N.A.S.A.

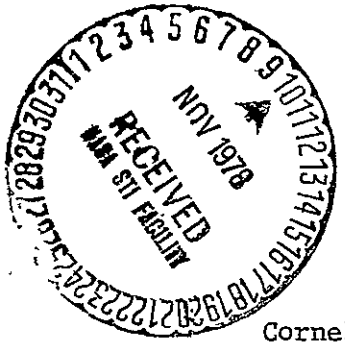
ASTROPHYSICAL MATERIALS SCIENCE:

THEORY

FINAL TECHNICAL REPORT

NEIL W. ASHCROFT
PROFESSOR OF PHYSICS

August 1972 to September 1978



Laboratory of Atomic

and

Solid State Physics

Cornell University, Ithaca, N.Y. 14853

(NASA-CR-157782) ASTROPHYSICAL MATERIALS SCIENCE: THEORY Final Technical Report, Aug. 1972 - Sep. 1978 (Cornell Univ., Ithaca, N. Y.) 192 p HC A09/MF A01 CSCL 03B

N79-10978

G3/90 Unclas 36040

GRANT NUMBER

NGR-33-010-188

ASTROPHYSICAL MATERIALS SCIENCE: THEORY

PROJECT SUMMARY: 1972-1978

Since the initial award of Grant NGR-33-010-188 in summer of 1972, the aim of the project "Astrophysical Materials Science: Theory" has been to develop analytic methods to better our understanding of common astrophysical materials particularly those subjected to extreme physical conditions. The program has been administered in the past by the staff of the Lewis Research Center, National Aeronautics and Space Administration, Cleveland, Ohio. Beginning Oct. 1, 1978 the project will be administered by N.A.S.A. Washington, re-appearing under the same title as NSG-7487.

This document briefly summarises the research discoveries and work carried out over the last six or so years. Hydrogen and helium constitute by far the most abundant of the elements and it is no accident that the research has focussed heavily on these elements in their condensed forms, both as pure substances and in mixtures. It will be seen below that the research has combined the fundamental with the pragmatic. The proper and complete understanding of materials of astrophysical interest requires a deep appreciation of their physical properties, especially when taken into the unusual ranges of extreme conditions. Fundamental theoretical condensed matter physics has played a very important part in the research to date, and will continue to be a dominant element in the research carried out under NSG-7487. The collaborati with the experimentalists (Prof. Ruoff and his group) have also been exceedingl beneficial, and this too will continue in the future.

The research will now be summarized. (Notice that publication #3(a) on aluminum under high-pressure is discussed in the Final Technical Report on NGR-33-010-189.)

Paper #1, on the ground state energies of simple metals developed the method of structural expansions for use in determining the equation of state of metallic hydrogen (and indeed other metals) up to 4th order in perturbation theory. Previously, work in the Soviet Union and elsewhere had made predictions on the nature of the structure of metallic hydrogen based on lower order perturbation theory. Paper #1 called this into question, at least for static lattices.

Paper #2 concerned itself with nature of the deep interior of Jupiter, particularly with respect to the transport properties. We were able to calculate both the electrical and thermal transport properties of the planetary interior and hence comment on the origin of the Jovian magnetic field.

Paper #3 is devoted to a problem in molecular hydrogen, specifically the nature of the interaction between molecules at short range and the importance of multi-center terms in arriving at an adequate description of the thermodynamic functions of condensed molecular hydrogen.

Paper #4 returned to the subject of Paper #1 and took up the question of proton dynamics, again arriving at a method applicable to many metals. In accounting for the structural energies in a dynamic lattice we also obtained a method for determining x-ray structure factors (particularly diffuse thermal scattering) which has been very useful.

Paper #5 addresses a problem raised in Paper #2, namely are metallic hydrogen and metallic helium mutually soluble under the conditions prevailing in the deep interior of Jupiter? The results of the calculations presented in Paper #5 show fairly convincingly that almost complete phase separation is to be expected and this has interesting consequences in the transport propertie as a function of depth into the planet.

Paper #6 tackles a question emerging from Paper #4, namely, can the proton and electron degrees of freedom really be separated when dealing with the

thermodynamic functions of hydrogen, or should they be treated as coupled systems? The latter is found to be the case and the structural consequences are really quite important. Simple structures are favored by this approach, rather than the grossly anisotropic structures proposed by the Soviet groups.

Paper #7 continues the work of Paper #5, but continued into the domain of liquid rather than solid solutions of hydrogen and helium. The miscibility gap in the solid is found to persist in the liquid alloys unless the temperature gets exceedingly high. This has application in some stellar exteriors.

Paper #8 begins a study of molecular hydrogen and its band-structure and continues the work begun in Paper #3. The ultimate intent is the determination of the thermodynamic functions of the molecular phase, and then the estimation of the metallization pressure. The results of the calculation introduce the notion that metallization by isostructural band-overlap may be a possibility.

Paper #9 deals with the quantum aspects of ground state defects in hydrogen and asks whether "quantum-defectons" can be present in metallic hydrogen crystals, and if so whether they can co-agulate into macroscopic voids whose surfaces may then be unstable to molecule formation. This prospect is ruled out by calculation: again a general method for dealing with systems other than hydrogen is introduced.

Paper #10 introduces a new idea: that the ground state of metallic hydrogen might be a quantum liquid. To obtain the ground state energy of such a system it is then necessary to extend the theory of liquids somewhat and the paper deals with a method for obtaining the necessary distribution functions.

Paper #11 then takes up the idea of Paper #10 to calculate the ground state energy of a proposed liquid phase of metallic hydrogen and indeed finds

that up to third order at least (in the electron-proton interaction) such a state is a very strong possibility. It also examines the likelihood of partially ordered magnetic phases, and notes that some of the ordering energies are quite characteristic of superconducting ordering energies.

Paper #12 extends the notions discussed in Paper #2 and discusses both the metallic and insulating form of hydrogen and helium in the context of models of the interior of Jupiter and Saturn.

Paper #13 is also concerned with Jupiter and Saturn, but from the standpoint of dynamic aspects, specifically convection and the influence on it of composition gradients in the mixture of hydrogen and helium.

In concluding this report, it is worth recording that the systems studied so far have yielded a richness in their physical properties that considerably exceeded the initial expectations. There is every reason to believe that this situation will continue, and that the low temperature highly quantum aspects of both high density hydrogen and helium will remain fascinating systems for further study.

N.W. Ashcroft

Ithaca, N.Y.-Fall 1978

N.W. Ashcroft

CUMULATIVE PUBLICATIONS NGR-33-010-188

(Supported by NASA)

1. "Ground-State Energies of Simple Metals" by J. Hammerberg and N.W. Ashcro Phys. Rev. B 9, 409, (1974).

2. "Conduction in Fully Ionized Liquid Metals" by D.J. Stevenson and N.W. Ashcroft, Phys. Rev. A 9, 782 (1974).

3. "Short Range Interaction Between Hydrogen Molecules" by A.K. McMahan, H. Beck and J.A. Krumhansl, Phys. Rev. A 9, 1852 (1974).

3a. "Aluminum Under High Pressure. I. Equation of State" by Carlos Friedli and N.W. Ashcroft, Phys. Rev. B 12, 5552 (1975). [NGR-33-010-189]

4. "Thermal Diffuse X-ray Scattering in Simple Metals", by David M. Straus and N.W. Ashcroft, Phys. Rev. B 14, 448 (1976).

5. "Phase Separation of Metallic Hydrogen-Helium Alloys" by David M. Straus N.W. Ashcroft, and H. Beck, Phys. Rev. B 15, 1914 (1977).

6. "Self-Consistent Structure of Metallic Hydrogen", by David M. Straus and N.W. Ashcroft, Phys. Rev. Lett., 38, 415 (1977).

7. "Thermodynamics of Thomas-Fermi Screened Coulomb Systems", by B. Firey and N.W. Ashcroft, Phys. Rev. A 15, 2072 (1977).

8. "Combined Representation Method for Use in Band-Structure Calculations: Application to Highly Compressed Hydrogen" by Carlos Friedli and N.W. Ashcroft, Phys. Rev. B 16, 662 (1977).

9. "Einstein-Kanzaki Model of Static and Dynamic Lattice Relaxation: Application to Vacancies in Metallic Hydrogen" by J.F. Dobson and N.W. Ashcroft, Phys. Rev. B 16, 5326 (1977).

10. "Analytical Solution of Percus-Yevick and Hypernetted Chain Equations for Quantum Liquids" by Sudip Chakravarty and N.W. Ashcroft. (Submitted to Physical Review)

11. "On the Ground State of Metallic Hydrogen" by Sudip Chakravarty and N.W. Ashcroft (submitted to Physical Review).

12. "The Phase Diagram and Transport Properties for Hydrogen-Helium Fluid Planets" D.J. Stevenson and E.E. Salpeter, Ap. J. Suppl. 35, 221 (1977).

13. "The Dynamics and Helium Distribution in Hydrogen-Helium Fluid Planets", by D.J. Stevenson and E.E. Salpeter, Ap. J. Suppl. 35, 239 (1977).

THE ASTROPHYSICAL JOURNAL SUPPLEMENT SERIES, 35:239–261, 1977 October

THE DYNAMICS AND HELIUM DISTRIBUTION IN HYDROGEN-HELIUM FLUID PLANETS

D. J. STEVENSON* AND E. E. SALPETER
Center for Radiophysics and Space Research and Physics Department, Cornell University
Received 1976 June 23; accepted 1977 April 13

ABSTRACT

In the preceding paper (Paper I) we discussed the thermodynamic and microscopic transport properties of hydrogen-helium fluid mixtures. These results are used in the present paper for a semiquantitative analysis of the thermal and compositional history of an evolving hydrogen-helium planet such as Jupiter or Saturn.

First, the evolution of a homogeneous planet with no first-order phase transitions or immiscibilities is considered. The temperature gradient is at least adiabatic (since thermal conduction cannot transport a sufficient heat flux) and is also large enough to ensure that the fluid state prevails everywhere. Convection is therefore uninhibited by molecular viscosity, and the fractional superadiabaticity is very small, despite the inhibitory effects of rotation and magnetic field. Adiabatic, evolutionary models are discussed. The times taken for Jupiter and Saturn to reach their observed luminosities are about 4×10^9 and 2×10^9 years, respectively, essentially independent of formation details. The result for Saturn appears to be inconsistent with its actual age, assumed to be $\sim 4.5 \times 10^9$ years.

Next, the effects of a first-order molecular-metallic hydrogen transition are discussed for a pure hydrogen planet: A well-defined interface between the phases persists, despite the presence of convection. The temperature is continuous at the interface and the entropy is discontinuous, the change in entropy being equal to the latent heat of transition. Consequently, the heat content and derived "age" differ from that determined for a purely adiabatic model (by a factor between 1 and 2, depending on the unknown latent heat).

Convection in the presence of a composition gradient is discussed, and the importance of overstable modes and diffusive-convective equilibria established. The convective transport of helium away from a localized helium source is shown to be inefficient because helium diffusivity is much less than heat diffusivity.

Evolutions with helium immiscibility (but *no* first-order molecular-metallic hydrogen transition) are discussed. Helium droplets nucleate from the supersaturated mixture, grow to ~ 1 cm radius, and fall under the influence of gravity, despite the convection. Most of the energy release from this differentiation is available for radiation, and the decay time for the planet's excess luminosity is increased, typically by about a factor of 5.

Finally, more complicated cases are discussed which include both immiscibility and the first-order character of the molecular-metallic hydrogen transition. The Gibbs phase rule leads to a discontinuity of the helium fraction at the transition, the formation of a helium-rich core, and an energy release comparable to that for immiscibility. This core can grow at the expense of the helium content in either the metallic or molecular region. In some cases, the molecular envelope helium content is actually enhanced by upward convective transport of helium.

The various parameters (especially the critical temperature of the molecular-metallic hydrogen transition) are too uncertain for detailed quantitative conclusions. The success of adiabatic, homogeneous evolutionary calculations for Jupiter suggests that helium differentiation has not yet begun for that planet or has begun very recently ($\lesssim 10^9$ years ago), which in turn suggests that the critical temperature for the molecular-metallic hydrogen transition cannot greatly exceed 20,000 K. Helium differentiation in Saturn (and deviations from primordial abundance for helium and minor constituents in the atmosphere) appears to be required to explain the observed excess luminosity.

Subject headings planets: abundances — planets: interiors — planets: Jupiter

I. INTRODUCTION

Modeling of the giant planets is a well-constrained problem and has reached a quite high level of sophistication in recent years. Present models of Jupiter (Podolak and Cameron 1975, Zharkov and Trubitsyn 1976, Hubbard and Slattery 1976; Stevenson and Salpeter 1976; Podolak 1977) and Saturn (Podolak and Cameron 1974, Zharkov and Trubitsyn 1976) are substantially in agreement regarding the major

features of these planets However, none of these models systematically investigates the implications of the hydrogen-helium phase diagram. The hydrogen and helium are assumed to be uniformly mixed, and first-order phase transitions are either assumed to not exist, or are inadequately treated. In the preceding paper (Stevenson and Salpeter 1977, hereafter Paper I) the phase diagram was discussed in detail, and in this paper, those results are applied to the thermal and compositional history of the hydrogen-helium planets.

Before outlining our approach to this problem, we summarize the main features of Jupiter and Saturn which are common to all the models referenced above. For Jupiter, these features are (*a*) a composition that is roughly 65% H, 30% He, and 5% other elements by mass, the latter being somewhat concentrated toward the center of the planet, (*b*) an adiabatic temperature structure such that the temperature rises from about 180 K at $P = 1$ bar, to about 10,000 K at $P \approx 3$ Mbar (the molecular-metallic hydrogen transition) and 20,000 K at the innermost hydrogen-helium region ($P \approx 45$ Mbar); (*c*) a metallic hydrogen-helium core that is 3 or 4 times more massive than the molecular envelope.

The main features for Saturn are less well established (*a*) a composition of 50–55% H, 20–25% He, and 15–20% other elements by mass, but with wider variations conceivable, (*b*) an adiabatic temperature structure such that the temperature rises from about 140–150 K at $P = 1$ bar to about 8500 K at $P \approx 3$ Mbar (the molecular-metallic hydrogen transition) and a central temperature of perhaps ~11,000 K; (*c*) a metallic hydrogen-helium core that is as little as one-third or as much as equal in mass to the molecular hydrogen envelope. For more details and comparisons for Jupiter and Saturn, see Stevenson (1977).

The main question we address in this paper is, Are the above models consistent with the hydrogen-helium phase diagram? In attempting to answer this, the following subsidiary questions necessarily arise:

1. Under what circumstances does a hydrogen-helium planet have an adiabatic thermal structure? Since the discovery of the excess infrared emission of Jupiter (Aumann *et al* 1969, Ingersoll *et al.* 1976) and Saturn (Aumann *et al.* 1969, Nolt *et al.* 1974; Rieke 1975), it has been assumed that these planets are convective almost everywhere and hence adiabatic However, this is not correct if there are first-order phase transitions or composition gradients

2. Under what circumstances is a hydrogen-helium planet homogeneous? It is inevitable that some part of the planet will eventually evolve into a phase excluded region of the hydrogen-helium phase diagram, either because of the immiscibility or because of the Gibbs phase rule requirement that the helium content be discontinuous at the molecular-metallic hydrogen phase transition The only doubt is whether this has occurred already, is occurring now, or will only occur in the future evolution of Jupiter or Saturn Inhomogeneity is ensured for a temperature less than about 10,000 K at the molecular-metallic transition. The similarity between this and the actual temperature predicted by homogeneous models may not be a coincidence.

3 What implications does inhomogeneity have for the thermal evolution? Recent evolutionary calculations for Jupiter (Graboske *et al* 1975, Hubbard 1977) appear capable of explaining the excess infrared emission as the release of primordial heat content from a homogeneous planet. A similar calculation for Saturn (Pollack *et al.* 1977) appears to be incapable of predicting sufficient heat flux after 4.5×10^9 years. However, if gravitational layering is possible, with the more dense helium separating toward the center of the planet, then a large energy source becomes available to augment the primordial heat content (Kiefer 1967, Salpeter 1973). Helium differentiation always occurs eventually, but the details are found to be quite complicated, in general Approximate calculations indicate that the present luminosity of Saturn is readily explained by helium differentiation during the last 2×10^9 years

4. What implications do the phase transitions have for the distribution of minor constituents (e g , H_2O, CH_4, NH_3)? Although we will not attempt a quantitative answer to this question, it is found from quite general considerations that the atmospheric composition is *not* in general representative of the bulk composition of the planet, even at levels deeper than any possible clouds In view of the difficulty of estimating atmospheric helium abundance from remote observations, this fact may be the best observational test of our theory

5. Can atmospheric observations be used to determine constraints on the thermal evolution of a fluid planet? The present distribution of constituents depends in a complicated way on the previous evolution of the planet. Unfortunately, we find that the current uncertainties in the hydrogen-helium phase diagram and transport properties preclude any firm predictions that relate the present compositional distributions to the past thermal evolution

In this paper we proceed from the simple to the complex. In § II we discuss the particularly simple case of a homogeneous planet in which there are no first-order phase transitions. The assumption of homogeneity is common to almost all recent models of the evolution and internal structure of Jupiter. In this particular case, convective heat transport dominates almost everywhere, and the specific entropies of the atmosphere and deep interior are almost equal. Homogeneous, adiabatic evolutionary calculations then indicate that the times taken for Jupiter and Saturn to reach their observed excess luminosities are about 4×10^9 years and 2×10^9 years, respectively, essentially independent of the details of planetary formation

In § III we discuss a pure hydrogen planet in which there is a fluid molecular hydrogen to fluid metallic hydrogen first-order phase transition It is assumed that convection dominates the heat transport everywhere, except possibly near the pressures and temperatures corresponding to the phase transition. This general situation was considered in detail by Salpeter

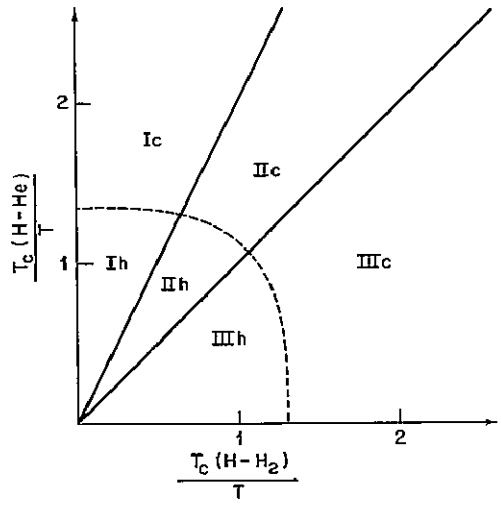


FIG. 1.—Various possible evolutionary regimes depending on the relative values of T_c(H-He), T_c(H_2-H), and T. This figure assumes $T_c(H_2\text{-He}) = 1/2T_c(\text{H-He})$ and is the analog of Fig. 6 in Paper I. In Sector I, immiscibility effects dominate. In Sector III, the effects of the molecular-metallic hydrogen transition dominate. Sector II is intermediate and complicated (see text for discussion). The dashed line separates "hot" evolutions from "cold" evolutions.

and Stevenson (1976). We apply those considerations to Jupiter and Saturn, and conclude that a well-defined interface exists between the phases, strongly inhibiting convective flow in its vicinity. Since the temperature is essentially continuous across the interface, the entropies of the two phases are found to differ by the latent heat of the transition. Under these circumstances, the temperature in the metallic core can differ by up to a factor of 2 from that predicted for a fully adiabatic planet (but the actual factor is probably nearer unity than 2). A similar effect on the derived "age" of the planet is also predicted.

In § IV we discuss some general aspects of convection in the presence of compositional gradients. Particular attention is given to the most relevant case, in which thermal diffusion is greater than particle diffusion. Overstability and the convective transport of solute are discussed.

Sections V and VI are devoted to particular evolutionary sequences. In Figure 1, the various possibilities are characterized by the critical temperatures $T_c(\text{H-H}_2)$ and T_c(H-He), for the molecular-metallic hydrogen transition and the metallic hydrogen-helium mixture, respectively. This figure is directly analogous to Figure 6 of Paper I. As in that paper, we set $T_c(H_2\text{-He}) = 1/2T_c(\text{H-He})$, where $T_c(H_2\text{-He})$ is the critical temperature for the molecular mixture. The evolution of a planet can be characterized in Figure 1 by a straight line segment, the extension of which passes through the origin. Thus the evolution lies in one of the three sectors shown. For the purposes of our considerations, the starting point of the evolution is defined as the temperature of the central hydrogen-helium region of the planet, when that region first becomes degenerate (i.e., reaches megabar pressures). The dashed line in Figure 1 further subdivides the sectors according as to whether that starting point is "hot" or "cold." A "cold" situation is one in which a phase excluded region is encountered at the beginning of the evolution. A "hot" situation is one in which the evolutionary starting point is inside the dashed boundary. It is necessary to consider several possibilities, primarily because $T_c(\text{H-H}_2)$ is so uncertain (see the discussion in Paper I). There is also considerable uncertainty as to the starting temperature for the evolution.

In § V, Sector I of Figure 1 is considered. Since the immiscibility of helium in hydrogen is the main consideration here, this section assumes, for simplicity, that there is no first-order molecular-to-metallic hydrogen transition. It is also assumed that the starting point is "hot," since the starting temperature is likely to be well in excess of $T_c(\text{H-He}) \approx 1 \times 10^4$ K. As the planet cools down, it becomes possible for droplets of helium-rich fluid to nucleate from the mixture, grow rapidly, and drift downward. The subsequent inhomogeneous evolution is discussed, using parameters appropriate to Jupiter and Saturn. Once this differentiation is initiated, a large energy source becomes available. Most of this energy is available for radiation. The rate at which the excess luminosity decreases with time is found to decrease by typically a factor of 5 relative to homogeneous evolution, once differentiation begins.

In § VI we discuss Sector III of Figure 1. The main consideration here is the first-order character of the molecular-metallic hydrogen transition, but helium insolubility is also an important consideration. Both "hot" and "cold" starting points are considered. In the "cold" case, the evolution depends on the relative densities of the coexisting helium-rich molecular phase and helium-poor metallic phase. If the former is more dense then there is a net downward transport of helium, if the latter is more dense then there is initially a small net upward transport of helium. We also discuss the "hot" case, in which there is always a net downward transport of helium.

Sector II in Figure 1 is not discussed in detail since there are no new effects in this sector that are not already present in Sector I or Sector III. The results for Sector II are, however, summarized in the concluding § VII. There, we summarize the various possible cases and their implications. A brief discussion of the disposition of minor constituents (such as water) is given, and some possible inadequacies in our analysis are assessed. Unfortunately, the uncertainties in the phase diagram and transport properties are still so great that we are unable to predict, say, the helium abundance in the Jovian and Saturnian atmospheres. However, the success of adiabatic, homogeneous evolutionary calculations for Jupiter suggest that helium differentiation has not yet begun for that planet, or has begun very recently ($\lesssim 10^9$ years ago). Helium differentiation in Saturn appears to be required to explain its observed excess luminosity, but the uncertainties are large.

II THE THERMAL EVOLUTION OF A HOMOGENEOUS PLANET

We consider first the unlikely case where the molecular metallic hydrogen transition is *not* first-order and there is unlimited solubility of helium in hydrogen. The infrared excesses of Jupiter and Saturn led Hubbard (1968, 1973) to propose that such planets are convective almost everywhere, with the consequence that the specific entropies of the deep atmosphere and metallic interior are equal (i.e., the temperature and pressure are adiabatically related). This "adiabatic hypothesis" is based on three assertions: (i) The internal heat flux is too high to be transported by conduction (electronic, molecular, or radiative) at a subadiabatic temperature gradient. (ii) The resulting internal temperature is therefore high enough to ensure that the fluid state prevails everywhere. (iii) Convection is therefore not inhibited by viscosity and readily transports the required heat flux with only a very small superadiabaticity.

The inadequacy of electronic conduction has been discussed elsewhere (Stevenson and Ashcroft 1974, Stevenson and Salpeter 1976; Stevenson 1976) for the particular case of Jupiter. Similar calculations can be made for Saturn In both cases, the thermal conductivity in the metallic core is about 2×10^8 ergs cm^{-1} s^{-1} K^{-1} (eq. [11], Paper I) and the adiabatic temperature gradient is typically 2×10^{-6} K cm^{-1}, so the conductive heat flux is typically 400 ergs cm^{-2} s^{-1}. The total internal heat flux that emerges into the atmosphere is about $(7 \pm 2) \times 10^3$ ergs cm^{-2} s^{-1} for Jupiter (Ingersoll *et al.* 1976) and $(4 \pm 1\ 5) \times 10^3$ ergs cm^{-2} s^{-1} for Saturn (Aumann *et al* 1969, Nolt *et al* 1974; Rieke 1975). In each case, the energy source must be gravitational (Hubbard and Smoluchowski 1973), but the distribution of the energy source is not accurately known. However, even for a highly decentralized energy source such as primordial heat, the heat flux at the molecular-metallic hydrogen transition is comparable to (and may even be larger than) the heat flux emerging into the atmosphere, because of the smaller surface area. In both planets, the inequality between conductive and total heat flux in the metallic region is not enormous, but is nevertheless strong enough to be almost certain. A smaller, purely conductive region near the center of each planet is not excluded.

In the molecular region, electronic or molecular conduction is negligible but radiative opacity could conceivably be low enough to allow a radiative rather than adiabatic thermal structure. However, the discussion in Paper I indicates that the opacity of pure hydrogen alone is sufficient to ensure convection, except at temperatures where the 1500 cm^{-1} to 3000 cm^{-1} window is important (i.e., 400 K $\lesssim T \lesssim$ 700 K) In this region, a solar abundance of "ices" (H_2O, CH_4, NH_3) will probably "block" the window in the pure hydrogen spectrum. It follows that a deep radiative layer, almost immediately below the observable atmosphere, cannot be discounted until we know the abundance of minor constituents in such planets. It should be noted, however, that a radiative layer is not compatible with the interpretation by Gulkis and Poynter (1972) of the thermal radio emissions from Jupiter and Saturn It would also be very difficult to reconcile with the inversion of the higher gravitational moment J_4, made by Anderson, Hubbard, and Slattery (1974).

The fluid state of these planets is assured by showing that the adiabatic temperature profile which matches the deep atmosphere gives a temperature that exceeds the melting point of hydrogen (or the liquidus of a hydrogen-helium mixture) at each depth. To a very crude approximation, the Jovian adiabat is

$$T \approx 10{,}000 \rho^{1/2} \text{ K}, \quad (1)$$

where ρ is in g cm^{-3}, and the Saturnian adiabat has the same form but is 10–20% colder. This temperature is comfortably in excess of the melting temperatures estimated in § II, Paper I. The fluid state ensures that convection is readily initiated once the adiabatic temperature is slightly exceeded, and is not inhibited by molecular viscosity

To confirm the adiabatic hypothesis, it remains to be demonstrated that the thermal convection requires only a very small fractional superadiabaticity Stevenson and Salpeter (1976) have discussed this for Jupiter, but almost identical numbers apply for Saturn. Even if allowance is made for the strongly inhibiting effect of rotation, the fractional superadiabaticity is found to be much smaller than unity. The effect of rotation has recently been analyzed in more detail (Gierasch and Stevenson 1977), and the same conclusion was reached The inhibiting effect of the magnetic field is not expected to be greater than that of rotation, if a dynamo is operating, since the Lorentz force will be at most comparable to the Coriolis force (Hide 1974) Apparently, the only other conceivable inhibition of the convection is the molecular-metallic transition, but if this is continuous, then an element of fluid can change smoothly from one phase to the other as it moves through the pressure region of the transition No supercooling or superheating would be possible, and a rising fluid element would always be only slightly less dense than the surrounding field. Of course, the region of the transition will in general have an "anomalously" large or "anomalously" small adiabatic temperature gradient In the case where the adiabatic gradient is much larger in magnitude within the transition region than elsewhere, electronic conduction can become important and the adiabatic assumption could break down This possibility is too unlikely to merit a discussion.

Provided there exist minor constituents to block the window in the molecular hydrogen opacity spectrum, the adiabatic approximation is valid for a homogeneous planet with no first-order phase transitions or immiscibilities.

Evolutionary calculations for Jupiter (Graboske *et al* 1975; Hubbard 1977) and Saturn (Pollack *et al.* 1977) have been made only for this homogeneous, adiabatic case The major part of the evolution is then

the gradual loss of primordial heat during the degenerate cooling phase. To an adequate first approximation, the luminosity is then equal to the rate of change of internal thermal energy:

$$L = 4\pi R^2\sigma(T_e^4 - T_0^4) \approx -\frac{d}{dt}\left(\tfrac{4}{3}\pi R^3 \bar{C}_v T_i\right), \quad (2)$$

where L is the excess luminosity, R is the radius, σ is the Stefan-Boltzmann constant, T_e is the actual effective temperature, T_0 is the effective temperature in the absence of an internal heat source, $\bar{C}_v$ is the average specific heat per unit volume, and T_i is some average internal temperature. Since the entire interior is assumed to be convective, T_i is related to T_e by being on the same adiabat

$$T_i \approx T_e\left(\frac{P_i}{P_e}\right)^n, \quad (3)$$

where P_i is a characteristic internal pressure, P_e is the effective pressure (i.e., the pressure at optical depth unity in the atmosphere) and $n \approx 0.25$ is the average adiabatic index. From the virial theorem (Clayton 1968),

$$P_i \approx \frac{GM^2}{4\pi R^4}, \quad (4)$$

while optical depth unity corresponds to

$$P_e \approx \frac{g}{\kappa}, \quad (5)$$

where g is the acceleration due to gravity and κ is the effective transmission opacity of the atmosphere. In the degenerate cooling phase, T_i changes more rapidly with time than C_v or R. Furthermore, the atmospheric models of Graboske *et al.* (1975) indicate that κ changes little, even as T_e changes by an order of magnitude. It follows that P_i and P_e can be regarded as constant during most of the evolution, so that $T_i \propto T_e$. The solution of equation (2) is then

$$t_0 = \frac{(\alpha)\,(\text{present heat content})}{(\text{present excess luminosity})},$$

$$\alpha \equiv (1 - q^4)\int_1^{x_m} \frac{dx}{x^3 - q^4}, \quad (6)$$

where t_0 is the "age" of the planet (the time that has elapsed since it first became degenerate), $q = T_0/T_{e,f}$ where $T_{e,f}$ is the present effective temperature, and $x_m = T_{e,f}/T_{e,i}$ where $T_{e,i}$ is the effective temperature at the beginning of the degenerate cooling. The value of α is insensitive to x_m for $x_m \gtrsim 3$. In the limit as $x_m \to \infty$,

$$\alpha = \frac{1}{3}\left[1 - \frac{4q^4}{7} - \frac{12q^8}{77} + O(q^{12})\right] \quad (7)$$

For both Jupiter and Saturn at present, $q^4 \approx 0.5$ and $\alpha \approx 0.25$. The value of α is substantially less than unity because the luminosity increases rapidly as one goes back in time. For "typical" adiabatic, homogeneous models of Jupiter (Stevenson and Salpeter 1976) and Saturn (Podolak 1974), one finds $t_0 \approx 4 \times 10^9$ for Jupiter and $t_0 \approx 2 \times 10^9$ years for Saturn, each with about 1×10^9 years' uncertainty. The more precise evolutionary calculations for Jupiter (Graboske *et al.* 1975; Hubbard 1977) and Saturn (Pollack *et al.* 1977) do not differ greatly from the above crude analysis. The major uncertainties are the present luminosity, the transmission opacity, the specific heat in the deep interior, and the average adiabatic gradient. The calculation suggests that a homogeneous Jupiter with no first-order phase transitions is consistent with the *assumed* age of about 4.5×10^9 years. (There is no direct evidence relating to the ages of the major planets, but neither is there any reason to believe that they differ greatly in age from the terrestrial planets.) The uncertainties (especially in the present luminosity) are greater for Saturn, but the small value of t_0 derived for that planet suggests that Saturn may not be homogeneous, or at least may have a different evolution from Jupiter. In "natural" (i.e., gravitational) units, Saturn has an "anomalously" large excess luminosity (see Stevenson 1977). The two most likely explanations are either that Saturn is inhomogeneous or that observers have overestimated the excess luminosity. This dilemma may be resolved with the flyby of Saturn by *Pioneer 11* in 1979. In §§ IV and V, we examine the hypothesis that inhomogeneity is the explanation. We are not precluding inhomogeneity in Jupiter either, since the uncertainties are still large in the homogeneous evolution. Furthermore, even if the planets were pure hydrogen, the adiabatic assumption would not be valid if the molecular-metallic transition were first-order. At the end of the next section we discuss how this can also affect the evolutionary time scale.

