$  

We have in (8.1) a collective term $F_e$ dependent only on electron density, a sum of monatomic terms $F_C = \sum_l U_l$ which depend on electron density as well as the nature of each atom, and finally a sum of pairwise, interatomic interactions. Johnson and March have attempted to infer $U_{lm}$ from the experimental radial distribution for liquid Na. Harrison has calculated it explicitly for Zn.

8. The model (8.1) breaks down when the pseudopotential is large so that $E_g/E_F > 0.1$ or when the free
electron Fermi sphere would be in nearly critical contact with a Brillouin zone face. Near critical contact has such consequences for the theory in its present form (logarithmic divergences) that it is quite clear that the familiar lattice parameter anomalies \(^4\) and order-disorder transformations such as those described by Sato in this special session \(^4\) can be explained in terms of contact, as has always been proposed, \(^1\) by a straightforward extension of the present theory.

9. Great caution is necessary in applying the model (8.1) to the theory of lattice vibrations because \(U_m\) is volume dependent. Thus, to get the same value of the velocity of long-wavelength longitudinal sound waves from a compressibility calculation and from a lattice vibration calculation, it is necessary to include \(W_{\nu m}\) and \(W_{\nu \mu m}\) in the latter. This difficulty does not arise for shear waves or for short-wavelength lattice vibrations.

10. Another cautionary note should be struck regarding the separation of properties \(P\) into structure-dependent and structure-independent parts, \(P_e + P_c\) vs \(P_s\). For example, the structure-dependent terms in the model (8.1) arise from a sum of pairwise interactions. Such sums are themselves only slightly dependent on structure for fixed atomic density. Thus one cannot infer from the empirical generalization regarding the structural insensitivity of electronic properties that \(\sum U_m\) is necessarily small
in magnitude relative to $F_e + F_c$; one can infer only that its structure-dependent part is small. The relative smallness of $\sum_{l<m} U_{l,m}$ follows only from the excellence of the cancellation in the pseudopotential. There may be metals, e.g., the noble metals, for which the pseudopotential is not small, the interatomic interactions are relatively large, and higher-order interactions are important, but yet the truly structure-dependent part of what we have called $F_s$ is relatively small.

The basic point of all of the preceding is that in addition to the structure-independent part of the sum of interatomic interactions common to many classes of materials, there are two additional, large structure-independent terms, $F_e$ and $F_c$, for metals.

IX. ACKNOWLEDGMENT

I am grateful to M. H. Coopersmith for those discussions out of which the theorem (7.16) and my understanding of its use grew.

This work has been supported in part directly by the ONR and the NSF. In addition, it has benefited indirectly through partial support of solid state theory by NASA, of the Low Temperature Laboratory of the Institute for the Study of Metals (ISM) by the NSF, of Materials Sciences at ISM by the AEC, and of Materials Sciences at the University of Chicago by ARPA.
REFERENCES

12. R. Pick and G. Sharma, to be published.
13. A review of the pseudopotential method as applied to metals has been given by J. M. Ziman, Adv. in Physics (to be published).
   [Reprinted, Dover, N.Y. 1958.]
38. R. Pick and G. Sarma, to be published.
42. H. Sato (this paper).
FIGURE CAPTIONS

1. FIG. 1. The structure of a Landau quasi-particle in a metal. An electron moving with velocity $v$ is surrounded by a comoving electron deficiency, or correlation hole, and by the backflow of the other electrons associated with the motion of the correlation hole.

2. FIG. 2. Variation of effective charge, or radius times potential, for the ion core Si$^{++++}$. $Z$ corresponds to the core Coulomb potential, and $Z'$ corresponds to the pseudopotential produced by the cancellation of the core potential by the repulsive potential associated with the 1s and 2s core states.

3. FIG. 3. Dependence of an electronic property $P$ on atomic concentration in alloys of noble metals with elements in the IIIB through VB columns of the periodic table. Schematic.

4. FIG. 4. Dependence of the same electronic property $P$ as in FIG. 3 on electron-to-atom ratio e/a for the same set of alloys. Schematic.
Electron Deficiency • Backflow

Fig. 1
Fig. 3
Fig. 4