III. THE MOLECULAR-METALLIC HYDROGEN TRANSITION

We consider now a pure hydrogen planet in which the molecular-to-metallic hydrogen transition is first-order at the temperatures of interest, but in which the conductivity is always low enough (or the opacity high enough) to ensure convection everywhere well away from the transition. In a recent paper, Salpeter and Stevenson (1976) consider a self-gravitating fluid, stratified into two phases of appreciably different densities and heated from within. It is assumed that, away from the interface between the phases, the heat flux is mainly carried by turbulent convection with a very small superadiabaticity. Different modes are investigated for transporting the heat flux across the interface, and both possible signs for the phase-transition latent heat L are considered. Under a wide range of conditions, it is found that the transition region near the interface is thin, with a small change in temperature across it. The entropy difference between the two phases is then L/T, where T is the temperature at the transition. In reaching this conclusion, the following assumptions were needed: (1) a fractional density change at the transition that is not enormously less than

unity, (ii) a substantial positive surface energy σ between the phases, at both microscopic and macroscopic levels, (iii) a substantial latent heat L, with magnitude of order k_BT per particle, where k_B is Boltzmann's constant; (iv) a heat flux which is determined by conditions elsewhere, and whose average is not affected by the dynamics of the phase transition (in the case of Jupiter and Saturn, the heat flux is determined by conditions in the surface layers of the planet and its central temperature); (v) a Prandtl number (defined as $Pr \equiv \nu/\kappa$, where ν is the kinematic viscosity and κ is the thermal diffusion coefficient) that is not so enormously greater than unity that large-scale convective flows are inhibited by viscosity

Of all these conditions, (ii) and (v) are particularly crucial. If the molecular-metallic hydrogen transition is indeed first-order (see the discussion in Paper I), then these conditions are probably satisfied.

This conclusion is in contrast to that reached by Schubert, Turcotte, and Oxburgh (1970) in their discussion of the olivine-spinel solid-state phase transition in the Earth's mantle. They propose no entropy discontinuity, but rather a "two-phase" region where the two phases are intermingled and neither phase predominates. To understand why their conclusion is not incompatible with ours, two aspects of the problem must be considered: the predictions of linear stability analysis, and the nature of the finite amplitude flow

A linear stability analysis was carried out for $L > 0$ by Busse and Schubert (1971) They found that a state in which the phases are stratified with a well-defined interface becomes unstable to mixing when the superadiabaticity becomes so large that an upward-moving parcel of fluid can change phase, cool down (because of the latent heat), and yet still remain buoyant For $L \approx k_BT$, this requires a fractional superadiabaticity of order unity This instability criterion is apparently satisfied in the Earth, where viscosity greatly inhibits the flow in the *solid* phases, and the superadiabaticity must be large. This criterion is *not* satisfied for fluid phases in Jupiter or Saturn, where the superadiabaticity has a very small average value.

The second aspect of the problem is the nature of the finite amplitude flow. Turcotte and Schubert (1971) consider a simple, one-dimensional model for the flow and deduce a "two-phase" region. Since the two phases have different densities, there is a tendency for them to separate under the action of gravity. However, in the high viscosities prevailing in the Earth's mantle, the rate of separation is no greater than convective speeds elsewhere, so a dynamic steady state can be envisaged in which a two-phase region persists. In our situation, where molecular viscosity is essentially irrelevant, no two-phase region is conceivable in steady state, since it would separate almost at sound speed, on a time scale much less than typical convective time scales. To summarize, the most important difference between the Earth's mantle and the interiors of fluid hydrogen-helium planets is the factor of $\sim 10^{24}$ difference in Prandtl numbers.

This does not prove that our conclusion of an essentially "isothermal" (rather than "adiabatic") interface is correct To prove that, we would need to consider all possible modes for finite-amplitude disturbance of the interface This has not been done, but those modes that were considered were found to be stable (Salpeter and Stevenson 1976). Turner (private communication) has pointed out that a major (possibly *the* major) source of mass transfer between the phases was not considered in Salpeter and Stevenson (1976). Experiments on turbulent entrainment across density interfaces (between fluids of different composition) in the large Reynold's number limit (Turner 1968*b*; Linden 1973; Long 1975) indicate that a small amount is ejected at high speed from one fluid into the other during the recoil of a large eddy that has hit the interface The ejection velocity is comparable to $\mathscr{U}_W$, the wave velocity on the interface

$$\mathscr{U}_W = \left(gl \frac{\Delta\rho}{\rho} \right)^{1/2}, \tag{8}$$

where g is the acceleration due to gravity, l is a length scale characterizing the turbulence (i.e, eddy size), $\Delta\rho$ is the density contrast at the interface, and ρ is the average fluid density. The amount ejected (in each direction) can be expressed as an entrainment velocity $\mathscr{U}_e$ (the ejected volume per unit interface area per unit time) given by

$$\frac{\mathscr{U}_e}{\mathscr{U}} = \left(\frac{\mathscr{U}}{\mathscr{U}_W} \right)^n, \tag{9}$$

where $\mathscr{U}$ is a characteristic turbulent (convective) velocity for eddy size l, and $n = 3$ according to Turner (1968*b*) and Linden (1973) Neglecting rotation, $\mathscr{U} \approx 10$ cm s^{-1} for Jupiter and Saturn and $l \approx 10^9$ cm (Hubbard and Smoluchowski 1973), so that $\mathscr{U}_e \approx 10^{-14}$ cm s^{-1} The latent heat flux $\mathscr{U}_eL$ is therefore $\lesssim 10^{-2}$ ergs cm^{-2} s^{-1} in magnitude, and negligible compared with the sensible heat flux. Unlike the experiments, the two fluids are phases of the same substance and the net effect of ejection is zero. (There is, however, a small but finite probability of encountering a macroscopic amount of the "wrong" phase at large distances from the phase boundary.)

Experiments by Long on shear-induced turbulence (1975) have been interpreted as implying $n = 2$ In this case, both the latent heat and sensible heat fluxes are proportional to $\mathscr{U}^3$, but the latent heat flux is nevertheless smaller by $|L|/\rho\mathscr{U}_W^2 < 1$. In this case, the entrained fluid, although small in total volume, can have a thermal effect comparable to the sensible heat flux. Even if Long's experiments are applicable (which they probably are not), the interface would still be well defined, although the convection would be substantially different from the "normal" ($n = 3$) case

An "isothermal" interface appears to be ensured provided $\mathscr{U} \ll \mathscr{U}_W$ and $R_e \equiv \mathscr{U}l/\nu \gg 1$, where ν is the kinematic viscosity The conclusions of Salpeter and Stevenson (1976) can be applied to Jupiter and Saturn as follows In the molecular-metallic hydrogen transition, the metallic phase is about 30% more dense than the molecular phase. The sign of L is not known, but

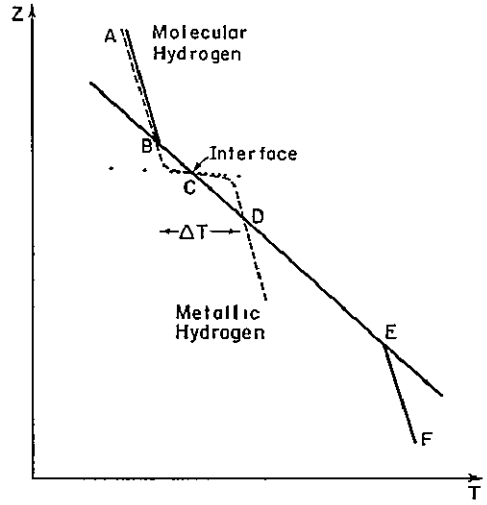


FIG. 2.—Temperature versus vertical coordinate z, for positive latent heat and no nucleation. $BCDE$ is part of the phase boundary, while AB and EF are adiabats corresponding to the same specific entropy. In a fully adiabatic case, the temperature profile would be $ABEF$, with a two-phase region between B and E. The actual temperature profile (---) is almost adiabatic except for a thin region near the interface. This region, labeled by ΔT, is exaggerated for clarity. The temperature profile for pure conduction (·) is also shown.

$|L|/k_BT$ is probably slightly less than unity (Stevenson and Salpeter 1976). Consider the case where $L > 0$ and no nucleation of one phase within the bulk of the other is possible. We predict the formation of a thermal boundary layer between the phases, in which heat conduction dominates (small-scale convection is inhibited by heat leakage or molecular viscosity). A simple mixing-length analysis yields a boundary layer thickness of order 10 cm, across which there is a very small temperature drop $\Delta T \approx 10^{-2}$ K, as shown in Figure 2 (ΔT is enlarged for clarity). Flow across the phase boundary is inhibited by the density difference and the inability of a macroscopic volume of fluid to change phase instantaneously. Instead, there are gravity waves on the interface, with amplitudes as great as 10^3 cm for the longest wavelengths $\lambda \approx 10^9$ cm. This mainly represents a moving up and down of the boundary layer, with the actual thickness of the boundary layer itself being appreciably less.

Suppose, now, that nucleation is possible. It is evident from Figure 2 that the fluid between B and the interface C is supercooled and molecular, while the fluid between C and D is superheated and metallic. At $T \approx 10^4$ K in Jupiter or Saturn, homogeneous nucleation is probably the only nucleation mechanism. Using a surface energy comparable to that of pure metallic hydrogen relative to vacuum (about 0.1 eV per surface atom, according to the theory of Lang and Kohn 1970), Salpeter and Stevenson find that the amount of superheating or supercooling is never enough to initiate significant nucleation. If heterogeneous nucleation were somehow possible, then only infinitesimal superheating or supercooling might be needed. However, it is *still* not possible for a large amount of fluid to rapidly change phase, since the superheating (or supercooling) is generally much less than the latent heat. Consider, for example, a crest of metallic hydrogen on the wavy interface. Since the interface itself can be neither superheated nor supercooled, the interface itself lies on the phase boundary. However, the fluid just below the crest is superheated and metallic. If nucleation seeds are available, then bubbles of the molecular phase begin to grow at a rate determined by the diffusion of heat onto the bubble. However, only a small amount of fluid has changed phase before the entire crest has cooled to the local phase boundary, and superheating no longer exists. This nucleation process cools the metallic hydrogen and thus contributes to an *upward* heat flux. Since the total heat flux must be constant, it follows that the thermal profile will rearrange itself so that the interface is actually more hydrodynamically quiescent than it would be in the absence of nucleation.

In the case $L < 0$, no supercooled or superheated regions arise, and the thermal boundary layer is similar to that for $L > 0$ if there are no waves at the interface. The phase change of fluid at the interface in a wave crest or trough might enhance the upward heat flux, so a temperature inversion may be needed to inhibit excessive heat flow. This temperature inversion is at most about $\Delta T \approx 10^{-3}T \approx 10$ K.

The effect of planetary rotation on these considerations is small. Far from the interface, the superadiabaticity is much larger in the presence of rotation than in its absence, but it is still much less than unity. Simple mixing-length theory (without rotation) predicts a fractional superadiabaticity $\epsilon \approx 10^{-3}$ in Jupiter or Saturn, if the mixing length is of the order of the pressure scale height. Allowance for rotation (Stevenson and Salpeter 1976, Gierasch and Stevenson 1977) yields $\epsilon \approx 10^{-4}$, in similar circumstances. As one approaches the interface, a point is reached at which rotation is no longer important (i.e., Coriolis force becomes smaller than buoyancy force). This occurs at a distance z from the interface, given by

$$\frac{v(z)}{z\Omega} \approx 1 , \tag{10}$$

where $v(z)$ is the convective velocity appropriate to a mixing length z, and Ω is the planetary angular velocity. This is satisfied in Jupiter or Saturn by $z \approx 10^5$ cm, within an order of magnitude. Since the thermal boundary layer is much thinner than this, rotation is not rapid enough to change its structure.

The effect of magnetic fields on the structure of the interface is difficult to assess, especially if there is a large discontinuity in electrical properties across the interface. According to most dynamo theories (Stevenson 1974) the Lorentz force is no greater than the Coriolis force, so it seems likely that magnetic field effects are unimportant, if rotation is unimportant. Magnetic "buoyancy" of the metallic fluid immediately below the interface may enhance the amplitude of interfacial waves, but since magnetic pressure is probably many orders of magnitude less than the

hydrostatic pressure, this should not be an important consideration.

To summarize If the molecular-to-metallic transition is first-order, and the conclusions of Salpeter and Stevenson (1976) are applicable, then large deviations from full adiabaticity may result In contrast to Hubbard's hypothesis, which states that

$$S_c = S_{\rm atm}\,, \tag{11}$$

where S_c, $S_{\rm atm}$ are the specific entropies of the central and atmospheric regions of the planet, respectively, we have instead

$$S_c + \Delta S = S_{\rm atm}\,, \tag{12}$$

where $\Delta S = L/T$ is the entropy change at the transition It follows that a central temperature T_c evaluated according to equation (11) could be wrong by as much as a factor of 2 (Stevenson and Salpeter 1976) in either sense. This is an extreme upper bound, and it is more likely that T_c determined by equation (11) is wrong by only 10% or 20%, but even this is not negligible in an accurate interior model. (The uncertainty in ΔS is essentially the uncertainty in the adiabat for molecular hydrogen at $\rho \gtrsim 0.1$ g cm^{-3}, since the adiabats are well known at lower densities and at metallic densities. All models of Jupiter and Saturn—except Stevenson and Salpeter [1976]—implicitly assume $\Delta S = 0$.)

The existence or absence of a well-defined interface is a qualitative feature which may have observable consequences for the multipolarity of the magnetic field, the large-scale convective pattern (Busse 1976), or the normal modes of the planet, in addition to modifying the compositional and thermal structure

We consider now the effect of this first-order phase transition on the cooling of the planet For simplicity, we assume that the actual temperature at the phase boundary is much less than the critical temperature for the first-order character of the transition, and we assume that the entropy change and volume change at the transition are independent of temperature. There are two ways in which the cooling rate differs from that for an adiabatic, homogeneous planet. First, the present heat content is different since the specific entropy in the metallic core is no longer equal to the specific entropy in the atmosphere (eq. [12]). This is a primordial latent heat effect (i e , the nonadiabatic structure resulted during the formation or very early evolution of the planet) Second, the phase boundary is evolving as the planet cools, because of the temperature-dependence of the transition pressure. This is a contemporary latent heat effect.

The primordial latent heat effect is readily evaluated by noting that the age of the planet is proportional to its present heat content (eq. [6]), provided the planet is homogeneous. In Jupiter, most of the present heat content is in the metallic core, and the temperature in this core differs from that for an adiabatic homogeneous planet by a multiplicative factor exp $(-\Delta S/2)$, where ΔS is the entropy change at the transition in k_B per proton (Stevenson and Salpeter 1976) The age of the planet is therefore modified by roughly the same multiplicative factor. This factor could be as small as 0.5 or as large as 2.0, but is probably closer to unity. The effect on Saturn is smaller, since a smaller fraction of the total heat content resides in the metallic core or in very dense molecular hydrogen.

The contemporary latent heat effect is much smaller. As the planet cools, one phase grows at the expense of the other. This leads to gravitational and internal energy changes that almost compensate, the net effect being the purely thermal one of latent heat release (Flasar 1973) According to the Clausius-Clapeyron equation,

$$\left(\frac{dP}{dT}\right)_{\rm ph} = \frac{\Delta S}{\Delta v} \tag{13}$$

where the derivative is evaluated along the phase boundary, and $\Delta v \approx 3a_0^3$/proton (Stevenson and Salpeter 1976) is the volume change at the transition. The additional luminosity from latent heat generated at the boundary, Q_L, is

$$Q_L \approx -\frac{4\pi R^2 L}{g}\left(\frac{dP}{dT}\right)_{\rm ph}\left(\frac{dT}{dt}\right)\,, \tag{14}$$

where $L = T\Delta S$ is the latent heat per gram, and (dT/dt) is the rate at which the temperature is changing at the phase boundary. Assuming $dT/dt \approx -2 \times 10^{-14}$ K s^{-1}, which is appropriate to adiabatic, homogeneous models of Jupiter (see § II), one finds that for $T \approx 10^4$ K,

$$Q_L \approx 6 \times 10^{23}(\Delta S)^2 \text{ ergs s}^{-1}\,, \tag{15}$$

where ΔS is in k_B per proton Since $|\Delta S| < 1\ k_B$ per proton (Stevenson and Salpeter 1976), it follows that Q_L is at most 10% of the total heat flux of 5×10^{24} ergs s^{-1}. In Saturn, the inequality is even greater because of the smallness of the metallic core Note that Q_L is positive regardless of the sign of ΔS. (If $\Delta S > 0$, then the metallic core grows at the expense of the molecular mantle. If $\Delta S < 0$, then the molecular mantle grows at the expense of the metallic core. In either case, heat is released)

These calculations are of limited usefulness for Jupiter and Saturn, which are *not* pure hydrogen. In fact, both planets contain a substantial mass fraction of helium The Gibbs phase rule enforces a discontinuity of helium fraction at a first-order molecular-metallic phase transition, and this can have a much larger effect on the cooling rate (see § VI).

IV CONVECTION IN THE PRESENCE OF A COMPOSITIONAL GRADIENT

Thermal convection in the presence of composition gradients is not a simple generalization of homogeneous thermal convection, because the additional available degrees of freedom can admit qualitatively new phenomena. There is an extensive literature on this problem (see, for example, Spiegel 1972), but we limit ourselves here to those conditions which arise in hydrogen-helium planets when the helium is nonuniformly

distributed In particular, we assume that $D < \kappa$ always, where D is the helium diffusivity and κ is the thermal diffusivity. We also assume that the temperature gradient is destabilizing. The first assumption is almost certainly valid for both molecular and metallic phases (see Paper I, §§ VII and VIII).

With these assumptions, it is possible to eliminate the "salt finger" modes (Turner 1967) The remaining steady states are: purely diffusive, overstable, and unstable. The purely diffusive solution is well understood and exactly solvable. It need not concern us further The unstable mode is a simple generalization of homogeneous thermal convection, and is highly efficient in the transport of heat or solute The overstable mode is qualitatively new and owes its existence to the presence of two diffusive processes of different efficiencies (Shirtcliffe 1967, Turner 1968*a*).

Consider, first, the unstable mode. In direct analogy to the well-known simple mixing-length theory, we can consider a parcel of fluid in equilibrium with the ambient medium, with composition and density given by x and ρ, respectively. The parcel is then displaced upward, expanding adiabatically and maintaining the same composition. The condition for instability is that the parcel must then have lower density than the ambient fluid, i e ,

$$\left(\frac{\partial \rho}{\partial p}\right)_{x,s} < \frac{d\rho}{dp}, \tag{16}$$

where s is the entropy, p is the pressure, and

$$\frac{d\rho}{dp} = \left(\frac{\partial \rho}{\partial p}\right)_{x,s} + \left(\frac{\partial \rho}{\partial x}\right)_{s,p}\left(\frac{dx}{dp}\right) + \left(\frac{\partial \rho}{\partial s}\right)_{x,p}\left(\frac{ds}{dp}\right),$$

which, after some elementary manipulation, becomes

$$\frac{d\rho}{dp} = \left(\frac{\partial \rho}{\partial p}\right)_{x,s} + \left(\frac{\partial \rho}{\partial x}\right)_{p,T}\left(\frac{dx}{dp}\right) + \left(\frac{\partial \rho}{\partial T}\right)_{x,p}\left[\frac{dT}{dp} - \left(\frac{\partial T}{\partial p}\right)_{x,s}\right].$$

If we define

$$\epsilon = \frac{1}{\rho}\left(\frac{\partial \rho}{\partial T}\right)_{x,p}\left[\frac{dT}{dp} - \left(\frac{\partial T}{\partial p}\right)_{x,s}\right]\frac{dp}{dz}H_p,$$

$$\chi = -\frac{1}{\rho}\left(\frac{\partial \rho}{\partial x}\right)_{q,T}\frac{dx}{dz}H_p, \tag{17}$$

where z is a vertical coordinate and H_p is the pressure scale-height, then

$$\epsilon > \chi \tag{18}$$

is the condition for instability. Generalizing the usual arguments of simple mixing-length theory, we can then derive a velocity v

$$v \approx v_s(\epsilon - \chi)^{1/2}(l/H_p), \tag{19}$$

where l is the mixing length, $v_s = (gH_p)^{1/2}$ is the sound speed, and g is the acceleration due to gravity. The heat flux F_T is of order

$$F_T \approx \gamma\rho v_s^3\epsilon(\epsilon - \chi)^{1/2}(l/H_p)^2, \tag{20}$$

where we have used the fact that

$$\frac{C_pT}{\alpha} = \gamma\frac{p}{\rho} \approx \gamma v_s^2,$$

$$\alpha = -\left(\frac{\partial \ln \rho}{\partial \ln T}\right)_{q,x}, \qquad \gamma = \left(\frac{\partial \ln P}{\partial \ln T}\right)_{s,x}, \tag{21}$$

and C_p is the constant pressure specific heat. We can also evaluate the solute mass flux F_x

$$F_x \approx \rho v_s\chi(\epsilon - \chi)^{1/2}(l/H_p)^2. \tag{22}$$

The rate at which work is done against gravity in redistributing the solute is of order $v_s^2F_x(l/H_p)$ An obvious consequence of these results is that a very small compositional gradient can have a large effect on the convection properties. For example, $\epsilon \approx 10^{-8}$ in Jupiter if $\chi \equiv 0$, and the effect of rotation is neglected (as it is above) Thus, if $\chi \gtrsim 10^{-8}$, the convection properties would be modified. In the next section, we consider situations in which $\chi \approx 1$. The effect of rotation is not negligible, of course, but it does not change the instability criterion, and roughly speaking just changes the right sides of equations (20) and (22) by the *same* multiplicative factor $\delta(l)$. [For Jupiter, $\delta(H_p) \approx 10^{-6}$, so that $\epsilon \approx 10^{-4}$ for $\chi \equiv 0$, $l = H_p$ (Gierasch and Stevenson 1977).]

Consider now the overstable mode. In this mode, the fluid is stably stratified ($\epsilon < \chi$), but small-scale fluid oscillations can grow because of the greater efficiency of heat diffusion relative to helium diffusion Consider a displacement of an element of fluid that is sufficiently small for molecular diffusion effects to be significant. In the displaced position, heat and solute diffuse from the fluid element into the surrounding ambient medium. If the density increase from this heat diffusion exceeds the density decrease from the solute diffusion, then the density contrast between the fluid element and the ambient medium is *enhanced*, and a growing oscillation is possible, driven by the thermal buoyancy force. In the absence of viscosity, the condition for overstability is

$$\kappa\epsilon > D\chi. \tag{23}$$

Molecular viscosity ν is always important, however, and the correct result incorporating ν is (Walin 1964)

$$(\kappa + \nu)\epsilon > (D + \nu)\chi \tag{24}$$

for overstability The regime of overstability is slice of (ϵ, χ)-space, bounded on one side ($\epsilon > \chi$) by the unstable region and on the other side by the stable (diffusive) regime. In Figure 3, the stability diagram is given for the situation of interest ($\kappa > D \approx \nu$, $\epsilon > 0$, $\chi > 0$).

The overstable mode is most efficient when the characteristic time for heat diffusion across a fluid

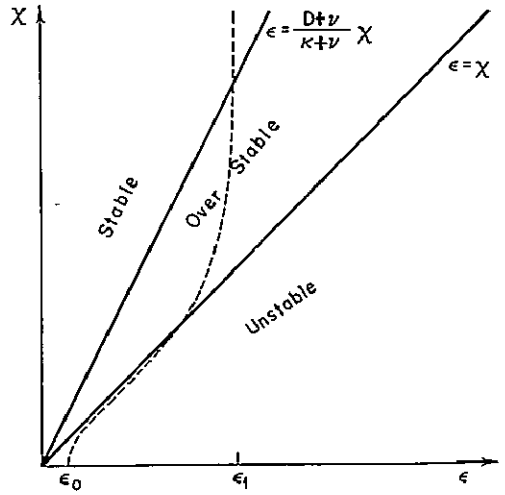


FIG 3.—The stability diagram for thermosolutal convection, assuming $\epsilon > 0$, $\chi > 0$, $\kappa > D$. The dashed line schematically represents a constant heat flux contour. For clarity, the $\chi = 0$ intercept ($\epsilon = \epsilon_0$) is shown well-displaced from the origin. Usually ϵ_1 (the value of ϵ for pure heat conduction) is many orders of magnitude larger than ϵ_0. The transition from unstability to overstability (at a given heat flux) is not well defined, but occurs in a region of ϵ that is not greatly less than ϵ_1.

element is comparable to the oscillation time

$$\lambda^2/\kappa \approx (H_p/v_s)(\chi - \epsilon)^{-1/2}, \qquad (25)$$

provided ν is not many orders of magnitude greater than κ. The characteristic horizontal length scale ("wavelength") λ is typically of order 10 cm in the situations of interest ($\chi - \epsilon \approx 1$). The vertical amplitude cannot be estimated from linear stability analysis, but experiments (Caldwell 1974) indicate that heat and solute fluxes are not very much greater than they would be from pure diffusion. This means that the amplitude of the oscillations is never enormously greater than the wavelength, a physically reasonable conclusion. Overstability should therefore be regarded as a mechanically enhanced diffusion process rather than a convective mixing process. This means that the ratio of thermal to solute fluxes should be roughly the same as it would be if only diffusion were acting. (This is only true for $\epsilon \gtrsim \alpha$ since thermal diffusion is driven by the *total* temperature gradient, not just the superadiabatic excess. This criterion is always satisfied in laboratory-sized experiments, and is satisfied in many of the situations that we consider in subsequent sections.)

In Figure 3, the dashed line schematically indicates a contour of constant heat flux. In the stable region, $\epsilon = \epsilon_1$ (a constant for all χ if we neglect the Soret effect—see Paper I, § VII). The onset of overstability is accompanied by a gradual reduction in ϵ for a given heat flux, but because of the inherent inefficiency of the overstable modes relative to normal convection, the reduction in ϵ is never very great, probably less than an order of magnitude. The transition from overstable to unstable behavior is complicated, and is not accurately represented in Figure 3. Once unstability predominates, equation (20) shows that $\epsilon - \chi \ll \epsilon$ until near $\chi = 0$, where $\epsilon \approx \epsilon_0 + \chi/3$. An interesting feature of the unstable regime in which $\epsilon - \chi \ll \epsilon$ is that equation (19) then predicts very slow convective velocities. Under these circumstances, convection is likely to be intermittent.

In thermosolutal convection, nonlocal (Turner and Stommel 1964) and time-dependent effects may occur. The following situation is of particular relevance in evolving hydrogen-helium planets.

Consider a semi-infinite pure fluid, bounded below by a rigid, perfectly conducting plate. Incident on this plate is a constant, given upward heat flux F_T. Experiments and theory (Howard 1964) indicate that an intermittent boundary layer is formed which grows by thermal diffusion until the local Rayleigh number is exceeded for a layer of thickness $\sim(\kappa t)^{1/2}$, where t is the elapsed time and κ is the thermal diffusivity. A thermal plume forms which removes the buoyant fluid from the plate, and the whole process is then repeated. Now suppose that solute is also introduced at the plane $z = 0$ at a constant mass rate F_x. Assume that at $t = 0$ there is no deviation from neutral stability in the fluid, and let $\Delta\rho_T$ and $\Delta\rho_x$ be the subsequent $z = 0$ density changes caused by heat and solute. (Both are defined to be positive, but the thermal effect is destabilizing and the solute effect is stabilizing.) The exact form of the subsequent diffusive solution need not concern us (see, for example, Jeffreys and Jeffreys 1950), but the general features are that (*a*) both $\Delta\rho_T$ and $\Delta\sigma_x$ increase as $t^{1/2}$ and their *ratio* is constant; (*b*) the characteristic distances over which the density changes extend are $(\kappa t)^{1/2}$ and $(Dt)^{1/2}$ for heat and solute, respectively. Let F_T^* and F_x^* be the respective $z = 0$ fluxes in density units. It follows that

$$F_T^* \approx \frac{\kappa\Delta\rho_T}{(\kappa t)^{1/2}}, \qquad F_x^* \approx \frac{D\Delta\rho_x}{(Dt)^{1/2}} \qquad (26)$$

These equations are approximate, but the ratio equation is *exact*.

$$\frac{F_x^*}{F_T^*} = \left(\frac{D}{\kappa}\right)^{1/2}\frac{\Delta\rho_x}{\Delta\rho_T}. \qquad (27)$$

Provided $\Delta\rho_T > \Delta\rho_x$, a thermal can still form at the plate surface, and all the introduced solute can be transported away by convection. However, if $\Delta\rho_x > \Delta\rho_T$, then a stable layer must form near $z = 0$. Experiment and theory (Linden 1974, Linden and Shirtcliffe 1976) show that a diffusive "core" forms. At the edge of this core there is a new intermittent boundary layer which has the property that $F_x^* = (D/\kappa)^{1/2}F_T^*$ locally. To conclude: If $F_x^* \le (D/\kappa)^{1/2}F_T^*$ at $z = 0$, then all the introduced solute can be transported away by convection. If $F_x^* > (D/\kappa)^{1/2}F_T^*$, then a stable diffusive layer grows, and the amount of solute transported away by convection is at most $(D/\kappa)^{1/2}F_T^*$ in density units. For relevant values of D and κ (see Paper I) this limits the work done in redistributing helium *upward* to $\sim 10\%$ of the thermal energy flux. This limit applies to initially *localized* perturbations of the helium frac-

tion (e.g , at an interface between phases, or an interface between convective and diffusive or overstable regions).

In addition to the diffusive-convective equilibrium described here, there is direct mixing of helium by entrainment (i.e., wave-breaking at the interface) This is negligible if convective speeds are more than an order of magnitude smaller than wave speeds (Linden 1974). This criterion is satisfied in most cases

Finally, we should consider whether more complicated *global* instabilities are favored relative to the simple steady states already considered. A common situation in experiments (Turner and Stommel 1964) is the formation of a steplike distribution of solute, in which uniformly mixed convective layers are separated by thin, diffusive layers where the temperature and solute concentration change rapidly. Experiment and theory (Linden and Shirtcliffe 1976) show that this is a possible steady state provided

$$\frac{\Delta\rho_x}{\Delta\rho_T} < \left(\frac{\kappa}{D}\right)^{1/2}, \qquad (28)$$

where $\Delta\rho_T$, $\Delta\rho_x$ (both positive) are now the total density drops across the fluid for the (destabilizing) superadiabatic temperature difference and (stabilizing) solute concentration difference, respectively. If this criterion is not satisfied, then the diffusive interfaces thicken with time and the system reverts to a purely diffusive or overstable state Equation (28) may not be satisfied in some of the situations considered in subsequent sections Furthermore, it is not clear whether layers could form at all. The usual laboratory and oceanographic situations in which layers form are *not* analogous to the planetary evolutions we consider in this paper.

V. HELIUM IMMISCIBILITY

In this section, we consider the effects of helium insolubility in a cooling hydrogen-helium planet. We assume throughout this section that the molecular-metallic hydrogen transition is *not* first-order. Nevertheless, the discussion of this section essentially corresponds to the "hot" case of Sector I in Figure 1.

The thermal energy content of Jupiter is about 3×10^{42} ergs at present. An even larger energy is available, in principle, if Jupiter changed from a chemically homogeneous structure to one where the denser helium resides in a central core (Kiefer 1967, Flasar 1973). Helium differentiation was originally invoked to explain the excess luminosity of Jupiter (Smoluchowski 1967), but appears to be even more desirable for Saturn (Pollack *et al.* 1977)

It might be supposed that chemical separation and gravitational layering are impossible in the presence of fully developed turbulent convection, because diffusion times are enormously large compared with convective times. Salpeter (1973) pointed out that layering may nevertheless take place in the presence of convection, if helium becomes insoluble in hydrogen.

Salpeter originally proposed that this insolubility would occur first in the metallic phase, but near the molecular-metallic transition. Our discussion in Paper I corroborates this guess At the molecular-metallic transition, helium mixed in solar proportions first becomes insoluble when the temperature drops below about 8000 K (see Fig. 3, Paper I). The critical helium

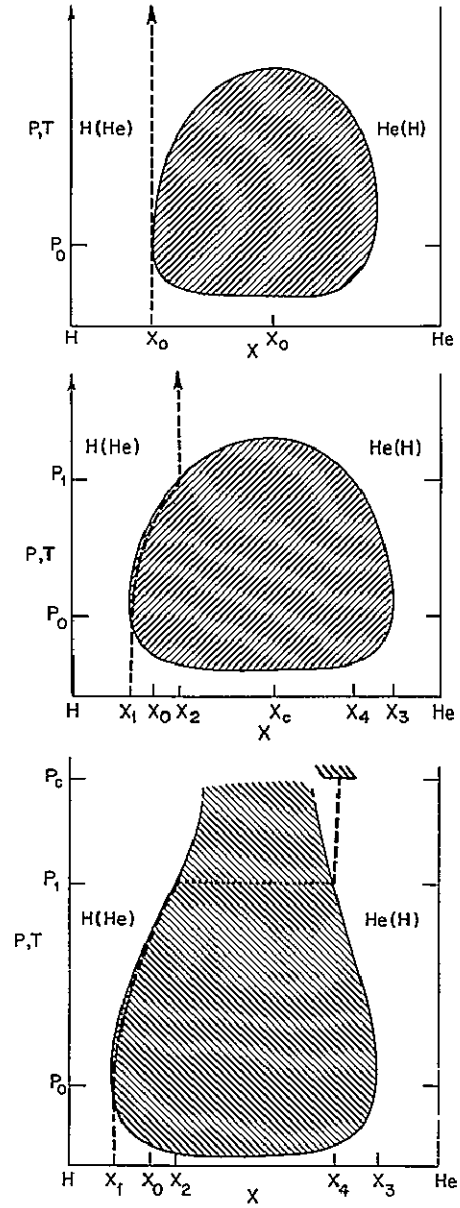


FIG 4.—The inhomogeneous evolution of a hydrogen-helium planet in which the only first-order transition is helium immiscibility The dashed line is the actual helium number fraction as a function of the actual pressure (or, equivalently, the actual temperature) within the planet The region of immiscibility is shaded The center of the planet (or the surface of a small rocky core) is $P = P_c$ In (*a*) (*top*), the planet is homogeneous, but phase separation is about to begin at $P = P_0$ In (*b*) (*middle*), the planet has cooled down more, and the region of immiscibility has expanded somewhat An inhomogeneous layer forms, but the helium-enriched central region is still predominantly hydrogen In (*c*) (*bottom*), the planet is cooler still, and now the inner region is predominantly helium

concentration x_c substantially exceeds the primordial solar abundance $x_0 \approx 0.1$ (Cameron 1973) where x is the helium number fraction. A supercooled mixture of primordial composition would therefore preferentially separate into hydrogen-rich and helium-rich phases.

Suppose $T(P)$ is the actual temperature within the planet, $x(P)$ is the helium abundance, and $T_{ph}(x, P)$ is the phase boundary temperature (the temperature below which the fluid would preferentially phase-separate). At an early stage in the degenerate cooling phase of the planet, $T(P) > T_{ph}(x_0, P)$ and $x(P) = x_0$ everywhere. Eventually, as the planet cools down, a time will be reached at which $T(P_0) = T_{ph}(x_0, P_0)$ for some pressure P_0, close to the molecular-metallic transition, as shown in Figure 4*a*. A slight further reduction in temperature leads to a macroscopic layer of supercooled metastable fluid. Droplets of helium-rich fluid begin to nucleate from the mixture and grow. We consider three important questions: What size droplets are needed for efficient helium separation? Can droplets of this size be grown? How much supercooling is needed?

First, we consider how large a helium-rich droplet must be to have a terminal velocity in excess of typical convective speeds (~ 10 cm s^{-1}). This convective speed is derivable from mixing-length theory (with the effects of planetary rotation incorporated [Gierasch and Stevenson 1977]). Let b be the radius of a droplet, let V_b be its terminal velocity, and let $\Delta\rho$ be the density difference between the helium droplet and the surrounding fluid. The velocity is found by equating gravitational and drag forces:

$$C_D \rho V_b^2 b^2 \approx \Delta\rho b^3 g , \tag{29}$$

where C_D is the drag coefficient. Assuming $\mathrm{Re} \equiv bV_b/\nu \gtrsim 10^3$, we can approximate $C_D \approx 0.05$ (Landau and Lifshitz 1959). It is also adequate to approximate $\Delta\rho \approx \rho$. Thus

$$V_b^2 \approx 20bg , \tag{30}$$

and $V_b \gtrsim 10$ cm s^{-1} provided $b \gtrsim 1$ cm. For $b \approx 1$ cm, $\mathrm{Re} \approx 10^4$, confirming our choice of C_D.

The diffusivity of helium in metallic hydrogen is roughly $D \approx 10^{-3}$ cm^2 s^{-1} (Paper I), so the characteristic diffusion time for the droplet is $b^2/D \approx 10^3$ s. This time is much less than 10^8 s, a typical large-scale convective time scale, so droplets can grow large enough to overcome convective motion before they are transported by convection to a region where they would preferentially dissolve. However, we must also consider whether droplets of this size are fragmented by the hydrodynamic pressure differences on the droplet surface. A measure of the distortion of the droplet from a sphere is the ratio of the work done by the hydrodynamic pressure in distorting a droplet to the additional surface energy created. This ratio is δ, where

$$\delta = \left(\frac{Pa_0^3}{\sigma}\right)\left(\frac{b}{a_0}\right)\left(\frac{v_b}{v_s}\right)^2 , \tag{31}$$

a_0 is a typical interparticle separation, $\sigma \approx 10^{-2}$ Ry is the surface energy per surface particle, and v_s is the sound velocity. For $b \approx 1$ cm, we find $\delta \approx 1$, so these droplets are near the maximum stable size. Regardless of the exact values of the parameters, it is clear that the downward flow of helium droplets is not *highly* inefficient.

Since the efficiency is not much less than unity, the gravitational energy release is at least of order $\rho b^3 gJH$, where J is the nucleation rate of droplets per unit surface area for the entire supercooled layer, and H is the typical distance a droplet falls. The energy release could be much larger because each droplet can produce a cascade of droplets by successive fragmentations, but an upper bound to the nucleation rate (and the supercooling) can be found by ignoring this complication. The homogeneous nucleation rate is given by Feder *et al.* (1966) as

$$J = \frac{Hv_s}{a_0^4} \exp\left[\frac{-\sigma^3 k_B T}{2\lambda^2 (k_B \Delta T)^2}\right] , \tag{32}$$

where λ is the latent heat per atom for the addition of helium-rich fluid to a droplet, and ΔT is the supercooling. For a rough estimate of ΔT, we equate the Jovian heat flux to $\rho b^3 gHJ$:

$$\left(\frac{T}{\Delta T}\right)^2 \approx \left(\frac{2\lambda^2 k_B T}{\sigma^3}\right) \ln \eta ,$$
$$\eta = \left(\frac{Hb^3}{a_0^4}\right) \epsilon_0^{-3/2} , \tag{33}$$

where $\epsilon_0^{3/2}$ is the ratio of the heat flux to ρv_s^3. For Jupiter, $\epsilon_0 \approx 10^{-8}$ and $\ln \eta \approx 100$. The theoretical calculations (Stevenson 1975) indicate that $\lambda \approx 0.5 k_B T$ at $T \approx 10^4$ K, so we finally get $\Delta T/T \approx 10^{-2}$. If heterogeneous nucleation is possible, then the required superheating would be even smaller. If the supercooling becomes larger, then more droplets are nucleated and more energy is released, heating up the fluid. This acts as a servomechanism, keeping the supercooling at just the right level to supply the required energy output. In our subsequent analysis, we neglect ΔT relative to T. It is almost certainly small enough to ensure that nucleation rather than spinodal decomposition occurs (see Paper I).

Once helium separation has been initiated, three regions are formed (see Fig. 4*b*): (i) $P < P_0$ and $x(P) = x_1 < x_0$; (ii) $P_0 < P < P_1$ where $T_{ph}[x(P), P] \approx T(P)$, (iii) $P > P_1$ and $x(P) = x_2 > x_0$. Regions (i) and (iii) are homogeneous and fully convective. The intermediate region is necessarily inhomogeneous because of the region of immiscibility. Consider, now, the life of a helium-rich droplet which nucleates out of the slightly supercooled mixture at $P = P_0$, $x = x_1$. According to Figure 4*b*, it has composition $x = x_3$. It eventually grows to about 1 cm size and begins to fall toward the center of the planet. Since diffusion times are much less than convective times, it will evolve along the right-hand-side boundary of the immiscibility region. At $P = P_1$, when the droplet has

composition $x = x_4$, the droplet merges with the inner homogeneous region. However, it must continue to evolve along the phase boundary until it either reaches the critical point ($x = x_c$) or the center of the planet. In Figure 4*b* the most likely case is shown, in which the critical point is reached first. The droplet then evaporates, enriching the inner region with helium. During this phase of the evolution the inner region is being enriched with helium, but is still predominantly hydrogen.

Later in the evolution, the innermost hydrogen-helium region reaches the critical composition x_c. After this, helium-rich droplets fall all the way without evaporating, and a predominantly helium core must begin to form This is indicated in Figure 4*c* Notice that a well-defined density discontinuity exists at $P = P_1$. The negative slope of the phase diagram on the right side ensures that the predominantly helium core is homogeneously mixed.

Consider, now, the thermal structure of the inhomogeneous intermediate layer The temperature drop ΔT, and pressure drop ΔP, across the layer are given by

$$\Delta T = T_{\rm ph}(x_2, P_1) - T_{\rm ph}(x_1, P_0)\,, \qquad (34)$$

$$\Delta P = P_1 - P_0\,.$$

Choice of x_1 (say) then gives a unique solution for the other parameters as a function of d, the layer thickness, given the phase diagram and the total helium content The thermal and solute gradients can then be evaluated from equation (17) In the limit where $d \ll H_p$, we find

$$\epsilon \approx 0\,05\left(\frac{\Delta T}{T}\right)\left(\frac{H_p}{d}\right)\,, \qquad \chi \approx 3\Delta x\left(\frac{H_p}{d}\right)\,, \qquad (35)$$

where $\Delta x = x_2 - x_1$. For the metallic hydrogen-helium phase diagram (Stevenson 1975) we typically find $\Delta T/T \approx 10\Delta x$ and $\epsilon \ll \chi$. This inequality arises because the fluid is degenerate and has a small thermal expansibility (i.e., $|(\partial \ln \rho/\partial \ln T)_{x,p}| \ll 1$) It would appear that unstable modes never exist for any layer thickness d. This could be misleading, however, since it does not take into account such nonlocal effects as "convective overshoot" (Gierasch 1971, Shaviv and Salpeter 1973) or the interaction of convection with the phase diagram itself.

Consider, for example, a fluid eddy of size l moving upward with velocity v_c This eddy impinges on the inhomogeneous layer from below, and begins to slow down as it loses buoyancy and penetrates the layer The uppermost parts of this eddy are then helium-rich relative to the phase boundary composition, and helium droplets can nucleate and grow We first evaluate the penetration of the eddy assuming that there is no nucleation Its penetration distance h can be found approximately by equating its initial kinetic energy to the work done against gravity in penetrating the lower density inhomogeneous layer:

$$\rho v_c^2 l^3 \approx \rho g_{\rm eff} h^2 l^2\,, \qquad (36)$$

where $g_{\rm eff}$ is the effective deceleration of the eddy

$$g_{\rm eff} = -gh\frac{dx}{dz}\,. \qquad (37)$$

Thus,

$$h \approx \left(\frac{v_c^2 l}{g|dx/dz|}\right)^{1/3}\,. \qquad (38)$$

For $v_c \approx 10$ cm s^{-1}, $l \approx 10^9$ cm (the largest conceivable eddies in Jupiter, say), $g \approx 10^3$ cm s^{-2} and $|dx/dz| \approx 10^{-9}$ cm^{-1}; we get $h \approx 10^5$ cm. This means that "waves" of this amplitude exist at the transition between homogeneous and inhomogeneous layers. Regardless of nucleation, it follows that if the layer thickness is less than about 10^5 cm, then convective overshoot can transport heat through the layer.

Suppose, now, that the ambient fluid is on the verge of nucleation Since nucleation is such a strong function of supercooling, nucleation would then begin immediately as the eddy began to penetrate the inhomogeneous layer. Droplets would grow at a rate limited by D, the helium diffusion coefficient (since heat diffusion is much more efficient). For $D \approx 10^{-3}$ cm^2 s^{-1}, droplets reach a size of 1 cm radius in 10^3 s. Since it takes $\sim 10^4$ s for the eddy to penetrate $h \approx 10^5$ cm, these droplets begin to separate out before the eddy comes to rest The droplet separation is inefficient, since the droplet velocity is only comparable to the convective velocity. Nevertheless, the theoretical calculations (Stevenson 1975) indicate that phase separation is accompanied by heating of the fluid (i e., the latent heat is "positive"), so part of the eddy might become buoyant if it loses some of its helium We shall now show that this instability does not in general occur, since it requires an unreasonably efficient separation process.

The uppermost portions of the eddy are helium-rich relative to the surrounding fluid by at most $h|dx/dz| = \Delta x$. Suppose a fraction δ of this excess is completely eliminated by nucleation, growth, and removal of droplets. Since the latent heat is of order $k_B T$ per particle, the fluid is hotter than the surroundings by roughly $T\delta\Delta x$. Consequently, it is more dense than the surrounding fluid by $\Delta\rho$, where

$$\frac{\Delta\rho}{\rho} \approx \Delta x\delta\left[\left(\frac{\partial \ln \rho}{\partial \ln T}\right)_{x,p} + (1-\delta)\left(\frac{\partial \ln \rho}{\partial x}\right)_{T,p}\right]\,, \qquad (39)$$

where the second term arises because the fluid is still more helium-rich than the ambient medium Since $(\partial \ln \rho/\partial \ln T)_{x,p} \approx -0.05$, whereas $(\partial \ln \rho/\partial x)_{T,p} \approx 2$, it follows that $\Delta\rho > 0$ provided $\delta < 0\,97$, which is most likely

In § II, the high-speed ejection of small volumes of fluid from one phase into the other during the recoil of an eddy was discussed for pure hydrogen. A similar effect probably occurs here, if the eddy is much larger

than the thickness of the inhomogeneous region (so that gravity waves at the now diffuse "interface" would be possible). However, the application of equation (9) indicates that the amount of ejected fluid would have no significant effect on the distribution of thermal energy or helium.

We conclude that the inhomogeneous layer is probably stable with respect to convective overshoot or entrainment Since the phase diagram (Paper I) predicts $\Delta T/T \approx 10\Delta x$, equation (35) predicts $\epsilon \approx 0.1\chi$ In Paper I, we found $(\kappa + \nu) \approx 0.5\,\mathrm{cm^2\,s^{-1}}$ and $(D + \nu) \approx 0.005\,\mathrm{cm^2\,s^{-1}}$, so $(\kappa + \nu)\epsilon > (D + \nu)\chi$ (eq [24]) and the condition for overstability is satisfied The criterion for layers (eq [28]) may be marginally satisfied, but even if it is, the temperature gradient in the inhomogeneous layer will not differ greatly from that predicted for overstability Overstable modes are inherently very inefficient, so the temperature gradient will be larger within the inhomogeneous layer than elsewhere. A consequence of this is that helium separation is accompanied by an *increasing* temperature in the innermost regions of the planet, despite the decreasing temperature externally This is illustrated schematically in Figure 5.

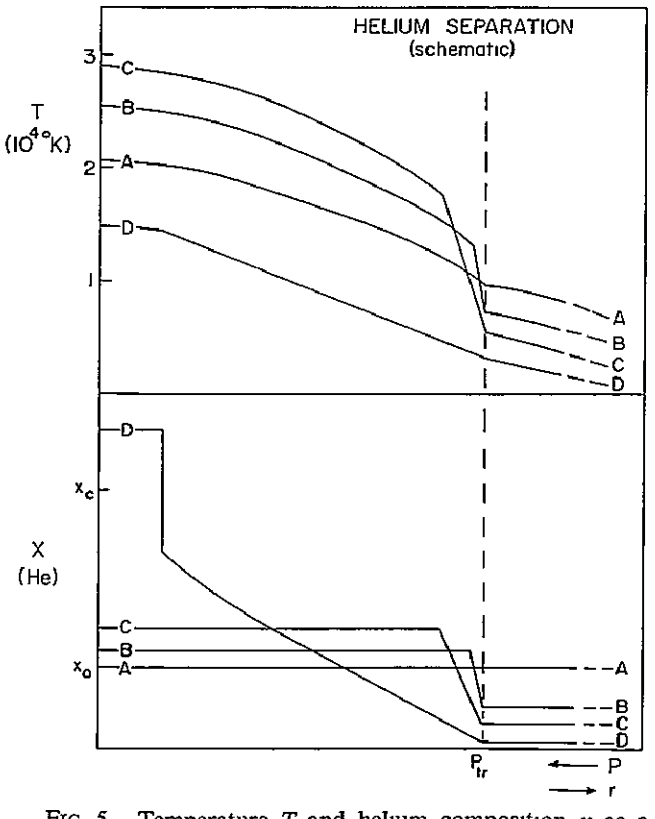


FIG 5 —Temperature T and helium composition x as a function of pressure P (or radial coordinate r) in a cooling hydrogen-helium planet The curves A, B, C, D are in order of increasing time Note that in the early stages of helium separation, the central temperature increases as the external temperature decreases Much later (D), a helium core begins to form, and the temperature gradient in the inhomogeneous region decreases because the total internal heat flux is lower

Assuming overstable modes, the thickness d of the inhomogeneous layer can be estimated For Jupiter, if we assume that the inner and outer helium fractions are $x_2 = 0\,12$ and $x_1 = 0\,06$, respectively, we find $d \approx 10^8$–10^9 cm, a significant fraction of the planetary radius (The precise value of d depends on the efficiency of the overstable modes.) As the planet cools, the heat flux becomes less, and this layer becomes even thicker. The discussion of § IV indicates that convection above the inhomogeneous layer transports some helium upward, but this is always counteracted by nucleation

To conclude, helium separation has the effect of prolonging the thermal evolution of the planet. Once it becomes thermodynamically favored, the separation proceeds with an efficiency that is neither very small nor very near 100%. It leads to depletion of helium from the atmosphere, and a thermal structure that is substantially different from that of an adiabatic, homogeneous planet An inhomogeneous layer is formed which is eventually stable with respect to large-scale convective flows, and which can encompass a significant fraction of the planetary mass.

The effect of helium differentiation on the cooling rate of the planet can be large. We shall estimate this for the early stages of differentiation, where no predominantly helium-rich region has formed (case B, Fig. 5) The correct procedure for constructing an evolutionary sequence is to compare total (gravitational and internal) energies for a sequence of models with gradually decreasing effective temperatures. However, an examination of the calculations of Kiefer (1967) and Flasar (1973) indicates that the energy release from differentiation that is available for excess luminosity or heating of the planet can be adequately approximated as Q_{Grav}, given by

$$Q_{\mathrm{Grav}} \approx \left(\frac{dM}{dt}\right)_{\mathrm{He}} gH, \tag{40}$$

where $(dM/dt)_{\mathrm{He}}$ is the rate at which a helium mass is moved down a distance H in a gravity field g. In our case, H is roughly the vertical separation of the centers of masses for the metallic and molecular fluids. Since differentiation *increases* the heat content of the core (even as the outer layers of the planet cool), we first consider what fraction of Q_{Grav} is required for this heating. Suppose the core composition changes from x_2 to $x_2 + \Delta x_2$. The mass of helium required to do this is

$$M_{\mathrm{He}} \approx \frac{4\Delta x_2 M_c}{(1 - x_2)(1 + 3x_2)}, \tag{41}$$

where M_c is the mass of the core. We assume that the mass of the inhomogeneous layer is negligible (a good approximation during the early stages of evolution). The gravitational energy release is therefore

$$E_{\mathrm{Grav}} \approx M_{\mathrm{He}} gH \tag{42}$$

However, T_2, the temperature at the boundary between the inhomogeneous layer and the metallic core, is related to x_2 according to the miscibility curve. Thus

T_2 must change to $T_2 + \Delta T_2$, where

$$\Delta T_2 = \left(\frac{dT}{dx}\right)_{\rm ph} \Delta x_2 . \quad (43)$$

According to the H-He phase diagram (Paper I),

$$\left(\frac{dT}{dx}\right)_{\rm ph} \equiv T_0 \approx 3 \times 10^4 \text{ K} \quad (44)$$

for $x_2 \approx 0.1$. The thermal energy increase of the (adiabatic) core, $E_{\rm th}$, is therefore

$$E_{\rm th} \approx \gamma C_v T_0 \Delta x_2 M_c , \quad (45)$$

where C_v is the specific heat per unit mass and γ is the ratio of the average core temperature to the boundary temperature T_2. The ratio of $E_{\rm th}$ to $E_{\rm Grav}$ is therefore

$$\frac{E_{\rm th}}{E_{\rm Grav}} \approx \frac{\gamma C_v T_0}{3gH} \approx 0.2 , \quad (46)$$

assuming $x_2 \approx 0.1$, $\gamma \approx 1.5$, $g \approx 3 \times 10^3$ cm^2 s^{-1}, $C_v \approx 2 \times 10^8$ ergs g^{-1}, and $H \approx 4 \times 10^9$ cm for Jupiter. (Similar figures apply to Saturn.) We conclude that most of the energy release from differentiation must be radiated. The ratio above is an *upper bound* corresponding to highly inefficient heat transport through the inhomogeneous layer.

We proceed now to evaluate the cooling rate during differentiation. (Cooling rate is here defined to mean dT_e/dt, where T_e is the effective temperature, since the total heat content of the planet may actually *increase* during the early stages of differentiation.) Let T_1 be the temperature at the boundary of the inhomogeneous region and the molecular envelope. We assume that T_1 and T_e lie on the same adiabat, so that

$$\frac{T_1}{T_e} \approx \left(\frac{P_1}{P_e}\right)^n , \quad (47)$$

where P_e is the pressure at T_e and n is the average adiabatic index. As the helium differentiation proceeds, P_1 changes much less rapidly than T_1 and can be assumed to be constant. The transmission opacity of the atmosphere is also only slightly affected by a change in helium content (Trafton and Stone 1974) and the gravitational acceleration also changes little, so P_e (eq. [5]) is approximately constant. The adiabatic index n is affected significantly by the helium content (especially in the outermost layers) because helium is monatomic whereas hydrogen is diatomic. Since n decreases as the helium content decreases, the decrease in T_e during differentiation is actually *less* than it would be if n were constant. (A change in n also indirectly changes P_e by changing the level in the atmosphere at which convective transport ceases to dominate.) Nevertheless, numerical calculations indicate that these effects are secondary and that P_1, P_e, and n can all be considered constant in the first approximation. Equation (47) then implies

$$\frac{d \ln T_1}{dt} \approx \frac{d \ln T_e}{dt} , \quad (48)$$

with a systematic error of typically 20–30%.

Let x_1 be the composition of the molecular envelope. Conservation of helium implies $M_c \Delta x_2 \approx -M_{\rm env} \Delta x_1$, where $M_{\rm env}$ is the mass of the molecular envelope. The gravitational energy release is therefore

$$Q_{\rm Grav} = \frac{4 M_{\rm env}}{(1 - x_2)(1 + 3x_2) T_0} \frac{dT_1}{dt} gH . \quad (49)$$

Equating $Q_{\rm Grav} = 5 \times 10^{24}$ ergs s^{-1} for Jupiter and $T_1 = 10^4$ K implies (from eq. [47]) that $dT_e/dt \approx -1.5$ K/10^9 yr. In contrast, Hubbard's homogeneous, adiabatic model for Jupiter requires $dT_e/dt \approx -7$ K/10^9 yr for the present era (Hubbard 1977). For Saturn, equation (49) with $Q_{\rm Grav} = 2 \times 10^{24}$ ergs s^{-1} implies $dT_e/dt \approx -1.3$ K/10^9 yr, whereas homogeneous evolution requires 4 or 5 times more rapid cooling. Differentiation, once initiated, therefore has the effect of dramatically changing the luminosity-time relationship and increases the Kelvin time by a factor of 4 or 5. In conjunction with the results of the homogeneous evolutionary calculations (§ II), these results suggest that Jupiter is not differentiating or at least has only recently (within the last 10^9 years) begun differentiation, whereas Saturn may already have been differentiating for $\sim 2 \times 10^9$ years. If Saturn's luminosity is indeed 2×10^{24} ergs s^{-1} at present, then the simple model outlined above suggests that the molecular envelope (and atmosphere) have already been depleted by 20–30% of its primordial helium (i.e., from $x_1 \approx 0.09$ to $x_1 \approx 0.07$).

The above calculations are applicable only if the molecular-metallic hydrogen transition is *not* first-order. In the next section, we consider the additional complications that arise in determining the helium distribution when this restriction is relaxed.

VI. MORE GENERAL CASES

In more general cases, both the first-order character of the molecular-metallic hydrogen transition and the limited solubility of helium in hydrogen must be considered. A qualitatively new feature is the Gibbs phase rule requirement that coexisting molecular and metallic phases must have different helium mass fractions. The discussion of Paper I indicates that helium would prefer to be mixed with molecular hydrogen. We consider in this section how that preference makes itself apparent in the helium distribution in a hydrogen-helium planet.

This section corresponds to Sector III of Figure 1. Both "hot" and "cold" starting points are considered because of the large uncertainty in T_c(H-H$_2$). The designation "hot" or "cold" need not imply anything about the actual central temperature of the planet. For example, a "cold" starting point corresponds to an evolution in which the actual temperature was less than the critical temperature for the molecular-metallic hydrogen transition, when the pressure first exceeded a few megabars.

a) The "Cold" Starting Point

Consider a hydrogen-helium planet in its early evolution, when the pressure in the innermost hydrogen-helium region still has not reached several

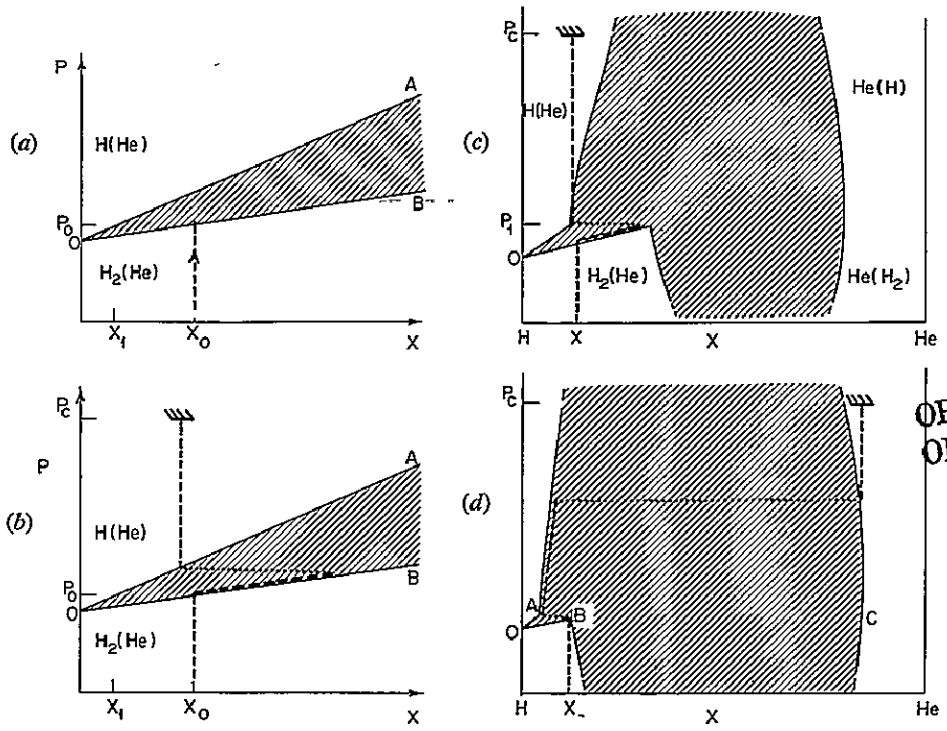


FIG 6a, b FIG 6c, d

FIG 6.—A sequence of phase diagrams for the "cold" stable case These diagrams are slices of the three-dimensional P, T, x diagrams, using the P-T relationship actually existing within the planet (The pressure coordinate P can equally well be labeled by temperature) The phase excluded region is shaded The actual helium composition is represented by the dashed line In (*a*), the innermost hydrogen-helium fluid is just beginning to be compressed into the phase excluded region Later, in (*b*), an inhomogeneous molecular layer is formed on top of a homogeneous metallic layer Subsequently, in (*c*), the molecular fluid evolves into the triple point B, and helium-rich droplets form at C An inhomogeneous metallic layer begins to form at A Even later, at (*d*), the triple point composition becomes equal to x_0, and the entire molecular layer begins to be uniformly depleted of helium The metallic hydrogen layer at A is inhomogeneous, while a homogeneous helium-rich core forms in the innermost region

megabars. We assume that the center of the planet is occupied by a small rocky core. This is a reasonable assumption from cosmogonic considerations (Podolak and Cameron 1974), but not crucial to our argument As the planet continues to contract, the pressure increases and any given element of fluid evolves upward along the dashed line in Figure 6*a* Eventually, in this "cold" case, a time is reached when the innermost hydrogen-helium fluid evolves into the phase excluded region (shaded in Fig 6*a*) This occurs at $P = P_0 \approx 3$ Mbar (see Paper I) Nucleation then becomes possible, and metallic droplets of lower helium content ($x = x_1$) form and grow Meanwhile, the molecular fluid becomes slightly helium-rich and evolves along the lower phase boundary. There are two very different cases, depending upon whether the helium-rich molecular phase is less or more dense than the helium-poor metallic phase.

Consider, first, the "stable" case in which the metallic phase is more dense. Once a macroscopic amount of this phase is formed, it settles into a layer covering the rocky core. The interface between this metallic layer and the molecular fluid is sharply defined, and lies exactly on the phase boundary for the relevant pressure. If no heat flux is transported through this interface, then the subsequent evolution is rather simple: the molecular fluid continues to evolve along the phase boundary toward a more helium-rich mixture The metallic phase remains uniformly mixed, since the new fluid added to this phase is always a little more dense than the fluid already present. Figure 6*b* shows the situation when the metallic-phase composition becomes almost the same as the original molecular-phase composition A steady-state configuration is then reached in which subsequent contraction and compression effectively process molecular hydrogen into metallic hydrogen without changing the helium content Only the rather thin intermediate molecular layer is inhomogeneous Notice that the outer molecular layer retains its primordial helium content. We have, of course, assumed that the molecular phase still remains less dense than the metallic phase, even at $P = P_1$ in Figure 6*c*

As the planet continues to cool, a time must be reached at which the molecular phase ceases to be less dense, or helium insolubility occurs The former case is discussed later In the latter case, the insolubility happens *simultaneously* in the molecular and metallic phases, as shown in Figure 6*c* (This is a general thermodynamic principle and not a consequence of

our phase diagram model) Notice that the innermost molecular region evolves into a triple point Droplets of helium-rich fluid nucleate from the molecular fluid at B and the metallic fluid at A These droplets form at C and are more dense than either of the other coexisting phases. The growth and separation of these droplets then proceeds exactly as we discussed in § V. Notice that an inhomogeneous layer begins to form in the metallic layer, but the atmospheric helium content remains primordial still.

Even later in the evolution, the triple point evolves toward the primordial helium content, and the inhomogeneous molecular layer is eliminated by helium separation Figure 6*d* shows the point beyond which the atmosphere begins to be depleted in helium The reason is that the innermost molecular region now begins to be depleted in helium relative to fluid above. This is an unstable situation, so the molecular layer remains *fully mixed* at the triple point composition, while the core becomes progressively more enriched The triple point continues to evolve to lower helium fraction as the immiscibility region expands to fill more of (P, T, x)-space The final (zero temperature) state is fully separated hydrogen above helium. If this case is applicable to Jupiter, then the current state of Jupiter is probably nearest to Figure 6*c* some helium separation may have occurred but there is no depletion from the atmosphere.

This rather simple picture can become more complicated when we consider (as we must) the transport of heat through the molecular-metallic boundary We assume a constant, given heat flux which is determined by opacity considerations in the atmosphere, but which is ultimately derived from adiabatic contraction, or helium separation, or latent heat, or even radioactive heat from the rocky core The question is whether the convective heat engine can do work transporting helium up into the atmosphere during the early evolution

Return, now, to Figure 6*a* where a metallic layer is just being formed, and the helium content of the molecular layer is beginning to be increased In the presence of a fixed heat flux F_T, this is directly analogous to the situation we discussed in § IV, in which solute is added at the lower boundary of a convecting fluid Provided the solute is added sufficiently slowly, we found that it would all be convected upward In our context, the criterion for complete mixing is that the work required to completely mix the helium upward be at most $(D/\kappa)^{1/2}F_T$, where D and κ are the helium and thermal diffusivities, respectively, for the *molecular* phase If, as seems likely, electronic degrees of freedom are available for heat conduction (see Paper I), $D/\kappa \approx 10^{-2}$, so the upward mixing of helium will be rather inefficient The actual amount of mixing depends on the value of F_T, which was surely many orders of magnitude larger during the early evolution than it is now (Graboske *et al* 1975) The amount of work required to redistribute helium upward in Jupiter is not prohibitive even now. For example, the present internal heat flux of Jupiter acting for 10^{10} years could, in principle, supply energy sufficient to double the helium content in the molecular envelope of the planet (at the expense of the metallic core) However, the small value of D/κ ensures that the actual amount of work done redistributing helium is small.

It seems likely, therefore, that the inhomogeneous layer (Fig 6*b*) will form even in the presence of the heat flux. An additional complication can then arise: since the temperature gradient must be very large in the inhomogeneous layer (with the heat flux carried by inefficient, overstable modes), it is possible (and, in fact, quite likely) that the self-consistently determined phase boundary OB in Figure 6*b* no longer has a positive slope! This can occur if the latent heat for the pure molecular-metallic hydrogen transition is negative (in the sense discussed in § III) What then happens is that the dashed line in Figure 6*b* ceases to follow the phase boundary but instead forms a purely diffusive-convective solution Helium transport in or out of the metallic phase is then maintained by diffusion at the molecular-metallic interface. The inhomogeneous layer, the thickness of which was previously determined by the slope of the phase diagram, is then a few times D/v_b, where v_b is the speed at which the molecular-metallic interface moves outward from the center of the planet. Typical values for Jupiter might be $D \approx 10^{-3}$ cm^2 s^{-1}, $v_b \approx 10^{-8}$ cm s^{-1}, and $D/v_b \approx 10^5$ cm The upward transport of helium will then be close to the upper limit of $(D/\kappa)^{1/2}F_T$ in energy units.

We now discuss the case where the molecular phase, by virtue of its helium excess, ceases to be less dense than the coexisting metallic phase The theoretical phase diagrams of Paper I indicate that this is quite likely We suppose that the early evolution is as in the stable case, but that somewhere between the Figure 6*a* and Figure 6*b*, the densities of the coexisting phases become equal. The planet continues to contract, so that at time t later, there exists a thin molecular layer of thickness $v_{ev}t$, which is more dense than the metallic fluid immediately beneath it. Here v_{ev} is a velocity characterizing the evolution rate, and is comparable to the velocity of the molecular-metallic boundary relative to the center of the planet

A Rayleigh-Taylor instability is now possible The time that disturbances of wavelength λ take to attain an amplitude $\sim\lambda$ is (Chandrasekhar 1961)

$$\tau_{RT} \approx \frac{4\nu\rho}{g\lambda\Delta\rho}, \tag{50}$$

where ν is the kinematic viscosity and $\Delta\rho$ is the density difference between the overdense molecular layer and the metallic fluid. Clearly,

$$\frac{\Delta\rho}{\rho} \approx \frac{v_{ev}t}{H_p}, \qquad \lambda \approx v_{ev}t, \tag{51}$$

since only the layer of thickness $v_{ev}t$ can participate in the instability. Equating t to τ_{RT} gives the time

for breakup of the layer:

$$t \approx \left(\frac{\nu H_p}{g v_{ev}^2}\right)^{1/3} . \quad (52)$$

For $v_{ev} \approx 10^{-4}$ cm s^{-1}, $t \approx 10^5$ s and $\lambda \approx 10$ cm. For $v_{ev} \approx 10^{-9}$ cm s^{-1} (a present-day value for the motion of the interface in Jupiter), $t \approx 10^9$ s and $\lambda \approx 0.1$ cm. Thus the instability is typically characterized by the breakup of a very thin layer of fluid into droplets of size 1 cm, to within an order of magnitude or 2. The helium diffusion time for such droplets is small (about 10^3 s) relative to the time they would take to fall a substantial fraction of a scale height, so these droplets remain in equilibrium with the phase boundary as they fall under gravity. They evolve in the direction of the arrow in Figure 7, becoming progressively more dense than the metallic phase. For the choice of phase diagram in Figure 7, these droplets merge in a helium-rich inner region at $P > P_2$. The helium-poor metallic region is shown as homogeneous in Figure 7, but it may actually tend to become stably stratified (with more helium in the innermost regions), for two reasons. First, the helium-rich and helium-poor fluids at $P = P_2$ are *not* in phase coexistence: there is a chemical potential difference tending to drive helium upward into the helium-poor fluid. Second, the shaded forbidden region in Figure 7 is actually expanding as the planet cools, so helium-poor droplets may nucleate from the helium-rich fluid and rise to merge with the helium-poor fluid above. These effects will not stop convection in the entire helium-poor layer ($P_1 < P < P_2$), but rather lead to a diffusive-convective solution of the type discussed in § IV. Except for a diffusive layer near $P = P_2$, most of the helium-poor layer continues to convect and transport some helium upward.

The subsequent evolution in this case is actually not much different from the stable case. The shaded region in Figure 7 will expand and form a diagram somewhat like Figure 6*c*. The molecular fluid at $P = P_1$ will then eventually evolve into the triple point. The situation will then be similar to Figure 6*c*, except that (*a*) a predominantly helium core has already formed, (*b*) the helium-poor metallic layer above this helium core will have a lower helium content than the primordial mixture, and (*c*) the coexisting phases at $P = P_1$ have the same density.

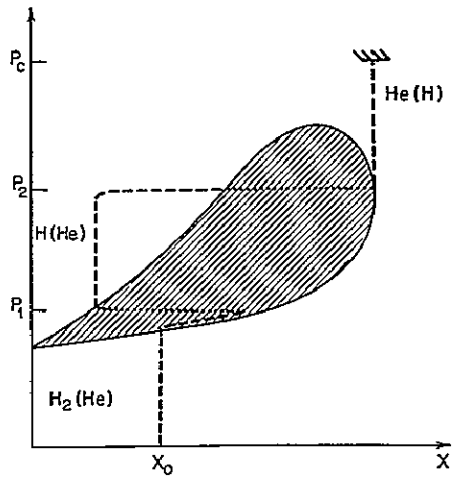


FIG. 7.—The unstable "cold" case. The coexisting phases at $P = P_1$ have the same density. Droplets break away from the molecular fluid at $P = P_1$, and evolve in the direction of the arrow to merge with the helium-rich core at $P = P_2$. Subsequent evolution of this figure is similar to Fig. 6.

The equality of densities at the molecular-metallic interface leads to another novel feature: large-amplitude gravity waves excited by convection. In Salpeter and Stevenson (1976), interfacial gravity waves were found to have small amplitude at a pure molecular-metallic interface, because of the substantial density difference between the phases. In the case where the densities are equal, however, the amplitude of the waves is limited only by the lower compressibility of the metallic phase relative to the molecular phase. Let Δz be the distance measured upward from the equal-density interface. The densities of the two phases (one stable, the other metastable) at this position are

$$\rho_{met} \approx \rho_0(1 - \alpha \Delta z/H_p) ,$$
$$\rho_{mol} \approx \rho_0(1 - \beta \Delta z/H_p) , \quad (53)$$

for the metallic and molecular fluids, respectively. The values of α and β are determined mainly by the properties of the pure hydrogen phases, rather than by the helium admixture, and are roughly $\alpha \approx 0.45$, $\beta \approx 0.55$. Consider an eddy of metallic fluid with velocity v_c and size l incident on the interface. It penetrates a distance h given by

$$\rho v_c^2 l^3 \approx \rho g(\beta - \alpha)(h/H_p)h^2 l^2 . \quad (54)$$

For simple mixing-length theory, $v_c \approx 10^{-6} v_s (l/H_p)^{1/3}$, whence we find

$$\frac{h}{l} \approx 10^{-4}\left(\frac{H_p}{l}\right)^{4/9} , \quad (55)$$

so $h \gtrsim l$ (i.e., wavelength exceeds wave amplitude) for $l \lesssim 10^2$ cm. At this size, molecular viscosity is not yet important, so it is possible for drops of size $\sim 10^2$ cm to break away from the interface and proceed a few times their own length into the opposite phase. Longer wavelengths have larger amplitude but are comparatively stable ($h/l < 1$).

Despite the larger distortions of the interface in this case, the interface will still not be completely destroyed. In other words, the considerations of Salpeter and Stevenson (1976; also see § III) still apply, and the interface is "isothermal."

b) The "Hot" Starting Point

We now consider a case in which the influence of phase transitions occurs much later in the evolution of the planet. Figure 8*a* shows one particularly likely situation in which the phase-excluded region begins small and then expands until it comes in contact with the actual (homogeneous) helium distribution at some

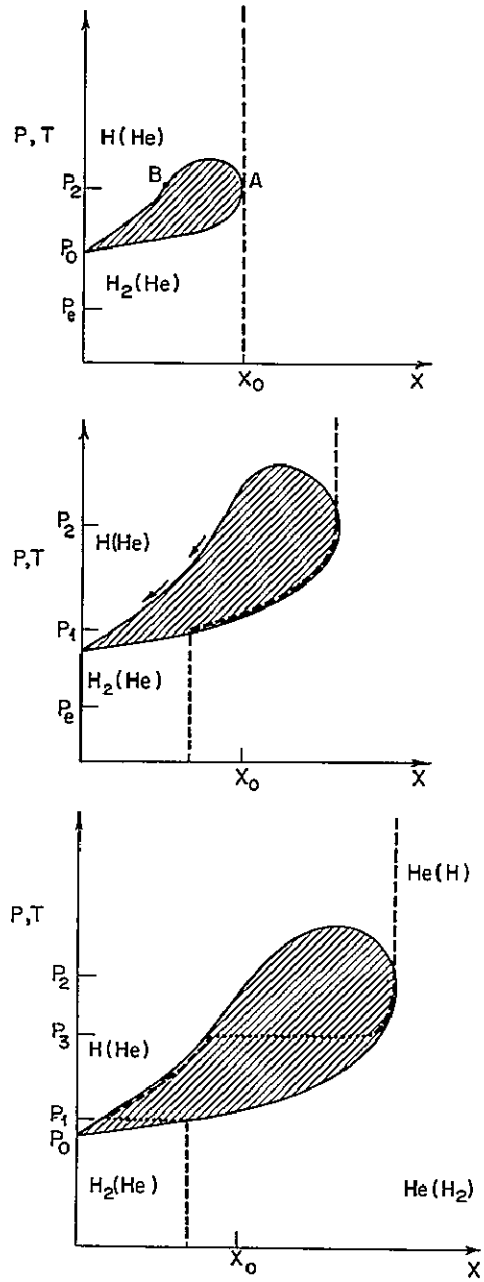


FIG 8.—The "hot" case The dashed line represents the helium concentration, and the phase excluded region is shaded In (*a*) (*top*), helium-poor droplets (*B*) nucleate from the fluid at *A* and rise along the phase diagram as indicated by the arrow These droplets eventually become pure metallic hydrogen, then evaporate at $P_e < P < P_0$ In (*b*) (*middle*), the subsequent evolution dilutes the helium content of the molecular envelope, while a helium-rich region forms In (*c*) (*bottom*), helium-poor metallic droplets at $P = P_1$ no longer have lower density than helium-rich molecular fluid, so a metallic layer forms The final state is not shown since it is equivalent to Fig 6*d*

pressure $P = P_2$. From the fluid at *A*, helium-poor metallic fluid droplets nucleate at *B* These droplets are always *less* dense for any plausible phase diagram like Figure 8*a*, so they begin to rise, maintaining equilibrium with the phase boundary, as shown by the arrow in Figure 8*a*. As usual (cf. § V), the droplets never grow much larger than 1 cm radius before fragmenting The droplets evolve to become essentially pure metallic hydrogen at $P = P_0$. They can now change phase, mainly by evaporation at the droplet surface, but also by nucleation within the droplet. In either case, the rate at which the droplet converts back to the molecular phase is determined by latent heat considerations We shall not discuss the details of this, but we assume that the resulting dilution of the molecular fluid at $P \approx P_0$ is sufficiently delocalized that convection maintains compositional uniformity. Presumably, microscopic droplets of metallic hydrogen have a very long lifetime, but even they cannot rise to pressures lower than P_e, the pressure at which the density of the droplet is the same as the ambient medium, unless they are transported by convection A steady-state metastable metallic hydrogen "mist" presumably exists, perhaps to quite low pressures, because of convective transport.

As the phase excluded region expands toward larger x, the region $P > P_2$ remains fully mixed since the region near $P = P_2$ is being continually enriched in helium Above this layer, an inhomogeneous molecular layer forms At even lower pressures, a homogeneous layer, extending up to the atmosphere, exists. This layer has a diluted composition relative to primordial, because of the continuous addition of pure metallic hydrogen droplets This is illustrated in Figure 8*b*.

The homogeneous molecular layer cannot evolve all the way to pure hydrogen because some level (labeled $P = P_1$ in Fig 8*b*) exists at which the coexisting phases now have equal density. Helium-poor metallic droplets at $P = P_1$ no longer rise, and begin to form a layer separating two molecular regions This is shown in Figure 8*c* There are now two interfaces, at $P = P_1$ and at $P = P_3$ The interface at $P = P_1$ is approximately a constant-density interface. It is rather unstable, since pieces of the metallic phase could break away and become buoyant by losing their helium as they continue to evolve along the phase boundary. The actual dynamic steady state presumably has the interface slightly displaced from the equal density level, so as to ensure greater stability The discussion earlier in this section on waves at a constant-density interface indicates that the instability is not catastrophic

The subsequent evolution is then quite straightforward Eventually an inflection develops in the molecular phase boundary in Figure 8*c*, and the phase excluded region evolves toward a diagram such as Figure 6*d* The helium distribution would then be the same as in the "cold" evolution. Thus the final state is similar for "hot" and "cold" starting points, but the *paths* by which this state is reached are different

VII DISCUSSION AND CONCLUSION

In Figure 1, the entire evolution of a hydrogen-helium planet can be characterized as a semi-infinite line segment, the extension of which passes through the origin We first summarize in qualitative fashion the

six possible evolutions corresponding to possible high-temperature starting points in Figure 1. Some of these evolutionary sequences also have further alternatives, depending on the relative densities of coexisting molecular and metallic phases.

Sector I (hot).—As the planet cools down, helium begins to separate out At first, a somewhat enriched metallic region and a somewhat depleted molecular region exist, separated by an inhomogeneous layer. Later, a predominantly helium core begins to form.

Sector I (cold).—During the early evolution, helium begins to separate out and (probably) forms a predominantly helium core. Depletion of helium from the atmosphere then begins very early in the evolution of the planet.

Sector III (hot).—As the planet cools down, helium-poor metallic droplets nucleate from the mixture, *rise*, and eventually lead to the dilution of the atmospheric helium abundance At first, a helium-poor molecular layer and a helium-enriched inner region exist, separated by an inhomogeneous layer Later, an inhomogeneous metallic layer also begins to form, while the inner region slowly evolves toward a predominantly helium composition

Sector III (cold).—(*a*) Stable If the metallic phase is more dense than the coexisting molecular phase, then initially an inhomogeneous molecular layer is formed, separating homogeneous molecular and metallic layers of essentially primordial composition A small amount of helium is transported upward into the uniformly mixed molecular envelope. Later, helium separation begins in the metallic layer and the inhomogeneous molecular layer Much later still, helium begins to be depleted from the homogeneous molecular envelope, and a predominantly helium core begins to form (*b*) Unstable If the metallic phase becomes less dense than the coexisting molecular phase, then formation of a helium core (or helium-enriched inner region) proceeds immediately, usually by depleting helium from the metallic phase Subsequent evolution is similar to the stable case, except that the molecular-metallic interface has no density discontinuity.

Sector II (hot).—This intermediate regime is difficult to characterize since it combines the effects of Sectors I and III A typical sequence of events would be that helium-poor metallic droplets nucleate from the mixture and rise to dilute the molecular envelope. Soon after, the helium-enriched inner region begins to phase-separate Subsequently, there can be as many as three inhomogeneous regions and four interfaces. These complexities arise because Sector II corresponds to a coincidental similarity of the values of T_c(H-H_2) and T_c(H-He)

Sector II (cold).—Similar complexities to the "hot" case. Figure 9 shows one possible helium distribution.

The complications of Sector II are not of concern except when T_c(H-H_2) is fortuitously very similar to T_c(H-He)

It is evident that detailed numerical calculations are premature at this stage. To give an indication of the impact of our considerations on the thermal history of a hydrogen-helium planet, we shall consider how

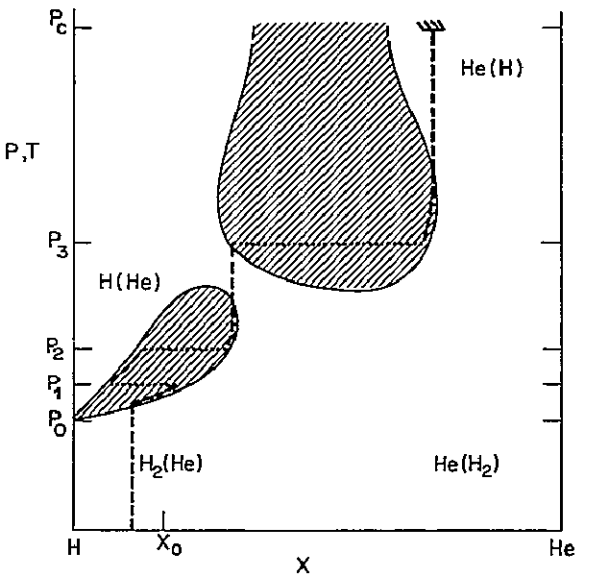


FIG 9.—An intermediate case (Sector II of Fig 1, cold starting point) This is essentially the sum of Fig 4*c* and Fig 8*c*.

the homogeneous, evolutionary models of Jupiter (Graboske *et al.* 1975; Hubbard 1977) are modified, using several possible choices of T_c(H-H_2), but for T_c(H-He) = 12,000 K and T_c(H_2-He) = 6000 K. Our considerations may actually be more relevant to Saturn, but we choose Jupiter because it is better understood and better constrained by current observations. We shall also neglect latent heat effects, since helium redistribution generally has the dominant effect on the planetary cooling rate.

In the homogeneous cooling models of Jupiter, models for the very early evolution are very speculative, and hydrodynamic effects may be important (Bodenheimer 1974), but this is of no concern here, since we consider only the evolution subsequent to the planetary center becoming degenerate ($P \gtrsim 1$ Mbar central pressure). The central temperature is then at most about 50,000 K, and the planet is probably only about 10^6 years old

Consider, first, T_c(H-H_2) = 60,000 K. In this case we have a "cold" starting point, and the first-order character of the molecular metallic transition is encountered as the center first becomes degenerate The phase diagram (Paper I) suggests that the unstable case is probably appropriate, so a helium-rich core immediately begins to form and grow at the expense of a helium-depleted metallic hydrogen region (Fig. 7). The gravitational energy release would prolong the high-luminosity phase of Jupiter, but since this phase lasts only a short time, it would not greatly affect the "age" (i e , the time taken to reach the observed excess luminosity). Nevertheless, the age is substantially affected since the phase excluded region in Figure 7 continues to expand throughout the evolution, and the helium core becomes progressively more helium-rich. The molecular envelope retains its primordial

helium abundance and is even slightly enriched with helium by upward convective transport If the effective temperature is decreased, then temperatures in the deep interior are *decreased* by a comparable fractional amount and the excluded region in Figure 7 expands slightly A rough calculation, analogous to that in § V, indicates that the rate of cooling is substantially less than that for a homogeneous planet, because of the continuing helium differentiation. (Unlike the simple calculation in § V, a precise calculation is difficult, since it necessarily depends on the efficiency of heat transport through the inhomogeneous layer) In other words, if upward convective transport of helium and latent heat effects are both negligible, then the present state of Jupiter is *not* compatible with $T_c(\text{H-H}_2) = 60{,}000$ K unless Jupiter is much older than 4.5×10^9 years

Consider, now, $T_c(\text{H-H}_2) = 20{,}000$ K The initial central temperature of 50,000 K then corresponds to a "hot" starting point in Sector III of Figure 1. Over 10^9 years elapse before the situation in Figure 8*a* occurs Helium-poor metallic droplets then form, and rise to lower pressures to dilute the molecular layer above In this case, the present state of Jupiter would have a helium-rich core which joins continuously with an inhomogeneous molecular layer and ultimately with a helium-poor molecular envelope The atmosphere would be depleted of helium, but no density discontinuity would exist anywhere in the planet (until much later in the evolution—about 10^{10} years from now) This is essentially as illustrated in Figure 8*b*. The gravitational energy released, integrated luminosity, and central temperature would all be larger than in an adiabatic, homogeneous model. Once again, it is clear that if helium differentiation is in progress, then the cooling rate would be much slower than for a homogeneous model, and the present luminosity o Jupiter would only be consistent with an age in excess of 5×10^9 years Nevertheless, $T_c(\text{H-H}_2) \lesssim 20{,}000$ K is consistent with observations, when allowance is made for all the uncertainties

Consider, finally, $T_c(\text{H-H}_2) = 0$ K In this case, the adiabatic, homogeneous evolutionary models are correct until immiscibility begins in the helium fluid (see § V). In Jupiter, immiscibility may have begun within the last 10^9 years, or is about to begin within about 10^9 years

Similar comments apply to Saturn, but with a lesser degree of certainty Present-day temperatures in Saturn's interior are lower than those at comparable pressures in Jupiter by perhaps 20% (see, for example, Podolak and Cameron 1974). Immiscibility has probably already been encountered, and this is an attractive explanation for the observed anomalously large excess luminosity (Pollack *et al.* 1977) A possible (but less likely) alternative is that the molecular-metallic transition is first-order in Saturn, but not in Jupiter (These conclusions assume that current estimates of the Saturnian excess luminosity are reliable)

In the preceding discussion we have not tried to keep account of the various latent heat effects associated with the various transitions and layer formations We predict (on the basis of the discussion of § III, and extensions thereof) that the following rules will apply. (i) In homogeneous layers, the temperature gradient is essentially adiabatic (ii) In inhomogeneous layers, the temperature gradient appropriate to overstable modes probably applies (iii) At each interface, the temperature (and not the entropy) is continuous (iv) No "two-phase" regions exist near first-order phase transitions (i.e., transitions are "abrupt").

These rules provide a unique prescription for evaluating the temperature everywhere.

We proceed, now, to a brief consideration of the distribution of minor constituents (such as water) In Paper I (§ VI) the partitioning of minor constituents among the various hydrogen-helium phases was discussed, but purely from a thermodynamic view. Thermodynamic equilibrium may not be achieved for two reasons. First, in the growth of droplets from a nucleation seed, any species which diffuses much more slowly than helium would not achieve equilibrium partitioning if the droplet moves to a region of substantially different thermodynamic environment during one diffusion time. For typical parameter values (§ V), droplets are 1 cm in radius and move at ~ 10 cm s^{-1} Except in special cases (such as at the beginning of differentiation), a droplet would have to move 10^8 cm or more to encounter a substantially different environment Nonequilibrium partitioning would therefore require a diffusivity less than $\sim 10^{-6}$ cm^2 s^{-1} This is unlikely in the fluid state (the helium diffusivity is $\sim 10^{-3}$ cm^2 s^{-1}, and larger molecules would not diffuse more than about one order of magnitude more slowly) The second and more important cause of nonequilibrium is the difficulty that we have already considered for helium: upward convective transport in cases where the solute would prefer to be mixed with the molecular phase (a likely situation, according to Paper I) If, as is likely, the solute diffuses less rapidly than helium, then it tends to be trapped in the helium diffusive layer (see § IV) which forms at interfaces. Any solute that diffuses more rapidly than helium probably achieves a distribution close to thermodynamic equilibrium. Unlike helium, the redistribution of minor constituents is not fundamentally limited by energy considerations (the convective heat engine could in principle transport several tens of Earth masses of material from the center to the atmosphere of Jupiter in less than 4.5×10^9 years). However, dynamic considerations may preclude efficient redistribution, just as they did for helium.

Nevertheless, any process which redistributes helium will also redistribute minor constituents The considerations of Paper I (§ VI) indicate that H_2O, NH_3, and CH_4 probably prefer molecular and helium-rich phases. An observational test of the considerations of this paper would be accurate determinations of the atmospheric compositions of the giant planets, especially Saturn. Unfortunately, the interpretation of such data is likely to be ambiguous.

We conclude by noting some of the inadequacies in the present analysis. First and foremost, our analysis lacks quantitative predictive power because the critical

temperature of the molecular-metallic hydrogen first-order transition is not known to better than an order of magnitude Further quantitative progress in the latent heats of transition and the molecular hydrogen-helium miscibility gap is also needed. Until these parameters are known, no detailed evolutionary model of Jupiter or Saturn can be very reliable. (Conversely, evolutionary calculations may be useful in imposing constraints on the various poorly known parameters)

Numerous assertions made in this paper about the properties of convection in turbulent, inhomogeneous fluids must be regarded as nonrigorous Even if we knew the hydrogen-helium phase diagram exactly, our predictions could be subject to error, simply because we may have overlooked some convective mode or instability.

Notwithstanding these admissions of ignorance, the following conclusions are indicated:

1 The major cause of deviations from homogeneous, adiabatic evolution is helium differentiation. Latent heat effects (either contemporary or primordial) are likely to be much less important (It is not possible to have latent heat effects without some helium differentiation and vice versa)

2 Helium differentiation can occur either because of immiscibility or because of the required discontinuity in helium fraction at a first-order molecular-metallic hydrogen transition.

3. Regardless of the cause of differentiation, it is almost invariably an ongoing process which, once initiated, has a dominant effect on the cooling rate of the planet for all subsequent time

4. The assumed age and known luminosity of Jupiter indicate that helium differentiation began $\lesssim 10^9$ years ago, or will begin in $\lesssim 10^9$ years from the present time This implies that the critical temperature $T_c(\mathrm{H\text{-}H_2})$ cannot greatly exceed 20,000 K

5. The assumed age and known luminosity of Saturn indicate that differentiation may have been proceeding for 2×10^9 years already, but the uncertainties are large and this conclusion is necessarily tentative

6 Helium differentiation is accompanied by a comparable (or even greater) redistribution of minor constituents This may provide an observational test of our theory.

We wish to thank P Gierasch, W. B. Hubbard, R Smoluchowski, and J. S Turner for discussions Support by National Science Foundation grants AST 75-21153 and MPS 74-17838 and National Aeronautics and Space Administration grant NGR 33 010-188 is gratefully acknowledged.

REFERENCES

Anderson, J D , Hubbard, W B , and Slattery, W L 1974, *Ap J (Letters)*, **193**, L149
Aumann, H H , Gillespie, C M., Jr., and Low, F. J 1969, *Ap. J (Letters)*, **157**, L69
Bodenheimer, P 1974, *Icarus*, **23**, 319
Busse, F H 1976, *Icarus*, **29**, 255.
Busse, F H , and Schubert, G 1971, *J Fluid Mech* , **46**, 801
Caldwell, D R 1974, *J Fluid Mech* , **64**, 347
Cameron, A G W. 1973, *Space Sci Rev* , **15**, 121.
Chandrasekhar, S 1961, *Hydrodynamics and Hydromagnetic Stability* (New York· Oxford University Press), chap. 4
Clayton, D. D. 1968, *Principles of Stellar Evolution and Nucleosynthesis* (New York· McGraw-Hill), p 138
Feder, J , Russell, K , Lothe, J , and Pound, G 1966, *Adv in Phys* , **15**, 111
Flasar, F. M 1973, *Ap J.*, **186**, 1097
Gierasch, P. 1971, *J Atm Sci* , **22**, 315.
Gierasch, P , and Stevenson, D J. 1977, unpublished
Graboske, H C , Pollack, J B , Grossman, A S , and Olness, R. J. 1975, *Ap J.*, **199**, 265.
Gulkis, S , and Poynter, R. 1972, *Phys Earth Planet Inter* , **6**, 36
Hide, R 1974, *Proc Roy Soc London*, **A366**, 63.
Howard, L N 1964, in *Proc. 11th Int. Congress Applied Mechanics* (Berlin: Springer-Verlag), p 1109
Hubbard, W B 1968, *Ap J* , **152**, 745.
——— 1973, *Ap J (Letters)*, **182**, L35
——— 1977, *Icarus*, **30**, 305
Hubbard, W B , and Slattery, W L 1976, in *Jupiter*, ed T Gehrels (Tucson: University of Arizona Press), p 176
Hubbard, W B , and Smoluchowski, R 1973, *Space Sci Rev* , **14**, 599
Ingersoll, A. P , Munch, G , Neugebauer, G , and Orton, G S 1976, in *Jupiter*, ed T Gehrels (Tucson: University of Arizona Press), p 197
Jeffreys, H., and Jeffreys, B S 1950, *Methods of Mathematical Physics* (2d ed ; Cambridge. Cambridge University Press), chap 20
Kiefer, H. H 1967, *J Geophys Res* , **72**, 3179
Landau, L , and Lifshitz, E M 1959, *Fluid Mechanics* (Reading, MA Addison-Wesley), p 171
Lang, N , and Kohn, W 1970, *Phys Rev* , **113**, 4555
Linden, P F 1973, *J. Fluid Mech* , **60**, 467
——— 1974, *Deep Sea Res* , **21**, 283
Linden, P. F , and Shirtcliffe, T. G L 1976, unpublished
Long, R R 1975, *J Fluid Mech* , **70**, 305.
Nolt, I G , Radostitz, J V , Donnelly, R J , Murphy, R. E , and Ford, H C 1974, *Nature*, **248**, 659
Podolak, M. 1977, *Icarus*, **30**, 155
Podolak, M , and Cameron, A. G W 1974, *Icarus*, **22**, 123.
——— 1975, *Icarus*, **25**, 627
Pollack, J B , Grossman, A S., Moore, R , and Graboske, H. C 1977, *Icarus*, **30**, 111
Rieke, G. 1975, *Icarus*, **26**, 37
Salpeter, E E 1973, *Ap J. (Letters)*, **181**, L83
Salpeter, E E , and Stevenson, D J 1976, *Phys Fluids*, **19**, 502
Schubert, G , Turcotte, D L , and Oxburgh, E R 1970, *Science*, **169**, 1075
Shaviv, G , and Salpeter, E E 1973, *Ap. J.*, **184**, 191.
Shirtcliffe, T. G. L 1967, *Nature*, **213**, 489
Smoluchowski, R 1967, *Nature*, **215**, 691
Spiegel, E A 1972, *Ann Rev Astr Ap* , **10**, 261.
Stevenson, D J. 1974, *Icarus*, **22**, 403
——— 1975, *Phys Rev* , **12B**, 3999
———. 1976, Ph D thesis, Cornell University
Stevenson, D J. 1977, in *The Origin of the Solar System*, ed. S F Dermott (New York Wiley)
Stevenson, D. J , and Ashcroft, N W 1974, *Phys Rev* , **9A**, 782.
Stevenson, D J , and Salpeter, E E 1976, in *Jupiter*, ed T Gehrels (Tucson: University of Arizona Press), p 85
———. 1977, *Ap J Suppl* , **35**, 221 (Paper I)
Trafton, L M , and Stone, P H 1974, *Ap J* , **188**, 649
Turcotte, D L , and Schubert, G 1971, *J Geophys Res* , **76**, 7980
Turner, J S 1967, *Deep Sea Res* , **16**, 497
——— 1968*a*, *J Fluid Mech* , **33**, 183.

Turner, J S 1968*b*, *J Fluid Mech*, 33, 639
Turner, J S, and Stommel, H 1964, *Proc Nat Acad Sci*, 52, 49
Walin, G 1964, *Tellus*, 16, 389
Zharkov, V. N, and Trubitsyn, V P 1976, in *Jupiter*, ed T. Gehrels (Tucson: University of Arizona Press), p 133

E. E. SALPETER: Newman Laboratory of Nuclear Studies, Cornell University, Ithaca, NY 14853

D. J. STEVENSON: Research School of Earth Sciences, ANU, P O Box 4, Canberra 2600, Australia

Phase separation of metallic hydrogen-helium alloys*

David M Straus and N W Ashcroft
Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853

H Beck†
Institut für Theoretische Physik, Universität Basel, CH 4056 Basel, Switzerland
(Received 10 September 1976)

Calculations are presented for the thermodynamic functions and phase-separation boundaries of solid metallic hydrogen-helium alloys at temperatures between 0 and 19000°K and at pressures between 15 and 90 Mbar Expressions for the band-structure energy of a randomly disordered alloy (including third order in the electron-ion interaction) are derived and evaluated Short- and long-range order are included by the quasichemical method, and lattice dynamics in the virtual-crystal harmonic approximation We conclude that at temperatures below 4000°K there is essentially complete phase separation of hydrogen-helium alloys, and that a miscibility gap remains at the highest temperatures and pressures considered The relevance of these results to models of the deep interior of Jupiter is briefly discussed

I INTRODUCTION

Knowledge of the phase diagram of hydrogen-helium alloys at high pressures (4–40 Mbar) is of importance in the study of the interior of the giant planets.[1–3] Phase separation of hydrogen and helium during the cooling process may partly account for Jupiter's excess emission of energy.[2] This paper presents a calculation of the thermodynamic functions and phase-separation boundaries of *solid* hydrogen-helium alloys at pressures between 15 and 90 Mbar, and at temperatures between 0 and 19 000°K These metallic systems are also of intrinsic interest, since the particles carry point charges, and the bare electron-electron, electron-ion, and ion-ion interactions are given exactly by Coulomb's law.

The calculations reported here supplement earlier results of Stevenson[3] on hydrogen-helium phase separation in the *liquid* phase. Present estimates of the melting curves of these materials[4] and of the temperature in the deep interior of Jupiter[3] indicate that both hydrogen and helium may well be liquid in the planet's interior, at temperatures far below 19 000°K. However, since the uncertainties in the calculated melting temperatures are quite large,[5] a solid-solid phase separation calculation remains of particular interest.

The phenomenon of solid-solid phase separation in alloys is not, of course, limited to the hydrogen-helium system, but is known to occur in many alloys.[6] For example, Li and Mg (both simple metals) form solid alloys at all concentrations except in the range of about (70–85)% Mg, where there exists a miscibility gap. An alloy formed in this concentration range will separate into two phases of different concentrations. It is noteworthy that the miscibility gap is still present at temperatures just below melting. The hydrogen-helium alloy is, however, different from many other alloys (such as Li and Mg) in one important respect. Whereas the difference between the Mg and Li electron-ion interactions (pseudopotentials) is small, hydrogen and helium have electron-ion interactions of very different strengths, and this difference is expected to play an important role in the thermodynamic properties of their alloys.

In Sec. II we discuss the general approach taken in formulating the Helmholtz free energy F for hydrogen, helium, and their alloys. The static internal energy $\overline{E}_s$ is calculated in Sec. III for any given configuration of hydrogen and helium (confined, however, to an underlying lattice), and is subsequently evaluated for a *randomly-disordered* configuration. Contributions to F arising from long- and short-range order are treated in Sec. IV, and the free energy associated with lattice dynamics in Sec. V. In Sec. VI we present the equations of state and the Gibbs free energy G per ion of hydrogen-helium alloys. Writing G as a function of its natural variables (pressure p, temperature T, and the relative concentration by number of helium c), we compute ΔG, which is defined by

$$\Delta G = G(p, T, c) - [cG(p, T, 1) + (1-c)G(p, T, 0)] \tag{1}$$

From ΔG we determine the curves describing solid-solid phase separation.

II HELMHOLTZ FREE ENERGY

For a system of volume Ω, the free energy F can be written

$$F(T,\Omega,c)=F_s(T,\Omega,c)+F_v(T,\Omega,c)\,, \tag{2}$$

where $F_s(T,\Omega,c)$ is the static free energy, and $F_v(T,\Omega,c)$ the vibrational free energy. In principle, F can be calculated from the partition function $\overline{Z}$, which is-the-sum-of $e^{-\beta E}$ over *all* degrees of freedom, electronic and ionic, and in particular over all configurations of hydrogen and helium on the assumed underlying lattice. (Here $\beta=1/k_BT$ and E is the total energy.) It is useful to introduce the following notation: Let $\langle A\rangle_s$ denote the ensemble average of the variable A for a *static* lattice. The electronic degrees of freedom and the configurational degrees of freedom remain summed over in obtaining $\langle A\rangle_s$. We use the symbol $\langle A\rangle_{s,0}$ to indicate the ensemble average of A for a static lattice in which the configurations summed over are restricted to be *randomly disordered*. We can now write $F_s(T,\Omega,c)$ of Eq. (2) as

$$F_s(T,\Omega,c)=\langle E\rangle_s-T\langle S\rangle_s\,, \tag{3}$$

where S is the entropy. We may also write Eq. (2) as

$$F=\langle E\rangle_{s,0}+(F_s-\langle E\rangle_{s,0})+F_v^0+(F_v-F_v^0)\,, \tag{4}$$

where F_v^0 is the vibrational free energy of a randomly disordered alloy.

We will ignore the last term in Eq. (4), and in Sec. V calculate only F_v^0. The validity of this approximation will be discussed in Sec. VI. The neglect of the term $(F_v-F_v^0)$, and the separation of the static free energy as shown in Eq. (4), are motivated by the fact that those temperatures for which hydrogen-helium alloys actually do form are sufficiently high as to favor such *random* disorder. (This point will be argued more fully in Sec VI.) Thus we expect that at these temperatures $\langle E\rangle_{s,0}$ will be the major contribution to $\langle E\rangle_s$. Note that the second term of Eq. (4) includes the configurational entropy, as well as corrections to the static energy due to correlations of the positions of hydrogen and helium on the lattice.

III STATIC ENERGY

In this section we calculate $\langle E\rangle_{s,0}$ by writing a general expression for $\overline{E}_s$, the static energy of *any* configuration of the ions, and then computing its average over randomly disordered configurations. The approach is to consider an alloy as consisting of hydrogen and helium ions, located on a lattice, and immersed in a responding electron gas of compensating density. The ion-ion, electron-electron, and electron-ion interactions are all given by Coulomb's law. The (divergent) long-wavelength limits of these interactions sum to zero, and are eliminated from the starting Hamiltonian.[7]

We can then write $\overline{E}_s$ as

$$\overline{E}_s=E^{(0)}+E_M+E_b\,. \tag{5}$$

Here $E^{(0)}$ is the energy (per ion) of a homogeneous interacting electron gas (in the presence of a positive, uniform background charge), the Madelung energy E_M is the electrostatic energy of the point ions (in the presence of a uniform negative background charge), and E_b is the energy due to the electrons' response to the *non*uniform component of the total ionic potential V. By treating V as relatively weak, E_b, which is known as the *band-structure* energy, can be calculated by perturbation theory. What we are describing is conventional pseudopotential theory,[8] applied to a system for which the electron-ion interaction is known exactly. This approach has been used extensively in the context of metallic hydrogen,[7,9] and is an important element in the alloy calculation of Ref. 3.

In Eq (5), $E^{(0)}$ is given by

$$E^{(0)}=Z^*(e^2/2a_0)\left[\tfrac{3}{5}(\tfrac{9}{4}\pi)^{2/3}1/r_s^{\,2}-(3/2\pi)(\tfrac{9}{4}\pi)^{1/3}\right.$$
$$\left.\times 1/r_s+(-0.115+0.031\ln r_s)\right], \tag{6}$$

where Z^* is the average ionic charge in units of $e(e>0)$. Since $Z_{\rm He}=2$ and $Z_{\rm H}=1$, $Z^*=cZ_{\rm He}+(1-c)Z_{\rm H}=1+c$. Note that r_s is the usual dimensionless electron spacing parameter

$$\tfrac{4}{3}\pi(r_s a_0)^3=\Omega/Z^*N\,, \tag{7}$$

where a_0 is the first Bohr radius. Since N is the number of ions (in Ω), NZ^* is the corresponding number of electrons. The first two terms in Eq. (6) are the kinetic and exchange energies. The last term is the correlation energy, and is only known approximately. We have used the approximation due to Nozières and Pines,[10] which is expected to be quite satisfactory in the r_s range considered here ($r_s\sim 1$). Note that $E^{(0)}$ is independent of both the configuration of hydrogen and helium ions on the underlying lattice, and of the lattice itself. Since we are interested in temperatures much less than the Fermi temperature

$$T_F=(5.82\times 10^5)/r_s^{\,2}\ ^\circ\mathrm{K}, \tag{8}$$

the electron system[11] is taken to be in its ground state.

The second term in Eq. (5) is the Madelung energy, and is given by[8]

$$E_M=\frac{1}{2\Omega N}\sum_{(i,j)}{}'\sum_{\vec{k}}{}'\frac{4\pi}{k^2}e^{i\vec{k}\cdot(\vec{R}_i-\vec{R}_j)}Z_iZ_j\,, \tag{9}$$

where Z_i is the charge of the ion at site i whose position is given by $\vec{R}_i$. The prime on the sum over i and j denotes the omission of the terms $i=j$. The prime on the $\vec{k}$ sum denotes the omis-

sion of $\vec{k}=0$.

The Madelung energy is generally large and negative, and for a given family of structures often assumes its lowest value for the most symmetric structure.

Using perturbation theory[7,9] E_b can be developed as a series in ascending orders of the electron-ion interaction

$$E_b = E_b^{(2)} + E_b^{(3)} + \cdots , \tag{10}$$

with

$$E_b^{(2)} = \frac{1}{2}\frac{\Omega}{N}\sum_{\vec{k}_1}{}' V(\vec{k}_1)V(-\vec{k}_1)\frac{k_1^2}{4\pi}\left(\frac{1}{\epsilon(\vec{q}_1)}-1\right), \tag{11}$$

and

$$E_b^{(3)} = \frac{1}{3}\frac{\Omega}{N}\sum_{\vec{k}_1}{}'\sum_{\vec{k}_2}{}' V(\vec{k}_1)V(\vec{k}_2)V(-\vec{k}_1-\vec{k}_2) \times\left(\frac{6}{k_F}\frac{1}{\epsilon(\vec{q}_1)}\frac{1}{\epsilon(\vec{q}_2)}\frac{1}{\epsilon(-\vec{q}_1-\vec{q}_2)}\right) 4H_s^{(3)}(-\vec{q}_1,\vec{q}_2), \tag{12}$$

where the primes denote the omission of $\vec{k}_1=0$, $\vec{k}_2=0$, *and* $\vec{k}_1=-\vec{k}_2$. In Eqs. (11) and (12), $V(k)$ is given by

$$V(\vec{k}) = \frac{1}{\Omega}\int d^3r\, e^{-i\vec{k}\cdot\vec{r}}V(\vec{r}), \tag{13}$$

$$V(\vec{r}) = -\sum_i \frac{Z_i}{|\vec{r}-\vec{R}_i|}, \tag{14}$$

where $V(\vec{r})$ is the total ionic potential as seen by the electrons. The restrictions on the sums in Eqs. (11) and (12) follow from the form of the Hamiltonian.[7] The vectors $\vec{q}$ are defined by $\vec{q}=\vec{k}/2k_F$, where the Fermi wave vector k_F is given by the relation

$$k_F^3 = 3\pi^2 Z^*N/\Omega$$

In Eqs. (11) and (12), $\epsilon(\vec{q})$ is the zero-frequency limit of the dielectric function of the homogeneous interacting electron gas, and $H_s^{(3)}$ is given in Eq. (C3) of Ref. 7. We use Hartree atomic units in the equations above (and throughout the rest of the paper).

It is important to note that Eq (11) is an exact result for $E_b^{(2)}$, for $(k_1^2/4\pi)[1/\epsilon(q_1)-1]$ measures the exact linear response of the number density of the homogeneous interacting electron gas to an external potential (in this case the potential due to the ions). In contrast, Eq. (12) is only approximate, as the corresponding second-order response function is not known exactly. The approximation used in Eq. (12) corresponds to treating the electrons as independent particles moving in a self-consistent potential constructed from a Hartree potential and the external potential, *provided* ϵ is taken to be the Lindhard dielectric function.[7,12] In the present calculations we have used the Geldart-Vosko[13] modified form of the Hubbard dielectric function, which includes effects due to exchange and correlation, and yields the correct $\vec{q}\to 0$ limit. It is certainly preferable to use this form (rather than the Lindhard function) in $E_b^{(2)}$, but it is technically inconsistent to use it in $E_b^{(3)}$ as written in Eq. (12). However, these two dielectric functions yield values of $E_b^{(3)}$ within 1% of each other, so that the effect on phase boundaries, which depend on differences of free energies, is inconsequential.

Although the hydrogen-helium alloys have been taken as metallic, the convergence of the perturbation series of Eq. (10) is not dependent on the existence of a metallic state, as discussed in Ref. 3. The point is that the perturbation series should be adequate as long as the one-electron band gaps are less than the bandwidths, which is the case for helium above 10 Mbar. Since actual metallic conduction may only occur[3] in helium at 70 Mbar, this distinction is of considerable importance. (Hydrogen, on the other hand, is expected[14] to be metallic at pressures of a few Mbar.)

Considerable progress[7] has been made in calculating $E_b^{(4)}$, which however, we do not include here. For metallic hydrogen $E_b^{(4)}$ is smaller than $E_b^{(3)}$ by roughly a factor of 10, and it includes the effects of the change in the chemical potential of the electron gas due to the presence of the ions. To correctly calculate $E_b^{(4)}$, one must use finite-temperature perturbation theory, as discussed in Ref. 7.

The terms E_M, $E_b^{(2)}$, and $E_b^{(3)}$ are valid as written for *any* configuration of hydrogen and helium, and contain contributions that depend both on the configuration and on the structure of the underlying lattice. More specifically, since the total potential $V(\vec{r})$ in Eq (14) takes the form of a sum over sites, E_b will contain the following classes of terms:

(i) *Structure-independent* terms, that is, terms independent of configuration and lattice structure. These arise from the terms in $E_b^{(2)}$ and $E_b^{(3)}$ in which all sites coincide.

(ii) *Two-body*, or ion-ion terms. These comprise the remaining terms in $E_b^{(2)}$, and the terms in $E_b^{(3)}$ for which only two site labels coincide.

(iii) *Three-body*, or ion-ion-ion terms. These arise from the terms in $E_b^{(3)}$ in which no site labels coincide.

There are, of course, four-body terms and terms involving more than four ions, but these originate in higher orders of perturbation theory. Recognizing that E_M is also a sum over ion-ion terms, we can group together contributions to $\bar{E}_s$

in Eq. (5) by the classes (i)–(iii) above, and obtain

$$\bar{E}_s = \bar{E}^{(0)} + \frac{1}{2N}\sum_{\vec{R}_i,\vec{R}_j}{}' \phi^{(2)}_{ij}(\vec{R}_i - \vec{R}_j) + \frac{1}{3N}\sum_{\vec{R}_i,\vec{R}_j,\vec{R}_k}{}' \phi^{(3)}_{ijk}(\vec{R}_i - \vec{R}_k, \vec{R}_j - \vec{R}_k) + \cdots . \tag{15}$$

Here the primes denote restrictions forbidding the terms $i=j$ in the two-body term, and the terms $i=k$ and $j=k$ (but not $i=j$) in the three-body term. Note that the two- and three-body potentials depend on density and on the identity of the ions at sites i and j (as well as on the separations $\vec{R}_i - \vec{R}_j$). All terms in $\bar{E}_s$ which are independent of configuration and lattice structure are included in $\bar{E}^{(0)}$. The point about rewriting Eq. (5) as in Eq. (15) is simply that by summing over the electron degrees of freedom (at $T = 0°K$), we have been able to write $\bar{E}_s$ as a sum over (density-dependent) effective pair and three-body potentials, plus a term dependent only on density. This recasting of Eq. (5) is clearly valid for any configuration of hydrogen and helium ions, and is a conceptually useful alternative to Eq. (5).

We now calculate the first term in Eq. (4), the static energy of a *randomly disordered* system:

$$\langle E\rangle_{s,0} = E^{(0)} + \langle E_M\rangle_0 + \langle E_b^{(2)}\rangle_0 + \langle E_b^{(3)}\rangle_0 + \cdots . \tag{16}$$

To do this we must first give the definition of *randomly disordered* To this end we introduce the quantity p_i:

$p_i = 1$ if site i is occupied by a helium ion,

$p_i = 0$ if site i is occupied by a hydrogen ion. (17)

From its definition,[15] we can see that p_i obeys the following relations:

$$(p_i)^n = p_i, \quad n = 2, 3, \ldots, \tag{18a}$$

$$\langle p_i\rangle = c, \tag{18b}$$

where the average in Eq. (18b) is over *all* configurations. Introducing the auxiliary variables d_i:

$$d_i = p_i - c, \tag{19}$$

we have

$$\langle d_i\rangle = 0. \tag{20}$$

Since p_i measures the probability that site i is occupied by a helium ion, d_i measures the deviation of that probability from its average value. In Eq. (9) for E_M, we write Z_i as

$$Z_i = p_i Z_{He} + (1-p_i)Z_H. \tag{21}$$

Thus E_M will clearly involve averages of the type $\langle p_i p_j\rangle$. In terms of these correlation functions we *define* a *randomly disordered* system as one for which the nth-order correlation-function factors according to[16]

$$\langle p_{i_1}p_{i_2},\ldots,p_{i_n}\rangle_0 = \langle p_{i_1}\rangle_0\langle p_{i_2}\rangle_0,\ldots,\langle p_{i_n}\rangle_0 = \langle p_{i_1}\rangle\langle p_{i_2}\rangle,\ldots,\langle p_{i_n}\rangle, \tag{22}$$

where

$$i_1 \neq i_2 \neq, \ldots, \neq i_n .$$

Thus for the two-site correlation function we obtain

$$\langle p_i p_j\rangle_0 = \langle p_i\rangle\langle p_j\rangle = c^2 \quad \text{if } i \neq j,$$
$$\langle p_i p_j\rangle_0 = \langle p_i^2\rangle_0 = \langle p_i\rangle_0 = c \quad \text{if } i = j. \tag{23}$$

Since $i=j$ is excluded from Eq. (9), we immediately have

$$\langle E_M\rangle_0 = \frac{Z^{*2}}{2\Omega N}\sum_{i,j}{}'\sum_{\vec{k}}{}' \frac{4\pi}{k^2} e^{i\vec{k}\cdot(\vec{R}_i - \vec{R}_j)} . \tag{24}$$

The Madelung energy of a randomly disordered alloy is that of a pure metal of ionic charge Z^* (corresponding to the so-called "virtual crystal"),[16] and can be calculated by well-known techniques.[8]

To calculate $\langle E_b\rangle_0$ we must first use Eqs. (13) and (14) to write $V(\vec{k})$ in terms of the variable p_i:

$$V(\vec{k}) = \sum_i e^{-i\vec{k}\cdot\vec{R}_i}\left[p_i\left(\frac{-4\pi Z_{He}}{k^2\Omega}\right) + (1-p_i)\left(\frac{-4\pi Z_H}{k^2\Omega}\right)\right], \tag{25}$$

where $\vec{R}_i$ is again the position vector of site i Introducing d_i via Eq. (19), we obtain

$$V(\vec{k}) = \sum_i e^{-i\vec{k}\cdot\vec{R}_i}[\bar{U}(\vec{k}) + d_i\,\Delta U(\vec{k})], \tag{26}$$

where

$$\bar{U}(\vec{k}) = -[c\,4\pi Z_{He}/k^2\Omega + (1-c)\,4\pi Z_H/k^2\Omega] = -4\pi Z^*/k^2\Omega, \tag{27a}$$

and

$$\Delta U(\vec{k}) = -(4\pi/k^2\Omega)(Z_{He} - Z_H) = -4\pi/k^2\Omega . \tag{27b}$$

From Eqs. (11) and (17), we find

$$\langle E_b^{(2)}\rangle_0 = \frac{1}{2}\frac{\Omega}{N}\sum_{\vec{k}_1}{}' \langle V(\vec{k}_1)V(-\vec{k}_1)\rangle_0 \frac{k_1^2}{4\pi}\left(\frac{1}{\epsilon(\vec{q}_1)} - 1\right) \tag{28}$$

and

$$\langle V(\vec{k}_1)V(-\vec{k}_1)\rangle_0=\sum_i e^{-i\vec{k}_1\cdot\vec{R}_i}\sum_j e^{+i\vec{k}_1\cdot\vec{R}_j}\langle[\overline{U}(\vec{k}_1)+d_i\,\Delta U(\vec{k}_1)][\overline{U}(-\vec{k}_1)+d_j\,\Delta U(-\vec{k}_1)]\rangle_0 . \tag{29}$$

From Eq. (20) we see that the cross terms in Eq. (29) vanish. Using the relation

$$\sum_i e^{-i\vec{K}\cdot\vec{R}_i}=N\delta_{\vec{K},\vec{K}} , \tag{30}$$

where $\vec{K}$ is any vector of the reciprocal lattice, we have

$$\langle V(\vec{k}_1)V(-\vec{k}_1)\rangle_0=N^2\delta_{\vec{k}_1,\vec{K}}\overline{U}(\vec{k}_1)\overline{U}(-\vec{k}_1)+\sum_i\sum_j e^{-i\vec{k}_1\cdot\vec{R}_i}e^{+i\vec{k}_1\cdot\vec{R}_j}[\Delta U(\vec{k}_1)\Delta U(-\vec{k}_1)]\langle d_i d_j\rangle_0 . \tag{31}$$

In the Appendix, we prove the relationship

$$\sum_i\sum_j e^{-i\vec{k}_1\cdot\vec{R}_i}e^{+i\vec{k}_1\cdot\vec{R}_j}\langle d_i d_j\rangle_0=N(c-c^2) . \tag{32}$$

Substituting Eqs. (31) and (32) into (28), and using

$$\lim_{\Omega\to\infty}\frac{1}{\Omega}\sum_{\vec{k}}\to\frac{1}{(2\pi)^3}\int d^3k ,$$

we have the final second-order result

$$\langle E_b^{(2)}\rangle_0=\frac{NZ^{*2}}{2\Omega}\sum_{(\vec{K})}{}'\frac{4\pi}{K^2}\left(\frac{1}{\epsilon(\vec{Q})}-1\right)+\tfrac{1}{2}(Z_{He}-Z_H)^2(c-c^2)\int\frac{d^3k}{(2\pi)^3}\frac{4\pi}{k^2}\left(\frac{1}{\epsilon(q)}-1\right) , \tag{33}$$

where $\vec{Q}=\vec{K}/2k_F$. In Eq (33), the first term is just the second-order band-structure energy[17] of a pure metal of ionic charge Z^*. This virtual-crystal result is *not* correct for a randomly-disordered system, because in Eq. (29) the terms in which the sites i and j coincide must be handled separately. However, it is worth noting that the virtual-crystal result correctly gives the *structural* dependence of $\langle E_b^{(2)}\rangle_0$, since the second term in Eq. (33) is clearly independent of both the lattice structure and the configuration of hydrogen and helium on the lattice.

We have written $\langle E_b^{(2)}\rangle_0$ in a form that is quite similar to other expressions in the literature,[8,18] and have used a rather indirect method to do so. This method, however, avoids much of the confusion that would otherwise arise in the calculation of $\langle E_b^{(3)}\rangle_0$, to which we now turn.

Equation (12) for $\langle E_b^{(3)}\rangle_0$ can be written in the following form[12]:

$$\langle E_b^{(3)}\rangle_0=\frac{\Omega}{N}\sum_{\vec{k}_1}{}'\sum_{\vec{k}_2}{}'\sum_{\vec{k}_3}{}'\langle V(\vec{k}_1)V(\vec{k}_2)V(\vec{k}_3)\rangle_0\chi_2(\vec{q}_1,\vec{q}_2,\vec{q}_3)\delta_{\vec{k}_1+\vec{k}_2+\vec{k}_3,0} , \tag{34}$$

where the function χ_2 is defined by direct comparison of Eqs. (34) and (12). However, we shall never need the explicit expression for χ_2, but only its symmetry properties The form of the function $H_s^{(3)}(-\vec{q}_1,\vec{q}_2)$ in Eq. (12) guarantees that χ_2 is symmetric with respect to the interchange of any two arguments.[7,12] Using Eqs (27) and (30), we have

$$\begin{aligned}\langle V(\vec{k}_1)V(\vec{k}_2)V(\vec{k}_3)\rangle_0=&N^3\delta_{\vec{k}_1,\vec{K}_1}\delta_{\vec{k}_2,\vec{K}_2}\delta_{\vec{k}_3,\vec{K}_3}\overline{U}(\vec{k}_1)\overline{U}(\vec{k}_2)\overline{U}(\vec{k}_3)+N\delta_{\vec{k}_3,\vec{K}_3}\overline{U}(\vec{k}_3)S_2(\vec{k}_1,\vec{k}_2)\Delta U(\vec{k}_1)\Delta U(\vec{k}_2)\\&+N\delta_{\vec{k}_1,\vec{K}_1}\overline{U}(\vec{k}_1)S_2(\vec{k}_2,\vec{k}_3)\Delta U(\vec{k}_2)\Delta U(\vec{k}_3)+N\delta_{\vec{k}_2,\vec{K}_2}\overline{U}(\vec{k}_2)S_2(\vec{k}_3,\vec{k}_1)\Delta U(\vec{k}_3)\Delta U(\vec{k}_1)\\&+S_3(\vec{k}_1,\vec{k}_2,\vec{k}_3)\Delta U(\vec{k}_1)\Delta U(\vec{k}_2)\Delta U(\vec{k}_3),\end{aligned} \tag{35}$$

where we have defined

$$S_2(\vec{k}_1,\vec{k}_2)=\sum_i\sum_j e^{-i\vec{k}_1\cdot\vec{R}_i}e^{-i\vec{k}_2\cdot\vec{R}_j}\langle d_i d_j\rangle_0 \tag{36a}$$

and

$$S_3(\vec{k}_1,\vec{k}_2,\vec{k}_3)=\sum_l\sum_m\sum_n e^{-i\vec{k}_1\cdot\vec{R}_l}e^{-i\vec{k}_2\cdot\vec{R}_m}e^{-i\vec{k}_3\cdot\vec{R}_n}\langle d_l d_m d_n\rangle_0 . \tag{36b}$$

These functions are shown in the Appendix to be

$$S_2(\vec{k}_1,\vec{k}_2)=N\delta_{\vec{k}_1+\vec{k}_2,\vec{K}}(c-c^2) \tag{37a}$$

and

$$S_3(\vec{k}_1,\vec{k}_2,\vec{k}_3)=N\delta_{\vec{k}_1+\vec{k}_2+\vec{k}_3,\,\vec{K}}(c-3c^2+2c^3). \tag{37b}$$

Substituting Eqs. (35)–(37) into Eq. (34), and making use of the symmetry of χ_2 we obtain

$$\begin{aligned}\langle E_b^{(3)}\rangle_0=\frac{\Omega}{N}\sum_{\vec{k}_1}{}'\sum_{\vec{k}_2}{}'\sum_{\vec{k}_3}{}'\big[&N^3\delta_{\vec{k}_1,\vec{K}_1}\delta_{\vec{k}_2,\vec{K}_2}\delta_{\vec{k}_3,\vec{K}_3}\overline{U}(\vec{k}_1)\overline{U}(\vec{k}_2)\overline{U}(\vec{k}_3)\\&+3N\delta_{\vec{k}_3,\vec{K}_3}\overline{U}(\vec{k}_3)N\delta_{\vec{k}_1+\vec{k}_2,\vec{K}}(c-c^2)\Delta U(\vec{k}_1)\Delta U(\vec{k}_2)\\&+N\delta_{\vec{k}_1+\vec{k}_2+\vec{k}_3,\vec{K}}(c-3c^2+2c^3)\Delta U(\vec{k}_1)\Delta U(\vec{k}_2)\Delta U(\vec{k}_3)\big]\chi_2(\vec{q}_1,\vec{q}_2,\vec{q}_3)\delta_{\vec{q}_1+\vec{q}_2,\vec{q}_3,0}\ .\end{aligned} \tag{38}$$

The first term in Eq. (38) is the third-order band-structure energy[17] of the virtual crystal. As before, there are corrections to the virtual-crystal result which have their origin in the coincidence of sites in Eq. (34). However, now the corrections are *structure dependent*. To see this more explicitly, we recast $\langle E_b^{(3)}\rangle_0$ in terms of the function $H_s^{(3)}$ of Eq. (12). By using the symmetry properties of $H_s^{(3)}$ with respect to interchange of arguments (see Ref. 7), we can rewrite Eq. (38) as

$$\begin{aligned}\langle E_b^{(3)}\rangle_0=-\frac{8}{9\pi}\left(\frac{4}{9\pi}\right)^{1/3}r_s\Big[&Z^*\sum_{\vec{Q}_1,\vec{Q}_2}{}'\left(\frac{1}{\vec{Q}_1^2\epsilon(\vec{Q}_1)}\,\frac{1}{\vec{Q}_2^2\epsilon(\vec{Q}_2)}\,\frac{1}{|\vec{Q}_1-\vec{Q}_2|^2\epsilon(\vec{Q}_1-\vec{Q}_2)}\right)H_s^{(3)}(\vec{Q}_1,\vec{Q}_2)\\&+\frac{9}{\pi}(c-c^2)(Z_{He}-Z_H)^2\sum_{\vec{Q}}{}'\int d^3q\left(\frac{1}{q^2\epsilon(\vec{q})}\,\frac{1}{Q^2\epsilon(\vec{Q})}\,\frac{1}{|\vec{q}-\vec{Q}|^2\epsilon(\vec{q}-\vec{Q})}\right)H_s^{(3)}(\vec{q},\vec{Q})\\&+\frac{9}{\pi^2}(c-3c^2+2c^3)(Z_{He}-Z_H)^3\int d^3q_1\int d^3q_2\left(\frac{1}{q_1^2\epsilon(\vec{q}_1)}\,\frac{1}{q_2^2\epsilon(\vec{q}_2)}\right.\\&\left.\times\frac{1}{|\vec{q}_1-\vec{q}_2|^2\epsilon(\vec{q}_1-\vec{q}_2)}\right)H_s^{(3)}(\vec{q}_1,\vec{q}_2)\Big]\ .\end{aligned} \tag{39}$$

As before, $\vec{Q}=\vec{K}/2k_F$, and the prime in the double sum means we omit $\vec{Q}_1=0$, $\vec{Q}_2=0$, and $\vec{Q}_1=\vec{Q}_2$. Since the second term in Eq. (39) involves a sum over the reciprocal lattice, it is clearly structure dependent. Equation (39) is our final result for $\langle E_b^{(3)}\rangle_0$.

The polynomials in c that appear in Eqs (33) and (39) (the basic results of this section) are cumulant polynomials $P_s(c)$, familiar from the theory of electron states in the tight-binding model of randomly disordered alloys.[19] They are defined by the generating function

$$\sum_{s=1}^{\infty}P_s(c)\frac{x^s}{s!}=\ln(1-c+ce^x), \tag{40}$$

which gives,

$$P_1(c)=c,\quad P_2(c)=c-c^2,\quad P_3(c)=c-3c^2+2c^3,\ldots\ . \tag{41}$$

The cumulants arise in both problems for the same reason, namely that the decoupling of the correlation functions, illustrated in Eq. (22), does not hold when two or more sites coincide. This point has been stressed previously in Refs. 20 and 21.

IV LONG- AND SHORT-RANGE ORDER

We now turn to the second term in Eq. (4), namely $F_s-\langle E\rangle_{s,0}$. In Sec. III we have summed over the electronic degrees of freedom to obtain an effective Hamiltonian for the ions [Eq. (15)]. The static partition function (and hence the static free energy) can be obtained by summing e^{-E/k_BT} over all (static) configurations of hydrogen and helium ions on the underlying lattice. To carry out this sum, we need a convenient language with which to describe the configurations. At high temperatures, this is achieved through the use of the correlation functions[22,23] $\langle p_i\rangle$, $\langle p_ip_j\rangle$, $\langle p_ip_jp_k\rangle$, etc., introduced in Sec. III. In general, a helium ion may be more likely to have a hydrogen ion as a nearest neighbor than another helium ion (or vice versa), but the probability (at high temperatures) of a very distant neighbor of the helium ion being another helium ion will depend *only* on the mean concentration of helium The correlation functions $\langle p_ip_j\rangle$, etc., are ideally suited to describe such short-range order,[24,25] for we expect the quantity $\langle p_ip_j\rangle-\langle p_i\rangle\langle p_j\rangle$ to become very small as $\vec{R}_i$ and $\vec{R}_j$ become increasingly well separated. On the other hand, at very low temperatures, and particularly

for stoichiometric compositions, the alloy, if it forms at all, is expected to take up an almost completely ordered state. (For example, if $c=0.5$, the alloy may have the CsCl structure at $T=0\,^{\circ}\mathrm{K}$.) It is clearly inappropriate to attempt to describe this situation with the correlation functions of the type $\langle p_i p_j\rangle$, since $\langle p_i p_j\rangle - \langle p_i\rangle\langle p_j\rangle$ is expected to be infinitely long ranged. Instead, it is convenient to introduce the notion of long-range order,[24,25] which for the example quoted above would be defined by the number of helium ions on "right sites," i.e., the number of He ions on the "helium ion" sublattice. The point is, of course, that this number is 1.00 at $T=0\,^{\circ}\mathrm{K}$. It also approaches rather abruptly the disordered value of 0.5 at the critical temperature (T_c), above which there is *no* long-range order.

Thus, any theory used to calculate $F_s - \langle E\rangle_{s,0}$ must be capable of describing these two very different types of behavior at low and high temperatures. More specifically, at low temperatures we have

$$\lim_{T\to 0}(F_s - \langle E\rangle_{s,0}) = \Delta E, \tag{42}$$

where ΔE is the energy difference between the completely ordered phase and its completely random counterpart. At extremely high temperatures we have[24]

$$\lim_{T\to\infty}(F_s - \langle E\rangle_{s,0}) = -T\langle S\rangle_{s,0} = k_B T[c\ln c + (1-c)\ln(1-c)], \tag{43}$$

where the expression on the right-hand side of Eq. (43) is simply the negative of the entropy of a randomly disordered alloy, weighted by the temperature.

The first step in formulating such a theory is drastically to simplify Eq. (15), and replace it by a nearest-neighbor model, viz.

$$\bar{E}_s = \frac{1}{2N}\sum_{l,m}^{nn}[p_l p_m \Phi_{\mathrm{He-He}} + 2p_l(1-p_m)\Phi_{\mathrm{He-H}} + (1-p_l)(1-p_m)\Phi_{\mathrm{H-H}}], \tag{44}$$

where the sum is over nearest neighbors only, and the pair interactions $\Phi_{\mathrm{He-He}}$, $\Phi_{\mathrm{He-H}}$, and $\Phi_{\mathrm{H-H}}$ will be *chosen* to satisfy Eqs. (42) and (43). Note that since we are computing only the *difference* between energies, the structure-independent term in Eq. (15) may be neglected. The appeal of the simple form in Eq. (44) is that it allows an exact mapping of the problem onto the antiferromagnetic Ising model.[26,27] In addition, the Hamiltonian of Eq. (44) has received a great deal of attention as a model Hamiltonian of an alloy.[24] Since we only need keep terms dependent on configuration, it is easy to show that the pair interactions do not enter separately, but only in the standard combination,

$$v = \Phi_{\mathrm{He-H}} - \tfrac{1}{2}(\Phi_{\mathrm{H-H}} + \Phi_{\mathrm{He-He}}), \tag{45}$$

where v is assumed to be negative.[28]

The energy difference ΔE, as calculated from the Hamiltonian of Eq. (44), is proportional to $-v$, with the proportionality constant depending on the (stoichiometric) composition and the assumed underlying lattice. It is therefore compelling to *choose* v so that the energy difference ΔE between ordered and disordered alloy will be the *true* static energy difference,[29] as calculated by the methods of Sec. III, i.e., with no restrictions to nearest neighbors. Providing our methods of solving the model problem defined by Eq. (44) satisfies the limit in Eq. (43), the resulting function $F(T,\Omega,c) - \langle E\rangle_{s,0}$ will then exhibit both the correct high- and low-temperature behavior.

Such a method of solution of the model problem is provided by the quasichemical approximation.[25,30] The basic idea of the method is to treat clusters of ions as independent units, subject only to the conservation of the number of each type of ion consistent with a *given* long-range order. The probability of cluster having a certain configuration of hydrogen and helium ions is then simply given by the standard Boltzmann factor. If the cluster is chosen to be the whole crystal, the result is exact. For smaller clusters, (in particular for a few atoms), error is introduced because the fact that a given site may be part of two (or more) clusters is ignored in assigning a probability that the site is occupied by (say) a helium atom. Nevertheless, the method does take into account correlation effects in a manner reminiscent of classical liquid theory. The free energy can be written down as a function of temperature and long-range order only, and is to be minimized with respect to the latter. The quasichemical approximation is thus able approximately to describe both long- and short-range order within one context.

The approximation is related to more accurate methods[31] in that it is the first of a hierarchy of approximations[32] which can be substantially developed, although the calculations become extremely involved. It is most readily applied in the following cases (i) $c=0.5$, where the underlying lattice is bcc, and the assumed ordered state is the CsCl structure, (ii) $c=0.75$ (or $c=0.25$), where the underlying lattice is fcc, and the assumed ordered state is the Cu_3Au structure. The method correctly predicts that for $c=0.25$ alloys (ii), the order-disorder transition is of first order,[25] that is, the long-range order drops discontinuously to zero at T_c. It also correctly predicts that the transition for alloys of type (i) is of second order, with the long-range order vanishing continuously at T_c. The existence of short-range order above the transition temperature, and hence a configurational contribution to the specific heat, is also described

by the method,[33] but the details of the experimental specific heats are reproduced only qualitatively.[25,27] When compared to more accurate solutions of the Ising model, the quasichemical method's prediction of T_c is only very roughly correct.[34,35] However, calculation shows that in the very low-temperature region the quantity $F_s - \langle E \rangle_{s,0}$ for $c = 0.5$ agrees fairly well with the low-temperature Ising model series expansion [26]

We have used the quasichemical approximation to calculate $F_s - \langle E \rangle_{s,0}$ for $c = 0.25$, $c = 0.50$, and $c = 0.75$ alloys by using the solutions corresponding to the categories (i) and (ii) above. The parameter v was chosen to yield the true static energy difference ΔE between ordered and disordered phases, as previously described. However, the assumed structures for the ordered and disordered phases in the calculation of ΔE were chosen by criteria to be explained in Secs. V and VI, and were *not* consistent with the structures for which the quasichemical method was evaluated [see (i) and (ii) above]. In addition, the contribution of lattice vibrations and the third-order band-structure energy to ΔE were neglected.[36] These approximations are expected to have a serious effect near T_c, but should make little difference well above or below T_c [37] Since ΔE is a function of r_s, we have constructed an approximate form for $F_s(T, r_s, c) - \langle E \rangle_{s,0}$ which has the correct high- and low-temperature limits. We have *not* assumed that the order-disorder transition occurs at constant volume, for the actual behavior of the alloys is determined in Sec. VI from the Gibbs energy G computed at constant pressure and temperature.

V LATTICE VIBRATIONS

To calculate the contribution to the free energy of the lattice vibrations we first assume that the alloy is randomly disordered The "phonon" spectrum of the random alloy is then calculated by replacing each ion with one of charge Z_{eff} *and mass* M_{eff}. The values of Z_{eff} and M_{eff} are chosen so that the long-wavelength limit of the phonon spectrum is given correctly.[38,39] This is readily seen to require

$$M_{\text{eff}} = M^* = cM_{\text{He}} + (1-c)M_{\text{H}}$$
$$\text{and} \qquad (46)$$
$$Z_{\text{eff}} = Z^* = cZ_{\text{He}} + (1-c)Z_{\text{H}}.$$

The force constants for an alloy of arbitrary configuration are defined (to second order in the electron-ion interaction) from Eq (15):

$$\tilde{\Phi}^{ij}_{\alpha\beta}(\vec{R}_i - \vec{R}_j) = \nabla_\alpha \nabla_\beta \phi^{(2)}_{ij}(\vec{R})\big|_{\vec{R}=\vec{R}_i-\vec{R}_j} \quad (i \neq j). \qquad (47)$$

There are three types of force constants (corresponding to hydrogen-hydrogen, hydrogen-helium, and helium-helium pairs), and from Eqs. (11)–(14) these are

$$\begin{aligned}
\tilde{\Phi}^{\text{H-H}}_{\alpha\beta}(\vec{R}_i - \vec{R}_j) &= Z_H^2 \tilde{\Phi}_{\alpha\beta}(\vec{R}_i - \vec{R}_j), \\
\tilde{\Phi}^{\text{H-He}}_{\alpha\beta}(\vec{R}_i - \vec{R}_j) &= Z_H Z_{\text{He}} \tilde{\Phi}_{\alpha\beta}(\vec{R}_i - \vec{R}_j), \\
\tilde{\Phi}^{\text{He-He}}_{\alpha\beta}(\vec{R}_i - \vec{R}_j) &= Z^2_{\text{He}} \tilde{\Phi}_{\alpha\beta}(\vec{R}_i - \vec{R}_j)
\end{aligned} \qquad (48)$$

Here $\tilde{\Phi}_{\alpha\beta}(\vec{R})$ depends on r_s and may be written

$$\tilde{\Phi}_{\alpha\beta}(\vec{R}) = \nabla_\alpha \nabla_\beta \int \frac{d^3k}{(2\pi)^3} \frac{4\pi}{k^2} \frac{1}{\epsilon(\vec{k})} e^{-i\vec{k}\cdot\vec{R}}. \qquad (49)$$

In terms of force constants, Eq. (46) is equivalent to the replacement of the three types of force constants with a particular type of "average" force constant

The concept of phonons in disordered systems in general, and more specifically the use of average masses and force constants, has met with some success when applied to alloys whose constituent elements have similar masses or force constants.[20,40,41] Clearly the masses and force constants of pure hydrogen and helium are *not* close to each other, but some justification for the replacement of an alloy by an "equivalent" pure system is given by the "virtual-crystal approximation" for the phonon Green's function [20] More specifically, if we start with a pure system of point ions having mass and charge given by Eq. (46), and introduce the difference between the physical charges and masses and the "average" ones as a perturbation,[42] then within this approximation the perturbation causes no change in the phonon Green's function.

We have evaluated the dynamical matrix of the pure system defined by Eq (46) in the adiabatic and harmonic approximations, with the electron-ion interaction taken into account up to second order. This has been repeated for a variety of crystal structures and concentrations, including pure hydrogen and helium. From the phonon frequencies, we calculate[43] the vibrational free energy F_v^0

$$F_v^0 = k_B T \sum_{\vec{q}j}^{\text{BZ}} \ln\{2 \sinh[\tfrac{1}{2}\beta\hbar\omega(\vec{q}j)]\}, \qquad (50)$$

where $\beta = 1/k_B T$, $\omega(\vec{q}j)$ is the phonon frequency of wave number $\vec{q}$ and branch index j, and the sum is over the first Brillouin zone. This zone sum was carried out using the special-point technique[44,45] with a modest number (~10) of special points.

Note that by using the harmonic approximation, the frequencies appearing in Eq. (50) depend on r_s but *not* on temperature In order for them to acquire a temperature dependence, a more sophisti-

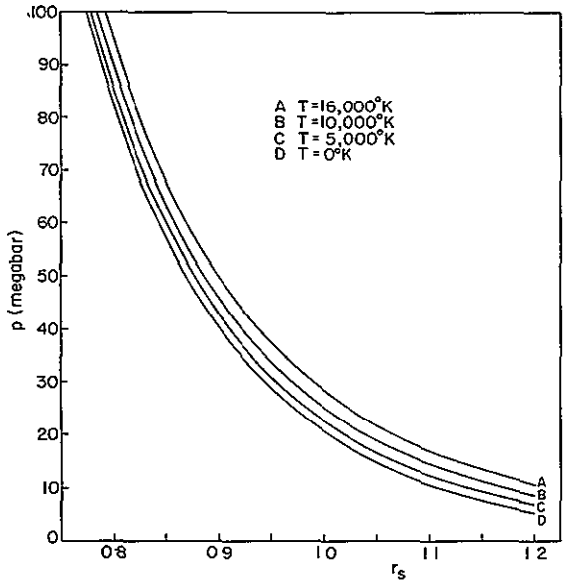


FIG 1 Equation of state of metallic hydrogen.

cated approximation, such as the self-consistent phonon theory,[43] would be needed However, some thermal expansion is included by using the harmonic frequencies, for the contribution of F_v^0 to the pressure is *not* negligible [see Figs. (1) and (2)].

The calculation of the phonon frequencies of the (randomly disordered) alloys and of hydrogen and helium was used as a guide in the choice of the lattice structure chosen for the calculations of Sec III The point is that these Coulomb systems (in the virtual-crystal phonon approximation) are very often harmonically unstable, as discussed by Beck and Straus.[29] (By an instability, we refer to the occurrence of imaginary phonon frequencies) The lattice structures used in the calculations of Sec. III, as described in detail in Sec. VI, were chosen to give *real* frequencies It should be noted, however, that the relationship between instabilities in the virtual crystal approximation and those in the real (randomly disordered) alloy is not clear. We shall assess the effect of our approximate treatment of the phonons on the phase boundaries in Sec. VI.

VI RESULTS AND DISCUSSION

A Choice of lattice structures

Here we discuss the lattice structures chosen to calculate the various contributions to Eq. (4) The static energy *differences* between lattices are in general very small,[7,9] especially when compared to the energy in the phonon system. (However, these energy differences may not be small compared to the *differences* in phonon energies between lattices.) This raises the question of whether these materials can ever solidify in the conventional sense It should be noted that the energy differences are also *not* necessarily small when compared to the difference ΔG of the Gibbs energies between the alloy and the pure hydrogen and helium systems, as Fig 3 illustrates An extensive search in Bravais lattice space for the structure of lowest energy (as carried out in Ref. 9) is not feasible for this problem we limited ourselves to the consideration of the bcc, fcc, and hcp (with variable c/a ratio) lattices in the calculations of $\langle E\rangle_{s,0}$ and F_v^0 in Eq (4) (Simple cubic lattices are harmonically quite unstable for these systems.)

For the randomly disordered alloys (and for pure hydrogen and helium), either fcc or bcc proved to be stable for all Z^* *except* in the range $1.20 \lesssim Z^* \lesssim 1.30$, and the stable lattice was chosen for the calculations. At $Z^* = 1.25$, hcp (with $c/a = 1.7$) was stable, and this structure was therefore chosen in the concentration range near $Z^* = 1.25$. The lattices used to compute $\langle E\rangle_{s,0}$ and F_v^0 are summarized in Table I The absence of an entry for a particular contribution to the energy indicates that the value of that contribution was obtained by interpolation from its values at other concentrations Note that $\langle E_b^{(3)}\rangle_0$ was calculated for fcc, *not* hcp, in the region $1.10 \lesssim Z^* \lesssim 1.35$. It is not expected that this procedure will cause any significant error in the phase separation curves. In addition, the designated phases for $Z^* = 1.00$ and 1.25 are harmonically unstable[46] at low densities (corresponding to pressures of less than 20 and 30 Mbar, respectively) Previous calculations[46]

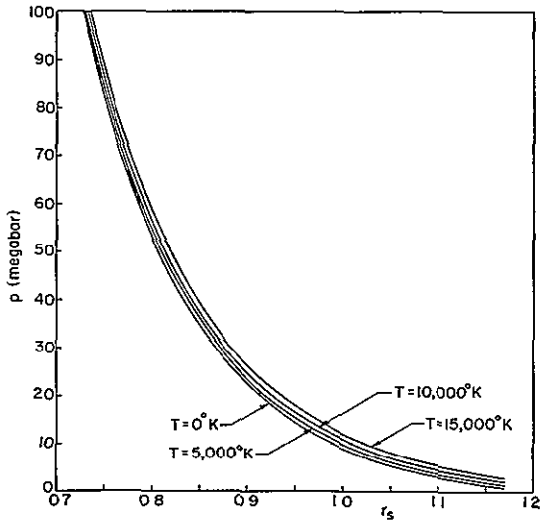


FIG. 2. Equation of state of helium

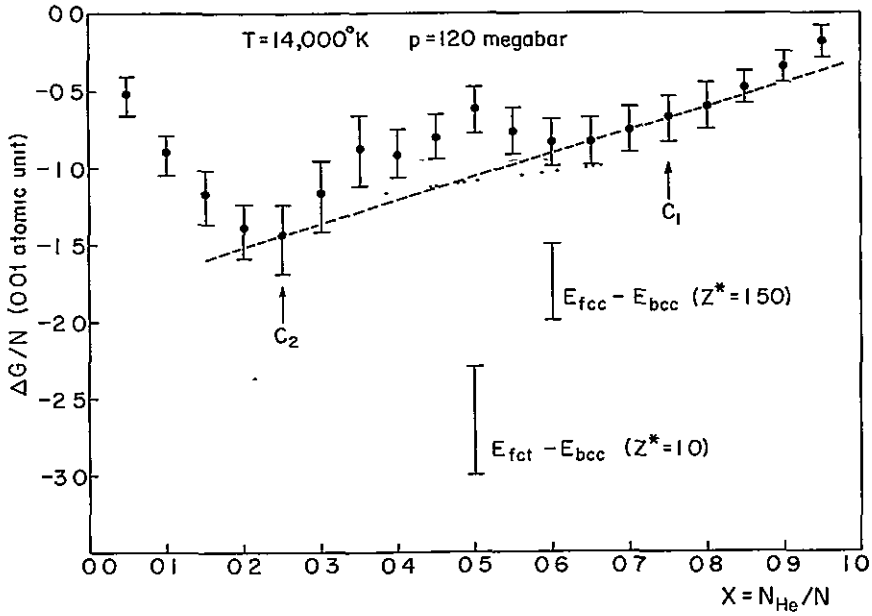


FIG 3 Typical results for ΔG vs c. The dashed line determines the phase separated region ($c_2 \lesssim c \lesssim c_1$) The dotted line shows another possibility for the phase-separated region consistent with the error bars. Typical static energy differences between lattices of randomly-disordered alloys are also shown (fct refers to face-centered tetragonal.)

show that such instabilities will only occur at much higher values of r_s (lower pressures) when the phonon spectrum is calculated in the self-consistent harmonic theory. Thus we adopted the procedure of extrapolating the phonon frequencies to lower density to calculate F_v^0 at low pressure.

We now discuss the lattice structure of the ordered alloys used in calculating $F_s - \langle E \rangle_{s,0}$ by the methods described in Sec. IV The energy difference ΔE between ordered and randomly disordered states was calculated for $c = 0.25$, 0 50, and 0 75 (For pure hydrogen and helium, ΔE, as well as $F_s - \langle E \rangle_{s,0}$, clearly vanishes.) For the alloy of $c = 0$ 50, we have considered two types of lattices (i) Simple tetragonal (st), with a basis of one helium and one hydrogen ion, situated so that when $c/a = 1.0$, this lattice has the CsCl structure [47] (ii) Face-centered tetragonal (fct), with a basis of one helium and one hydrogen ion, situated so that when $c/a = 1.0$, this lattice has the NaCl structure As the fct lattice proved unstable for a wide range of c/a values, we used the simple-tetragonal lattice at $c/a = 1$ 0, where it is stable.

We considered two structures for the ordered $c = 0$ 25 ($c = 0.75$) alloys (i) Simple tetragonal (st) lattice of helium (hydrogen) ions with a four-point basis. The helium (hydrogen) ion resides at the lattice point, and three hydrogen (helium) ions sit at the face centers. If all the ions were identical, the lattice would be face-centered tetragonal. (This is the generalization of the Cu_3Au structure to $c/a \neq 1$ 00.) (ii) Body-centered tetragonal (bct) lattice of helium (hydrogen) ions with a

TABLE I. Lattices used in computations for randomly disordered alloys, and for pure hydrogen and helium.

Z	1.00	1 05	1.10	1.15	1 20	1.25	1 30	1.35	1.40	1 45	1 50
$\langle E_M \rangle_0 + \langle E_b^{(2)} \rangle_0$	fcc	fcc	hcp[a]	hcp	hcp	hcp	hcp	hcp	fcc	bcc	bcc
$\langle E_b^{(3)} \rangle_0$	fcc	fcc	fcc	fcc	fcc	fcc	...	...	...	...	bcc
F_V	fcc	. .	...	...	...	hcp	...	...	...	...	bcc
Z	1.50	1 55	1 60	1 65	1.70	1 75	1 80	1 85	1.90	1.95	2 00
$\langle E_M \rangle_0 + \langle E_b^{(2)} \rangle_0$	bcc	bcc	bcc	bcc	bcc	bcc	bcc	bcc	bcc	bcc	bcc
$\langle E_b^{(3)} \rangle_0$	bcc	...	...	...	...	bcc	...	...	...	...	bcc
F_V	bcc	...	...	...	...	bcc	...	...	...	...	bcc

[a] hcp refers to the hexagonal close-packed lattice with $c/a = 1.70$.

TABLE II. Order-disorder critical temperature T_c (in units of 10^3 °K) as a function of pressure p (in units of Mbar) (pressures are approximate only).

$c=0.250$		$c=0\ 500$		$c=0\ 750$	
T_c	p	T_c	p	T_c	p
5.06	2.0	3.45	3 0	0.79	2 5
4.82	4 5	4 40	7.0	1 21	7.0
4.65	7 5	5 63	13.5	1.70	14.5
4 45	13.0	6.67	21.0	2.16	23.5
4.40	20 5	7.92	34.0	2.73	39.0
4.37	31.0	9 19	50.0	3.07	49.5
5.35	47 5	10 05	63.5	3 46	64 0
5 94	59.5	11 03	80 5	3.89	82.5
6 61	74 5	12.10	102 5	4.21	98 5
7 35	94 5	12 68	116 0	4.47	111 5
7 90	111 0	13 31	131.5	4 75	127 0
8 33	125.0				

four-point basis The helium (hydrogen) ion resides at the lattice point, and three hydrogen (helium) ions sit at the face centers and edge midpoints If all the ions were identical, the lattice would be simple tetragonal, with half the original lattice constant.[48] Of these two structures, the st lattice with $c/a=0.7$ proved, for $c=0.75$, to have the lowest static energy (to second order in the electron-ion interaction). Since this structure is harmonically stable, the difference between its static energy and that of the corresponding disordered alloy of Table I (bcc) was set equal to ΔE, as required in the application of the quasichemical theory of Sec IV. For $c=0.25$, *neither* of the two structures are harmonically stable (over a wide range of c/a values) This may be a *dynamic* indication[49] of immiscibility at $T=0$ °K, or alternatively it may indicate that these structures are energetically quite far from the structure an ordered alloy actually assumes Of these two structures, the bct lattice with $c/a=1\ 0$ has the lowest static energy for $r_s \geq 0.920$ ($p \leq 28\ 9$ Mbar at $T=0$ °K), but the st lattice with $c/a=1.0$ has the lowest energy for $r_s<0\ 920$ The static energy differences between these structures and the corresponding random alloy (hcp) were used for ΔE in the calculation of Sec. IV. In Table II we present the critical temperature T_c as a function of pressure for the order-disorder transition, as calculated from Sec IV.

In order to determine how serious an error was made in neglecting lattice vibrations in the computation of ΔE, we computed F_v for the CsCl-structure alloy at $T=0$ °K and $r_s=0\ 99$. The result is within 7% (0 001 a.u per ion) of the corresponding random alloy (bcc) result The difference is small, even on the scale of ΔG This also shows that our neglect of the term $(F_v-F_v^0)$ in Eq. (4) is quite justified.

B Phase separation

The equations of state of pure hydrogen and helium are presented in Figs. 1 and 2. For hydrogen, at $T=0$ °K, they agree well with Caron's results (see Ref 29).

Under conditions of constant temperature and pressure, the free energy to be minimized is the Gibbs free energy G:

$$G(p,T,c)=F(p,T,c)+p\Omega_0, \tag{51}$$

where p is the pressure and Ω_0 the volume per ion Stability of mixed phases is determined by ΔG

$$\Delta G=G(p,T,c)-[cG(p,T,1)+(1-c)G(p,T,0)] \tag{52}$$

Here $c=1$ refers to pure helium and $c=0$ to pure hydrogen In order for there to be any mixing, ΔG must be negative. A miscibility gap occurs when ΔG is negative but the system can lower its

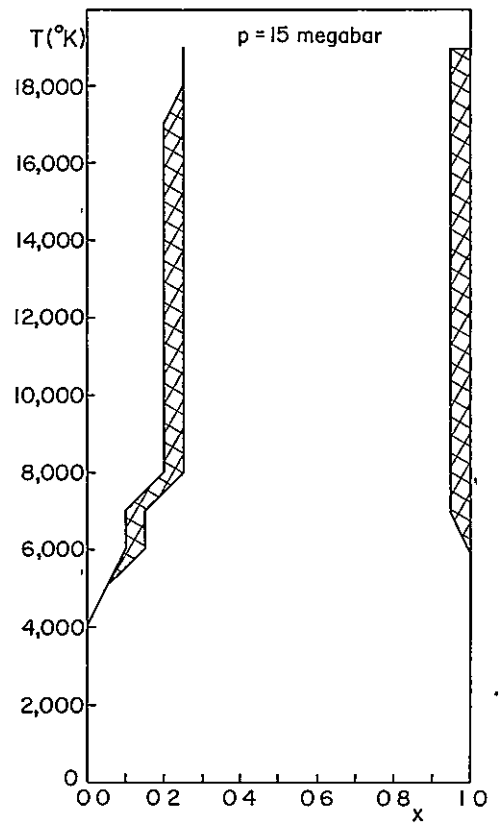


FIG 4 Phase separation curve at 15 Mbar x is the relative concentration (by number) of helium The cross-hatched regions show the uncertainty in the phase separation boundary.

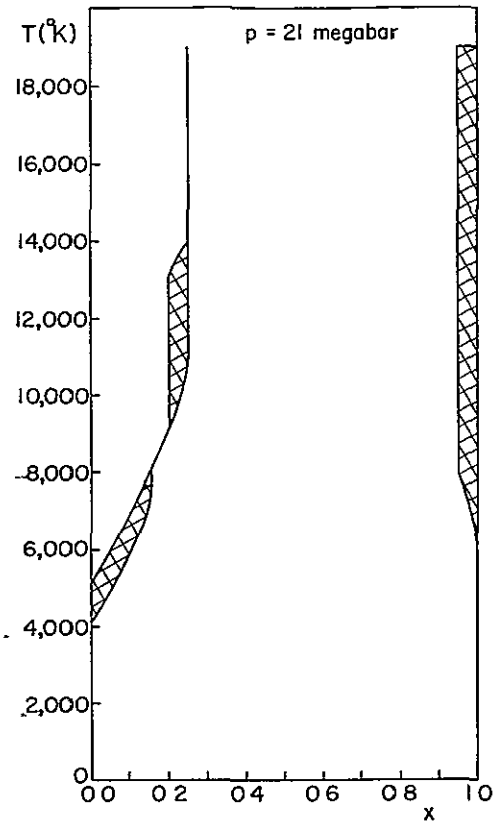


FIG. 5 Phase separation curve at 21 Mbar.

Gibbs energy by separating into a helium-rich phase and a hydrogen-rich phase [50] This is demonstrated in Fig. 3, where we present typical results for $\Delta G(p,T,c)$ at fixed p and T At any concentration between $c=c_1$ and $c=c_2$ the system can lower its Gibbs energy by separating into a helium-rich phase at $c=c_1$ and a hydrogen-rich phase at $c=c_2$, with the relative amounts of the two phases being given by number conservation For such a partially separated system, the Gibbs function is given by the dashed line in Fig. 3. The error bars in Fig. 3 refer to the estimated computational error,[51] *not* the error due to the various physical approximations made We have also shown typical static energy differences (to second order) between lattice structures in Fig. 3, from which the sensitivity of the phase boundaries to lattice structure can be estimated.

The phase separation curves themselves are presented in Figs. 4–8. Note that the temperatures for which mixing occurs are generally well above the order-disorder transition temperatures listed in Table II. Thus, as we have mentioned, the details of this transition are not very important in the calculation of the phase boundaries. The uncertainties in ΔG are the cause of the uncertainties in the phase boundaries, indicated by the cross-hatched regions. The most striking features of the results are (i) the persistence of a large miscibility gap at the highest temperatures and pressures, and (ii) the large temperatures necessary for *any* mixing to occur.[52] The occurrence of large mixing temperatures is not dependent upon the approximations we have used to take into account short-range order and lattice vibrations, although the precise values of the mixing temperatures clearly are. The prediction of *complete* phase separation[50] at temperatures below some temperature T_m reflects the large positive values of ΔG for the static alloys ($\Delta G \sim k_B T_m$). In contrast, the large miscibility gap is primarily due to the "pinning" of the phase boundary near $c=0.25$. This is caused by the exceptionally *low* values of ΔG for $c=0.25$ (see Fig 3) at high temperatures, an effect for which the lattice dynamics is entirely responsible

The relatively low phonon frequencies predicted by the virtual crystal approximation for the $c=0.25$ randomly *disordered* alloys should be compared with the imaginary frequencies found for the $c=0.25$ *ordered* alloys In both cases the alloy ex-

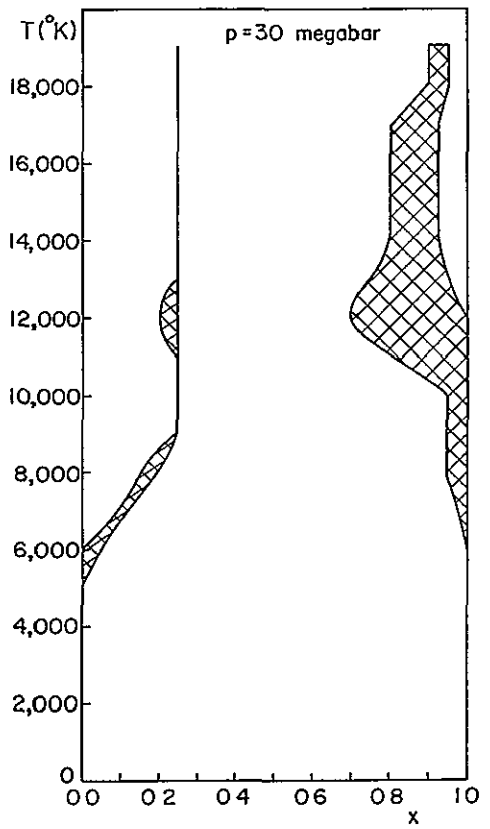


FIG 6. Phase separation curve at 30 Mbar.

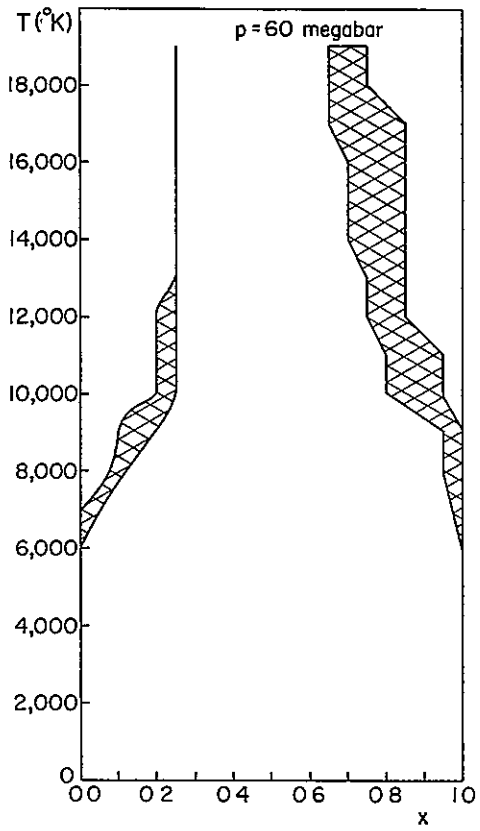


FIG. 7 Phase separation curve at 60 Mbar

hibits phonons whose frequencies squared are low. This results, in one case, in a true instability, and in the other case the low energy and high entropy resulting from these low frequencies greatly favor mixing. In respect of the $c = 0.25$ alloys, it appears that the treatment of the lattice dynamics may be quite crucial. A more correct treatment of the disordered alloy (within the harmonic theory), and the application on the *temperature-dependent* self-consistent (harmonic) phonon theory for example, may produce qualitative differences in the phase boundaries. One such difference might be the disappearance of the miscibility gap at temperatures below 19 000 °K.

In conclusion, the calculation predicts that until the temperature has reached a fairly high value, which will certainly depend upon pressure, there is essentially *complete* phase separation[50] in solid alloys of metallic hydrogen and helium. This may be regarded as a fairly firm result, since it is not dependent in any crucial way upon the approximations used to compute ΔG. If hydrogen and helium are solid in some region of the interior of Jupiter, these conclusions have a direct bearing on any phase separation model of energy emission.

We also predict a large miscibility gap that persists to $T = 19\,000$ °K and $p = 90$ Mbar. However, this prediction depends upon the approximations we have used in treating the lattice dynamics of the alloys, and might well be substantially modified by a more detailed treatment of the phonon spectrum. The third-order terms in the band-structure energy have little effect, tending to raise ΔG by only a small amount. Thus the approximate response function used in $\langle E_b^{(3)} \rangle_0$, as well as the neglect of $\langle E_b^{(4)} \rangle_0$, is not expected to have any important effect on the phase boundaries. The same is true of the use of the quasichemical approximation

ACKNOWLEDGMENTS

The authors wish gratefully to acknowledge very useful and stimulating discussions with D. J Stevenson One of us (H B.) wishes to acknowledge the support of the Swiss National Foundation

APPENDIX

The calculation of $\langle E_b^{(2)} \rangle_0$ and $\langle E_b^{(3)} \rangle_0$ in Sec III requires the evaluation of the following averages.

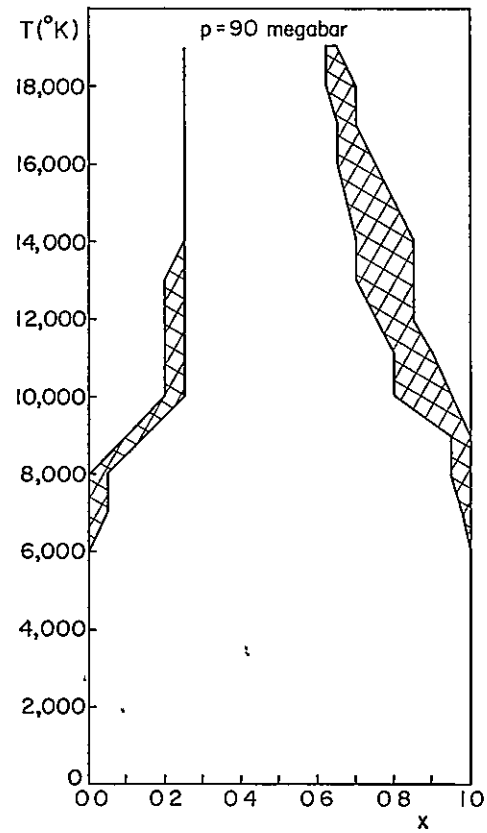


FIG. 8. Phase separation curve at 90 Mbar.

$$S_2(\vec{k}_1,\vec{k}_2)\equiv\sum_i\sum_j e^{-i\vec{k}_1\cdot\vec{R}_i}e^{-i\vec{k}_2\cdot\vec{R}_j}\langle d_i d_j\rangle_0 \quad \text{(A1)}$$

and

$$S_3(\vec{k}_1,\vec{k}_2,\vec{k}_3)\equiv\sum_l\sum_m\sum_n e^{-i\vec{k}_1\cdot\vec{R}_l}e^{-i\vec{k}_2\cdot\vec{R}_m}\times e^{-i\vec{k}_3\cdot\vec{R}_n}\langle d_l d_m d_n\rangle_0 \quad \text{(A2)}$$

We will freely make use of the definitions and properties of the variables p_i and d_i as presented in Sec III Expressing d_i in terms of p_i, we have

$$\langle d_i d_j\rangle_0=\langle(p_i-c)(p_j-c)\rangle_0=\langle p_i p_j\rangle_0-c^2=\delta_{ij}(c-c^2) \quad \text{(A3)}$$

Similarly,

$$\langle d_l d_m d_n\rangle_0=\langle(p_l-c)(p_m-c)(p_n-c)\rangle_0=\langle p_l p_m p_n\rangle_0-c\langle p_m p_n\rangle_0-c\langle p_l p_n\rangle_0-c\langle p_l p_m\rangle_0+3c^3-c^3. \quad \text{(A4)}$$

Note that if $l\neq m\neq n$ in Eq (A4), Eq. (22) guarantees that the average will vanish. If only two of the sites are equal, we use Eq (18) and again the average vanishes Thus

$$\langle d_l d_m d_n\rangle_0=\delta_{l,m}\delta_{m,n}(c-3c^2+2c^3) \quad \text{(A5)}$$

Substituting Eqs. (A3) and (A5) into (A1) and (A2), and using Eq (30),

$$S_2(\vec{k}_1,\vec{k}_2)=N\delta_{\vec{k}_1+\vec{k}_2,\vec{K}}(c-c^2) \quad \text{(A6)}$$

and

$$S_3(\vec{k}_1,\vec{k}_2,\vec{k}_3)=N\delta_{\vec{k}_1+\vec{k}_2+\vec{k}_3,\vec{K}}(c-3c^2+2c^3), \quad \text{(A7)}$$

where $\vec{K}$ is any vector of the reciprocal lattice

*Work supported by the National Aeronautics and Space Administration under Grant No. NGR-33-010-188
†Also supported by the Swiss National Foundation.

[1]W B Hubbard and R. Smoluchowski, Space Sci. Rev. 14, 599 (1973)
[2]E. E. Salpeter, Astrophys. J 181, L83 (1973).
[3]D. J. Stevenson, Phys. Rev. B 12, 3999 (1975).
[4]D J Stevenson and N W Ashcroft, Phys Rev A 9, 782 (1974).
[5]For instance, two different methods of calculating the melting temperature of hydrogen predict temperatures different by a factor of 4 at about 40 Mbar See Ref 4
[6]M Hansen, *Constitution of Binary Alloys*, 2nd ed. (McGraw-Hill, New York, 1958).
[7]J Hammerberg and N W Ashcroft, Phys Rev. B 9, 409 (1974).
[8]W. A. Harrison, *Pseudopotentials in the Theory of Metals* (Benjamin, New York, 1969).
[9]E G Brovman, Yu Kagan, and A Kholas, Zh Eksp. Teor Fiz 61, 2429 (1971) [Sov Phys.-JETP 34, 1300 (1972)]
[10]P Nozières and D. Pines, Nuovo Cimento 9, 470 (1958); P. Nozières and D Pines, Phys Rev 111, 442 (1958); P Vashishta and K S Singwi, Phys. Rev. B 6, 875 (1972)
[11]The first low-temperature correction to the free energy F of the free-electron gas can be shown to contribute negligibly to the phase separation boundaries.
[12]P Lloyd and C A. Sholl, J. Phys. C 1, 1620 (1968).
[13]D. J W. Geldart and S H Vosko, Can J Phys. 44, 2137 (1966).
[14]A. K. MacMahan, H Beck, and J. Krumhansl, Phys. Rev A 9, 1852 (1974)
[15]None of our final results depends upon the definition of p_i in terms of helium It might just as well have been defined in terms of hydrogen.
[16]D Stroud and N W Ashcroft, J Phys F 1, 113 (1971).
[17]Since the calculation is valid for all c, the $c=0$ (or $c=1$) limits of $\langle E_M\rangle_0$, $\langle E_b^{(2)}\rangle_0$, and $\langle E_b^{(3)}\rangle_0$ recover the pure crystal result.
[18]V. Heine and D. Weaire, in *Solid State Physics*, edited by H Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1970), Vol. 24
[19]F Yonezawa and T Matsubara, Prog Theor. Phys. 35, 357 (1966); R. Kubo, J Phys. Soc. Jpn. 17, 1100 (1962)
[20]R J Elliott, J. A. Krumhansl, and P L Leath, Rev Mod Phys. 46, 465 (1974).
[21]F. Yonezawa and K. Morigaki, Prog. Theor. Phys. Suppl. 53, 1 (1973).
[22]C G. Shirley and S Wilkins, Phys Rev. B 6, 1252 (1972).
[23]B. Taggart and R. A. Tahir-kheli, Phys. Rev 46, 1690 (1971), R A. Tahir-kheli, *ibid* 188, 1142 (1969)
[24]T Muto and Y Takagi, in *Solid State Physics*, edited by F Seitz and D. Turnbull (Academic, New York, 1955), Vol. 1.
[25]L Guttman, in *Solid State Physics*, edited by F. Seitz and D Turnbull (Academic, New York, 1956), Vol. 3
[26]C. Domb, in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M S Green (Academic, New York, 1974), Vol. 3
[27]A Bienenstock and J. Lewis, Phys Rev 160, 393 (1967).
[28]If Eq (44) is taken to define the complete Hamiltonian of the system, then positive v implies the occurrence of phase separation at $T=0$ °K (and zero pressure). Since we are using the Hamiltonian of Eq (44) only to describe the free energy involved in the ordering of an assumed alloy, it is necessary to take v as being negative.
[29]If only two-body interactions are kept in Eq (15), then such a choice of v is exact within mean-field theory (known as the Bragg-Williams approximation in the alloy context). Since mean-field theory is expected to be valid for very-long-range interactions [H E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Oxford U P , London, 1971), p 91], and since the pair interactions in these alloys have a range of at least 10 neighbors [H Beck and D Straus,

Helv Phys. Act 48, 655 (1975), L G. Caron, Phys. Rev B 9, 5025 (1974)], mean-field theory should be a reasonable approximation.

[30]R H Fowler and E. A Guggenheim, Proc. R Soc A 174, 189 (1940), C N Yang and Y. Y Li, Chim. J. Phys. 7, 59 (1947); Y. Y. Li, J. Chem Phys 17, 447 (1949)

[31]R Kikuchi, Phys Rev 81, 988 (1951), M Kurata and R Kikuchi, J. Chem Phys 21, 434 (1953).

[32]Each higher approximation consists of taking a larger group of ions as the basic cluster

[33]This is not true of mean-field theory

[34]C Domb, Adv. Phys 9, 245 (1960)

[35]D M Burley, in *Phase Transitions and Critical Phenomena*, edited by C Domb and M. S. Green (Academic, New York, 1972), Vol. 2

[36]A more subtle assumption made is that at every concentration, there is only one ordered phase For examples of other possibilities, see N S Golosov and A. M. Tolstik, J Phys Chem Solids 36, 899, 903 (1975), N S Golosov, A. M Tolstik, and L Ya. Pudan, *ibid* 37, 273 (1976); N S Golosov and A M Tolstik, *ibid* 37, 279 (1976)

[37]One should note that the quasichemical approximation itself is least accurate in the critical region.

[38]The long-wavelength limit of the vibrational spectrum will yield a compressibility which agrees with that calculated from the static energy (up to second order in the electron-ion interaction) *only* if some terms of *third* and *fourth* order in the electron-ion interaction are included in the dynamical matrix [C J. Pethick, Phys Rev B 2, 1789 (1970] Since we only keep second-order terms in the dynamical matrix, the replacement of Eq (46) is not exact, even in the long-wavelength limit The resulting error in the compressibility is of order 10% [E Stoll, P Meier, and T. Schneider, Nuovo Cimento B 23, 90 (1974)] This discrepancy is also present in the case of pure hydrogen and helium

[39]H Beck and D Straus (see Ref 29) define the "average mass" incorrectly However, since the mass of a pure system enters the dynamical matrix only as a multiplicative prefactor, none of their results are affected

[40]W A Kamitakahara and B N Brockhouse, Phys Rev B 10, 1200 (1974) Note that the "average" force constants used in this reference do not correspond to the average defined by Eqs (46) and (48)

[41]E. C. Svensson, B N. Brockhouse, and J M. Rose, Solid State Commun 3, 245 (1965); S C. Ng and B N Brockhouse, *ibid* 5, 79 (1967)

[42]This procedure is necessary to keep r_s constant

[43]P Choquard, *The Anharmonic Crystal* (Benjamin, New York, 1971)

[44]A Baldereschi, Phys Rev B 7, 5212 (1973), D. J Chadi and M L Cohen, *ibid* 8, 5747 (1973).

[45]D M Straus and N W Ashcroft, Phys Rev B 14, 448 (1976)

[46]The type of "Kohn anomaly" instability shown by these two substances is discussed in Beck and Straus (see Ref 29) The self-consistent phonon theory might well stabilize these substances at low density

[47]In the context of cubic lattices c/a is the ratio of the distance between equivalent planes to the distance between equivalent ions in a plane

[48]F Dyson, Ann Phys (N Y) 63, 1 (1971)

[49]Instabilities occur at long wavelength for both structures.

[50]We describe the criterion for *global* instability. The expected exponentially small limiting solubilities are not considered here

[51]The large error bars at higher temperatures and low concentrations of helium are largely due to the (estimated) error in using only a few special (hcp) points to calculate F_v^0 for $c = 0.25$ The fractional error $\Delta F_v^0/F_v^0$ is usually less than 5%, but F_v^0 can be large, on the scale of ΔG (F_v^0 for $c = 0.25$ in Fig. 3 is of order 0 1 a u per ion)

[52]These features should be contrasted with the phase separation curves of Ref. 3.

Self-Consistent Structure of Metallic Hydrogen*

David M. Straus† and N. W. Ashcroft
Laboratory of Atomic and Solid State Physics and Materials Science Center, Cornell University, Ithaca, New York 14853
(Received 23 November 1976)

A calculation is presented of the total energy of metallic hydrogen for a family of face-centered tetragonal lattices carried out within the self-consistent phonon approximation. The energy of proton motion is large and proper inclusion of proton dynamics alters the *structural* dependence of the total energy, causing isotropic lattices to become favored. For the dynamic lattice the structural dependence of terms of third and higher order in the electron-proton interaction is greatly reduced from static lattice equivalents.

Perturbation theory has been moderately successful in accounting for the structural dependence of the *static* energy in many simple crystalline metals.[1,2] In this method, the structural energy is obtained by expansion in orders of the effective conduction-electron–ion interaction (or pseudopotential), the expansion usually being truncated at the lowest term and resulting in what is referred to as the second-order band-structure energy For perfect lattices, this term reduces to a relatively simple sum over the sites of the reciprocal lattice.

In the case of metallic hydrogen, the electron-ion (electron-proton or electron-deuteron) interaction is exactly known, and it is partly for this reason that this system has attracted theoretical attention [3-7] Within the static-lattice approximation, perturbation theory for the structural energy has been carried through to fourth order,[7] and extensive scans of "Bravais lattice space" have been carried out in an attempt to determine, at zero pressure, the structures with lowest static energy.[3] In the latter calculations (which were at third order), Brovman *et al.*[3] concluded that static metallic hydrogen would take up structures which are so highly anisotropic that near the zero-pressure metastable density they would become "liquidlike" in certain crystal directions upon inclusion of the proton dynamics.

Since the ionic mass in metallic hydrogen is small, one expects on quite general grounds that the ionic degrees of freedom can play a rather significant role in determining the structure with lowest overall energy. It is known[3,7] that energy differences between different structures are small—much smaller, for example, than the estimate of the energy bound up in the zero-point motion of the protons. Evidently, what is required is a calculation of structural energies carried out self-consistently for various lattices disturbed by the presence of phonons. The purpose of this Letter is to report on the outcome of such an investigation: We have completed a series of calculations within the self-consistent harmonic phonon approximation[8,9] (SCHA) for a representative family of face-centered tetragonal (fct) Bravais lattices in their ground states at a density[10] of $r_s = 1.36$ [with $\frac{4}{3}\pi(r_s a_0)^3 = n^{-1}$, n being the electron density N/Ω]. Two important results emerge: First, the inclusion of ion dynamics radically alters the structural dependence of the energy so that, in the family which we consider, it is the *isotropic* lattice (fcc) that is ultimately favored. Second, by the inclusion of ion dynamics in the perturbation theory, the structural sensitivity of the terms *higher* than second order is greatly reduced from that appropriate to the static theory.

The arguments go as follows: To second order in the electron-proton interaction, the total *ground-state energy* per proton in the self-con-

sistent harmonic approximation can be written[11]

$$E(r_s)=\frac{1}{4N}\sum_{\vec{q},j}^{BZ}\hbar\omega(\vec{q},j)+\frac{1}{2}\sum_{\vec{X}\neq 0}\Phi(\vec{X})+(\text{terms independent of structure}). \tag{1}$$

Here the sum of frequencies $\omega(\vec{q},j)$ of polarization j is taken over the first Brillouin zone (BZ), and

$$\Phi(\vec{X})=\int\frac{d^3k}{(2\pi)^3}\frac{4\pi}{k^2\epsilon(k)}\exp[-\tfrac{1}{2}k_\alpha k_\beta\lambda_{\alpha\beta}(\vec{X})]\exp(i\vec{k}\cdot\vec{X}), \tag{2}$$

where

$$\lambda_{\alpha\beta}(\vec{X})=2[\langle u_\alpha(\vec{X})u_\beta(\vec{X})\rangle-\langle u_\alpha(\vec{X})u_\beta(0)\rangle]=\frac{\hbar}{MN}\sum_{\vec{q},j}^{BZ}(1-\cos\vec{q}\cdot\vec{X})e_\alpha(\vec{q},j)e_\beta(\vec{q},j)\omega^{-1}(\vec{q},j), \tag{3}$$

with the brackets indicating an average over *harmonic* states. In Eq. (2), $\epsilon(k)$ is the dielectric function of the interacting electron gas taken, as is customary, in its static limit. The small ionic displacements $\vec{u}(\vec{X})$ are defined by $\vec{u}(\vec{X})=\vec{R}-\vec{X}$, where $\vec{R}$ is the instantaneous position of the ion, and $\vec{X}$ the lattice site to which it is attached. Notice that the first term in (1) is the kinetic energy of the ionic system whereas the second is the potential energy averaged over the ions, motion. To carry out this averaging, we require both the frequencies $\omega(\vec{q},j)$ and the polarization vectors $\vec{e}(\vec{q},j)$ of the self-consistent phonons; and these are given by the solution of

$$M\omega^2(\vec{q},j)e_\alpha(\vec{q},j)=\left\{\sum_{\vec{X}\neq 0}(\cos\vec{q}\cdot X-1)\int\frac{dk}{(2\pi)^3}\frac{4\pi}{\epsilon(k)k^2}k_\alpha k_\beta\exp[-\tfrac{1}{2}k_\mu k_\nu\lambda_{\mu\nu}(\vec{X})]\exp(i\vec{k}\cdot\vec{X})\right\}e_\beta(\vec{q},j). \tag{4}$$

Evidently, the *static* energy can be formally recovered by setting $\lambda=0$ in Eqs. (1)–(3), and by omitting the phonon kinetic energy in Eq. (1). The harmonic approximation, on the other hand, can be obtained by expanding in powers of λ and retaining the terms linear in λ. In metallic hydrogen however, the root-mean-square proton displacement is substantial,[12] and such an expansion (implicit in Ref. 3) is open to question. The second-order static energies[13,14] (to which, in the harmonic approximation, the phonon energies are simply added) are shown in Fig. 1, plotted against c/a for the fct system (solid line). Note that there is noticeable structure in the curve not found, for example, in an ordinary simple metal (e.g.,[15] Al). In agreement with Ref. 3, we find a structure with $c/a<1$ to have the lowest static energy. However, when we compute the dynamic energy self-consistently, the situation changes markedly. It is important to note that the solutions of (4) do not always admit *real* frequencies: The arrows in Fig. 1 indicate three such lattices; the dashed line gives the *total* energy[16,17] [Eq. (1)] for the c/a values for which Eq. (4) *can* be solved. The reason for the apparent failure of the SCHA is simply that, for certain values of the parameter c/a, the small-oscillations problem is not well defined. For example, lattices corresponding to c/a values lying in the range $0.5<c/a<0.7$ are associated with a portion of the static-energy curve (Fig. 1) that is removed from a local minimum and for which the second derivative (with respect to c/a) is negative. In these lattices, the existence of *stable* small oscillations of the protons cannot be presumed, and the occurrence of imaginary frequencies in the SCHA is an indication that they do not. For values of c/a near 1.5,

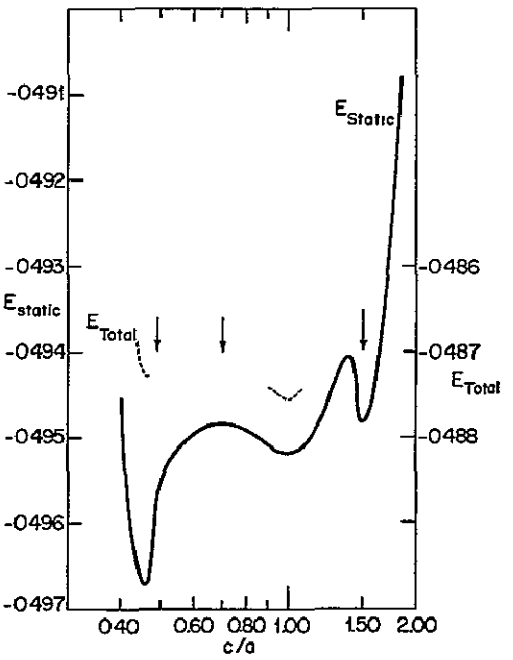


FIG. 1. Static energy and total self-consistent energy for fct metallic hydrogen (at $r_s=1.36$ and $T=0$°K) as a function of c/a (all energies are in hartree atomic units). Total (right-hand scale) is given by the dashed line. Arrows refer to particular values of c/a for which the crystal is unstable.

the absence of stable oscillations is already suggested by the results of the harmonic approximation, for which imaginary frequencies are found everywhere in the BZ. Although there is a minimum in the static energy near $c/a = 1.5$ (Fig. 1), the SCHA can still fail because in the wider Bravais lattice space referred to earlier this point can be situated at a saddle on the energy surface, in contrast to the regions corresponding to the dashed curves which evidently reflect local minima (as required for stability).

The total energy is minimized at $c/a = 1$ corresponding to the fcc structure, which is the most symmetric of the class considered. Since the sharp variations of static-lattice energy found in Fig. 1 and in the plots of Ref. 3 occur over values of c/a comparable to the ratio of $\{\langle \vec{u}^2 \rangle\}^{1/2}$ to a nearest-neighbor distance, it is not unreasonable to expect similar behavior for other families of Bravais lattices such as those investigated by Brovman, Kagan, and Kholas.[3] Evidently, we may conclude that in the metallic phase of hydrogen, lattice dynamical effects completely alter the structural dependence of the energy: In a self-consistent calculation, it is isotropic lattices that are favored. (Indeed, it is worth noting that *none* of the structures corresponding to the minima of the static energy in Fig. 1 is stable in the simple-harmonic approximation.) Finally, the energy of motion, defined by $E - E_{\text{static}}$, is[18] 0.0076 hartree units per proton for the fcc structure. This is a substantial fraction of the zero-pressure binding energy[3,7] which, depending on estimates of electron-gas correlation energy, is in the range 0.02 to 0.03 hartrees per proton.

We now come to the structural dependence of terms in the energy of third and higher order in the electron-ion interaction, which have been omitted from (1). In the SCHA the *total* second-order band-structure energy can be written

$$E_b^{(2)} = \frac{1}{2\Omega} \sum_{\vec{k} \neq 0} S(\vec{k}) \frac{4\pi}{k^2} \left[\frac{1}{\epsilon(k)} - 1 \right], \tag{5}$$

where the static structure factor $S(\vec{k})$ is given by[17]

$$S(\vec{k}) = \sum_{\vec{X}} e^{i\vec{k}\cdot\vec{X}} \exp[-\tfrac{1}{2} k_\alpha k_\beta \lambda_{\alpha\beta}(\vec{X})]. \tag{6}$$

This function is plotted in Fig. 2 for fcc metallic hydrogen ($r_s = 1.36$) with $\vec{k}$ along the [100] direction. The large weight between peaks (and the correspondingly sharp reduction in the strength of the Bragg peaks themselves) can be traced

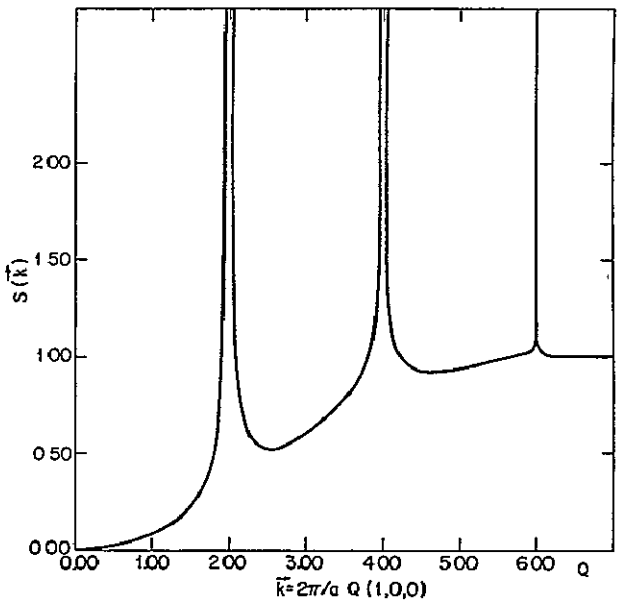


FIG. 2. Structure factor $S(\vec{k})$ for fcc metallic hydrogen (at $r_s = 1.36$ and $T = 0$°K) for $\vec{k}$ along [100]. The frequencies and polarization vectors used to compute $S(\vec{k})$ are the solutions of the self-consistent equations.

to the value of the Debye-Waller factor e^{-2W} where

$$2W = \langle \vec{u}^2 \rangle = \frac{\hbar}{MN} \frac{1}{2} \sum_{\vec{q},j}^{\text{BZ}} e_\alpha(\vec{q},j) e_\beta(\vec{q},j) \omega^{-1}(\vec{q},j) \tag{7}$$

is appreciable.[12] This transfer of weight from the Bragg peaks to the continuum in between means that the *dynamic* second-order energy is *less* sensitive to structure than the corresponding static lattice quantity. Now, in third and higher orders this effect is compounded: It is easy to show[15,18] that the dynamic third-order band-structure energy has three Debye-Waller factors, the fourth has six such factors, and so on. The extent to which the dynamics *reduces* the structural sensitivity is more marked at each successively higher order. Thus, for purposes of calculating the *structural* dependence of the energy, perturbation theory converges more quickly in the dynamic case than in the static counterpart. Perturbation theory does not, of course, say whether the assumption of a crystalline ground state for metallic hydrogen is valid. However within such an assumption, it offers a means for deciding on the preferred lattice; and in this context the calculations described above appear to be the first for a metal that go beyond the harmonic approximation.

*Work supported by the National Aeronautics and

Space Administration under Grant No. NGR-33-010-188 and in part by the National Science Foundation through the facilities of the Cornell University of Materials Science Center (Grant No. DMR-72-03029), Technical Report No. 2749, and Contract No. DMR 74-23494.

†Present address: Department of Meterology, Massachusetts Institute of Technology, Cambridge, Mass. 02139.

[1]V. Heine and D. Weaire, *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1970), Vol. 24, p. 250.

[2]W. A. Harrison, *Pseudopotentials in the Theory of Metals* (Benjamin, New York, 1966).

[3]E. G. Brovman, Yu Kagan, and A. Kholas, Zh. Eksp. Teor. Fiz. 61, 2429 (1971), and 62, 1492 (1972) [Sov. Phys. JETP 34, 1300 (1972), and 35, 783 (1972)].

[4]L. G. Caron, Phys. Rev. B 9, 5025 (1974).

[5]G. A. Neece, F. J. Rogers, and W. G. Hoover, J. Comp. Phys. 7, 621 (1971).

[6]T. Schneider, Helv. Phys. Acta 42, 957 (1969).

[7]J. Hammerberg and N. W. Ashcroft, Phys. Rev. B 9, 409 (1974).

[8]N. S. Gillis, N. R. Werthamer, and T. R. Koehler, Phys. Rev. 165, 951 (1968). For a recent review, see T. R. Koehler, in *Dynamical Properties of Solids*, edited by G. K. Horton and A. A. Maradudin (North-Holland, Amsterdam, 1975), p. 1.

[9]P. Choquard, *The Anharmonic Crystal* (Benjamin, New York, 1971).

[10]This corresponds to a representative pressure of about 1.7 Mbars.

[11]For a static lattice, the energy (to second order) is conventionally written as $E=E_0+E_M+E_b^{(2)}+\cdots$, where E_0 is the energy of the homogeneous interacting electron gas, E_M the Madelung energy, and $E_b^{(2)}$ the second-order band-structure energy. To obtain Eq. (1), we must not only include dynamics but also combine E_M and $E_b^{(2)}$ so that they each contribute to the potential Φ.

[12]The ratio of $\{\langle u^2\rangle\}^{1/2}$ to nearest-neighbor distance is readily calculated (from the self-consistent frequencies and polarizations) to be 0.1687 for the fcc structure at $r_s=1.36$ [see D. Straus, thesis, Cornell University Materials Science Laboratory Report No. 2739 (unpublished)]. This should be compared to the value for Na which for *melting* ($p=0$) is 0.123 [D. Stroud and N. W. Ashcroft, Phys. Rev. B 5, 371 (1972)].

[13]A plot of the static ground-state Gibbs energy G (at a pressure corresponding to the fcc crystal of Fig. 1) looks almost identical to the static-energy plot of Fig. 1; and minimization of G at constant p is, in this case, essentially equivalent to E at constant r_s.

[14]We have used for $\epsilon(k)$ the Hubbard-Geldart-Vosko form, and in the structure-*independent* terms of Eq. (1) we have used the Nozières-Pines approximation to the correlation-energy contribution.

[15]C. Friedli and N. W. Ashcroft, Phys. Rev. B 12, 5552 (1975).

[16]Full details of the calculational methods are found in Straus, Ref. 12. The important point is that the correlation functions $\lambda_{\alpha\beta}(\vec{X})$ were computed using the techniques of D. M. Straus and N. W. Ashcroft [Phys. Rev. B 14, 448 (1976)], except that the *full* directional dependence rather than the angular average of the $\vec{q}\to 0$ portion of the integrand in Eq. (3) has been completely included.

[17]D. M. Straus and N. W. Ashcroft, Phys. Rev. B 14, 448 (1976).

[18]Straus, Ref. 12.

Thermodynamics of Thomas-Fermi screened Coulomb systems*

B Firey and N W Ashcroft
Laboratory of Atomic and Solid State Physics and Materials Science Center, Cornell University, Ithaca, New York 14853
(Received 2 November 1976)

We obtain in closed analytic form, estimates for the thermodynamic properties of classical fluids with pair potentials of Yukawa type, with special reference to dense fully ionized plasmas with Thomas-Fermi or Debye-Huckel screening We further generalize the hard-sphere perturbative approach used for similarly screened two-component mixtures, and demonstrate phase separation in this simple model of a liquid mixture of metallic helium and hydrogen

I. INTRODUCTION AND FORMALISM

The variational procedure of Mansoori and Canfield[1] has proven to be a fruitful source of approximate thermodynamic information for dense classical fluids, liquid metals,[2] liquid alloys,[3,4] and, more recently, the pure classical Coulomb gas.[5] In this brief paper we apply the method to obtain analytic variational estimates analogous to those of Ref. 5 for the case of certain screened Coulomb systems.

We begin with the Hamiltonian for a system of ZN electrons (coordinates $\vec{r}_i$, momenta $\vec{p}_i$, mass m) and N fully ionized atoms [coordinates $\vec{R}$, momenta $\vec{P}(\vec{R})$, mass M and charge Ze]

$$H=\sum_i \frac{p_i^2}{2m}+\frac{1}{2}\sum_{i\neq j}\frac{e^2}{|\vec{r}_i-\vec{r}_j|}+\sum_{\vec{R}}\frac{P^2(\vec{R})}{2M}$$
$$+\frac{1}{2}\sum_{\vec{R}\neq\vec{R}'}\frac{Z^2e^2}{|\vec{R}-\vec{R}'|}-\sum_{i,\vec{R}}\frac{Ze^2}{|\vec{r}_i-\vec{R}|}.$$

Let V be the volume of the system, and let $\rho_e(\vec{k})=\sum_i \exp(i\vec{k}\cdot\vec{r}_i)$ and $\rho_i(\vec{k})=\sum_{\vec{R}}\exp(i\vec{k}\cdot\vec{R})$ be the Fourier transforms of the density operators for electrons and ions, respectively. In the limit $N\to\infty$, $V\to\infty$, $N/V\to n$, we may rewrite H as

$$H=H_{eg}+\sum_{\vec{R}}\frac{P^2(\vec{R})}{2M}+\frac{n}{2}\sum_{\vec{k}\neq 0}\frac{4\pi Z^2e^2}{k^2}[\rho_i(\vec{k})\rho_i(-\vec{k})-1]$$
$$-\sum_{\vec{k}\neq 0}\frac{4\pi Ze^2}{Vk^2}\rho_i(\vec{k})\rho_e(-\vec{k}),$$

where H_{eg} is the standard interacting electron-gas Hamiltonian. To obtain an approximate Helmholtz free energy for the ions, we follow the customary procedure of first calculating an adiabatic linear response of the electrons to the ionic potential, which leads to an ionic Hamiltonian in which the ions can be considered to move according to screened interactions. The variational procedure[1] can then be applied by comparing two isochoric systems, one a hard-sphere reference system and the other a system of particles interacting through a screened Coulomb force. Within the linear screening approximation the free energy is bounded by[2-4]

$$F(\sigma)=F_{ig}+F_{eg}+F_0(\sigma)+\frac{1}{2}\sum_{\vec{k}\neq 0}\frac{N}{V}\frac{4\pi Z^2e^2}{k^2}[S(\vec{k})-1]$$
$$+\frac{1}{2}\sum_{\vec{k}\neq 0}\frac{N}{V}\frac{4\pi Z^2e^2}{k^2}S(\vec{k})\left(\frac{1}{\epsilon(\vec{k})}-1\right), \quad (1)$$

where F_{ig} is the free energy of the corresponding ideal gas of ions, F_{eg} is that of the interacting electron-gas, $\epsilon(\vec{k})$ is the usual dielectric function of the electron gas [taken as $\lim_{\omega\to 0}\epsilon(k,\omega)$, in accordance with the conventional approximation of the theory of metal thermodynamics that the electrons follow the ionic motion adiabatically], $F_0(\sigma)$ is the excess free energy of a gas of hard spheres of diameter σ, and $S(\vec{k})$ is the structure factor of the hard-sphere gas.

With $n=N/V$, we may rewrite Eq. (1) as

$$F(\sigma)=F_{ig}+F_{eg}+F_0(\sigma)+\frac{n}{2}\sum_{\vec{k}\neq 0}\frac{4\pi Z^2e^2}{k^2}\left(\frac{1}{\epsilon(\vec{k})}-1\right)$$
$$+\frac{n}{2}\sum_{\vec{k}\neq 0}\frac{4\pi Z^2e^2}{k^2}\frac{1}{\epsilon(\vec{k})}[S(\vec{k})-1], \quad (2)$$

we now identify the fifth term as an effective-pair interaction between ions, and the fourth as the self-energy of the screened ions.

Our observation is that these terms can be obtained in closed analytic form for dielectric functions of the type

$$\epsilon(\vec{k})=1+q^2/k^2 \quad (3)$$

and the Percus-Yevick hard-sphere structure factor. Dielectric functions of this form are found in two physically significant limiting cases: the high-density low-temperature limit, in which the Thomas-Fermi dielectric function is appropriate [with $q=(6\pi e^2Zn/E_F)^{1/2}$, E_F being the Fermi energy]; and the low-density high-temperature limit, in which the Debye-Hückel form for $\epsilon(k)$ is suitable [i.e., $q=(4\pi e^2Zn/kT)^{1/2}$].

For dielectric functions of the form (3), the structure-independent fourth term of (2) becomes

$$\frac{n}{2}\sum_{\vec{k}\neq 0}\frac{4\pi Z^2e^2}{k^2}\left(\frac{-q^2}{q^2+k^2}\right)=-\frac{NZ^2e^2q}{2}, \tag{4}$$

and the fifth term may be written in coordinate space as

$$\frac{n}{2}\sum_{k}\frac{4\pi Z^2e^2}{k^2+q^2}[S(k)-1]$$
$$=Nn\frac{Z^2e^2}{2}\int_0^\infty dr\,4\pi r^2\frac{\exp(-qr)}{r}[g(r)-1], \tag{5}$$

where $g(r)$ is the standard radial distribution function for the fluid. It should be noted that we have here used the *liquid* structure factor, from which a δ-function term at $\vec{k}=0$ has been subtracted, removing a term associated with the bulk isothermal compressibility of the electron gas.

Now, the right-hand side of (5) is essentially the Laplace transform of $rg(r)$, and is available analytically for the hard-sphere fluid in the Percus-Yevick[6] approximation.[7,8] In a notation similar to that of Wertheim,[7] with $x=r/\sigma$

$$\int_0^\infty dx\,\exp(-\lambda x)[xg(x)-x]=G(\lambda)-1/\lambda^2,$$

where

$$G(\lambda)=\frac{\lambda L(\lambda)}{12\eta[L(\lambda)+\bar{S}(\lambda)e^{\lambda}]}$$

with

$$L(\lambda)=12\eta[(1+\tfrac{1}{2}\eta)\lambda+(1+2\eta)],$$
$$\bar{S}(\lambda)=(1-\eta)^2\lambda^3+6\eta(1-\eta)\lambda^2+18\eta^2\lambda-12\eta(1+2\eta).$$

Here η is the packing fraction, given by $\eta=(\pi/6)n\sigma^3$.

We thus achieve in closed form the following single-parameter expression for the free energy:

$$F(\eta)=F_{ig}+F_{eg}+F_0(\eta)-NZ^2e^2q/2+N(Z^2e^2/2r_0)$$
$$\times 12\eta^{2/3}[G(2\eta^{1/3}qr_0)-(2\eta^{1/3}qr_0)^{-2}] \tag{6}$$

where $(4\pi/3)r_0^3=V/N$. An appropriate expression for the excess free energy of the hard-sphere system is that of Carnahan and Starling[9]:

$$F_0(\eta)=NkT\eta(4-3\eta)/(1-\eta)^2.$$

An approximate lowest upper bound on $F(\eta)$ can now be obtained by appealing to the Gibbs-Bogolyubov inequality[10] and minimizing (6) in η; that is, for a fixed q and r_0, we impose $\partial F(\eta)/\partial\eta=0$. The resulting transcendental equation in η can be solved numerically to obtain the minimizing value of η, which we denote η^*, we then approximate the true free energy as $F(\eta^*)$. The thermodynamic derivatives can likewise be evaluated; we have, for example,

$$p=-\left(\frac{\partial F(\eta^*)}{\partial V}\right)_{T,\eta}. \tag{7}$$

Note that differentiations may be performed at fixed η by virtue of our variational condition.

We now illustrate the procedure for the case of Thomas-Fermi screening.

II ONE-COMPONENT THOMAS-FERMI GAS

For this case, we have $qr_0=(12Z/\pi)^{1/3}r_s^{1/2}$, where r_s is the usual electron-gas parameter

$$r_s=r_0/Z^{1/3}a_0.$$

Requiring $\partial F/\partial\eta=0$ imposes a transcendental equation of the form $f(T,r_s,\eta^*)=0$. It proves most convenient to solve this equation numerically for $r_s(\eta,T)$ and find the equation-of-state data in parametric form analogous to that of Ref. 5; we present in Fig. 1 results for r_s. The Thomas-Fermi approximation for the dielectric function is appropriate for systems satisfying $qr_0\ll 1$, i.e., (with $Z=1$, for which we have done all our calculations) for $r_s\ll 0.4$. In Fig. 2 we plot the plasma parameter $\Gamma_{TF}=(Ze)^2/r_0kT$ against η as well as the corresponding Γ_c for the unscreened Coulomb system of Ref. 5. As expected, Γ_{TF} approaches Γ_c as $r_s\to 0$.

Using for F_{eg} the zero-temperature RPA form,[11] we next compute the free energy and the pressure;

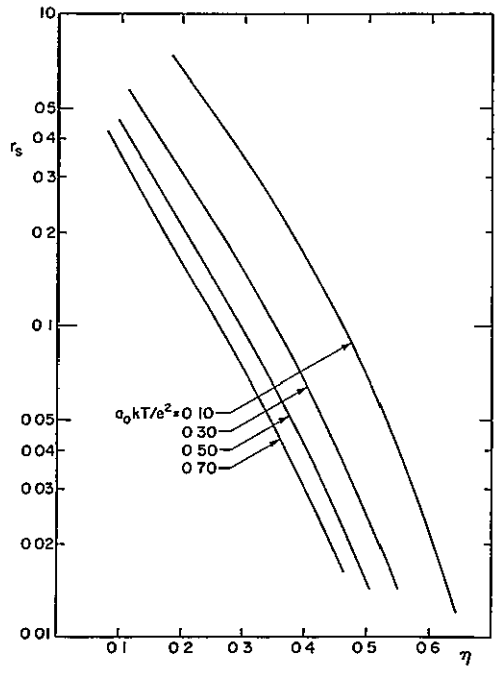


FIG. 1 r_s as a function of the minimizing value of η for the Thomas-Fermi case, from numerical solution of the transcendental equation of the variational condition

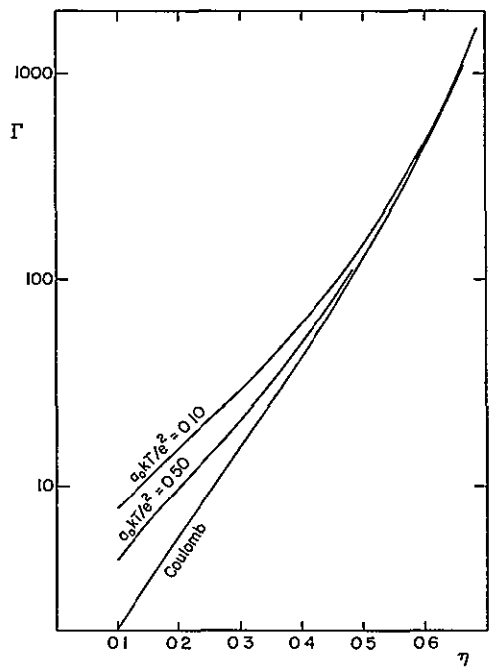


FIG. 2. Plasma parameter Γ versus the minimizing value of η, for the Thomas-Fermi case and the unscreened Coulomb gas, computed from the r_s of Fig. 1 and Ref 5

these we plot in Figs. 3 and 4, respectively.

A somewhat analogous computation has recently been carried out by Ross and Seale[12] using the RPA dielectric function (rather than the Thomas-Fermi form) and in which the second-order band-structure energy [the fifth term in (1)] is obtained

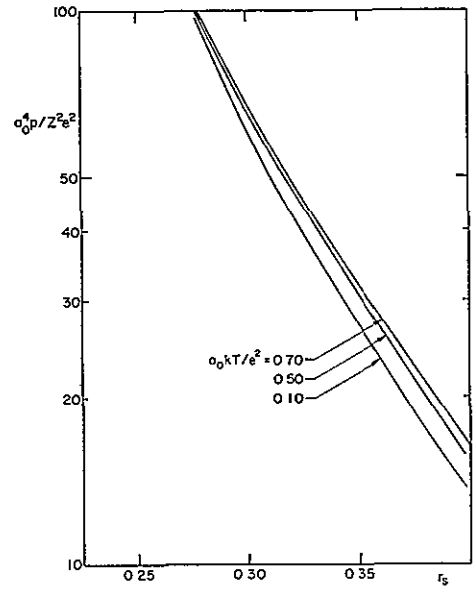


FIG 4 Pressure in atomic units as a function of r_s, from Eq. (7).

by numerical integration. We plot in Fig. 5 their excess free energy W [essentially the last three terms of Eq. (2)], and in Fig. 6 the excess pressure, for $r_s = 0.1$, together with our results. The agreement is seen to be excellent, especially in the excess pressure. Furthermore, in Figs. 7 and 8 we exhibit the corresponding plots at $r_s = 1$. We again see excellent agreement despite the fact that at this value of r_s one would not expect the Thomas-Fermi form of the dielectric function to

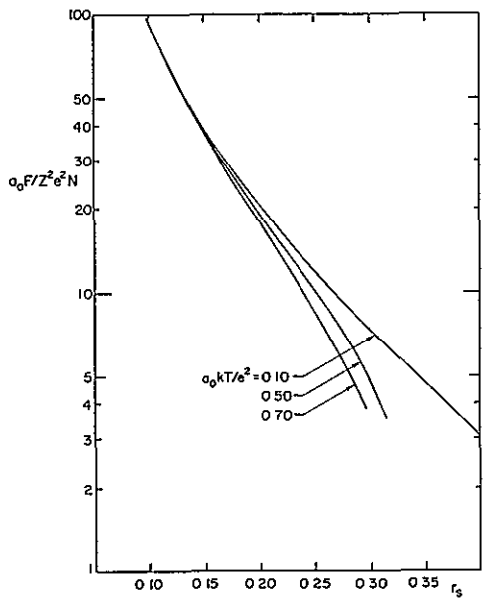


FIG 3. Free energy per ion as a function of r_s, in atomic units, from Eq (6).

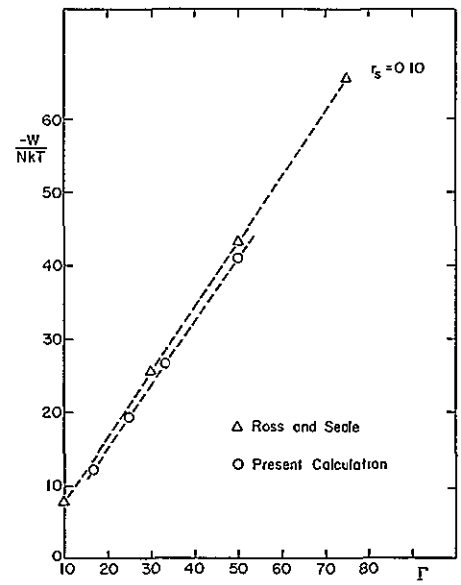


FIG 5 Excess free energy in temperature units for $r_s = 0.1$ as a function of Γ, compared with values from Ref 12

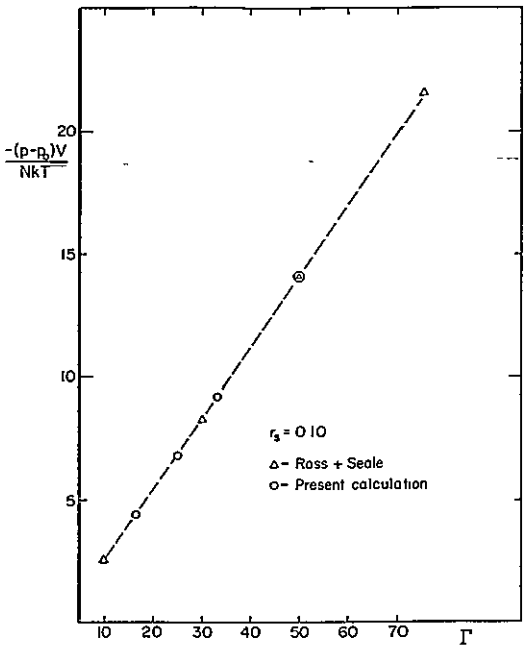


FIG 6 Excess pressure in temperature units for $r_s = 0\ 1$ as a function of Γ, compared with values from Ref 12

closely approximate the RPA form; the conclusion is perhaps that the two are essentially equivalent for the calculation of thermodynamic properties as the result of cancellations of somewhat fortuitous character in the integrals, at least at the larger value of r_s.

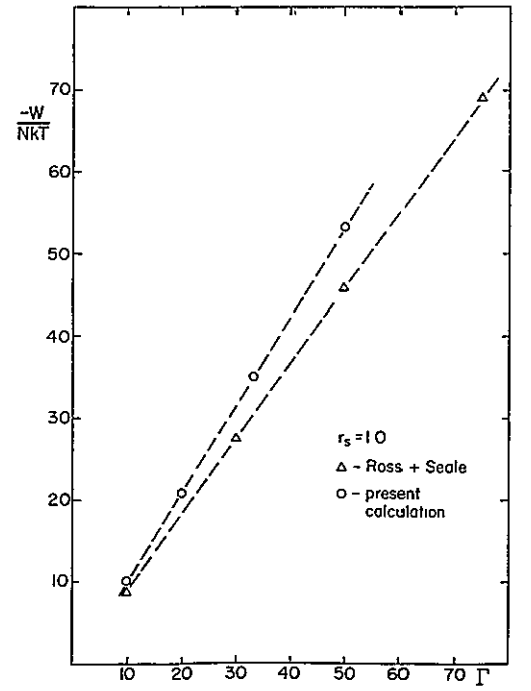


FIG 7 As in Fig 5, but with $r_s = 1\ 0$.

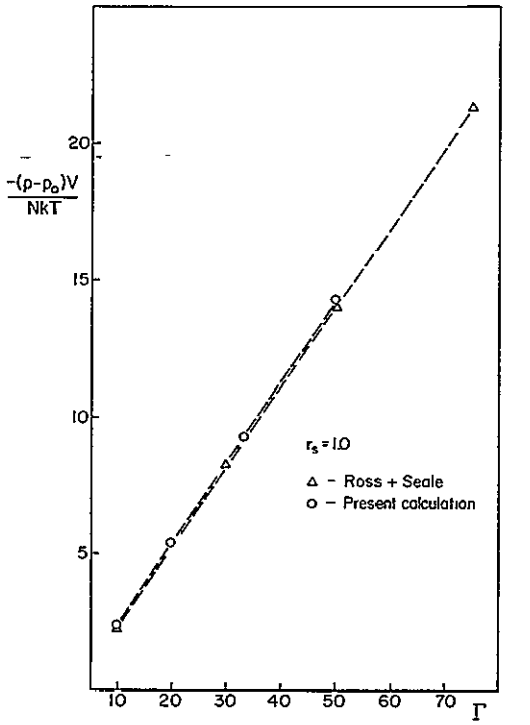


FIG 8 As in Fig 6, but with $r_s = 1\ 0$

It may be noted that modified forms of $g(r)$ have been proposed[13] which match the results of computer simulations somewhat better than the simple hard-sphere Percus-Yevick results, especially at higher densities. We have used the Percus-Yevick result here chiefly because of its analytic simplicity; we have, however, examined the corrections resulting from the use of the forms for $g(r)$ proposed in Ref. 13, and have found them to be only 1 or 2% of the pair energy, even at packing fractions as high as 0.68.

III TWO-COMPONENT THOMAS-FERMI SYSTEM

The above discussion can be readily generalized to the case of a system consisting of a mixture of fully ionized atoms of differing nuclear charges, together with their neutralizing background of electrons. In particular, the free energy of the two-component system with nuclear charges $Z_1 e$ and $Z_2 e$ and respective mole fractions x_1 and x_2 may be written in a manner similar to Eq. (6).

Let $\vec{r}_l$ now index electron coordinates, and $\vec{R}_\alpha^i$ ($\alpha = 1, \ldots, N_i$) the coordinates of ions of charge $Z_i e$ and mass M_i ($i = 1, 2$). Then the Hamiltonian for the assembly of electrons and fully ionized atoms is

$$H = \sum_l \frac{p_l^2}{2m} + \frac{1}{2} \sum_{l,m}{}' \frac{e^2}{|\vec{r}_l - \vec{r}_m|} + \sum_{\alpha,i} \frac{P^2(\vec{R}_\alpha^i)}{2M_i}$$

$$+ \frac{1}{2} \sum_{\alpha\beta,ij}{}' \frac{Z_i Z_j e^2}{|\vec{R}_\alpha^i - \vec{R}_\beta^j|} - \sum_l \sum_{\alpha,i} \frac{Z_i e^2}{|\vec{r}_l - \vec{R}_\alpha^i|},$$

or, introducing Fourier transforms of ionic densities

$$\rho_i^{\text{ion}} = \sum_{\alpha=1}^{N_i} \exp(i\vec{k}\cdot\vec{R}_\alpha^i),$$

we can rewrite the Hamiltonian as

$$H = H_{eg} + \sum_{\alpha,i} \frac{P^2(\vec{R}_\alpha^i)}{2M_i} + \frac{N}{2} \sum_{\vec{k}\neq 0} \sum_{i,j} \frac{4\pi Z_i Z_j e^2}{Vk^2} (x_i x_j)^{1/2} \times \{(N_i N_j)^{-1/2} \rho_i^{\text{ion}}(\vec{k}) \rho_j^{\text{ion}}(-\vec{k}) - \delta_{ij}\} - \sum_{\vec{k}\neq 0} \sum_i \frac{4\pi Z_i e^2}{Vk^2} \rho_i^{\text{ion}}(\vec{k}) \rho^e(-\vec{k}),$$

where $N = \sum N_i$.

The arguments leading to (1) are now repeated, the essential difference here being that the reference system is taken as a two-component (rather than one-component) hard-sphere fluid. The approximate free-energy is then

$$F = F_{ig} + F_{eg} + F_0(\sigma_1, \sigma_2) - \frac{Ne^2 q}{2}(x_1 Z_1^2 + x_2 Z_2^2) + Nn \frac{e^2}{2} \sum_{i=1,2} \sum_{j=1,2} \int_0^\infty dr\, 4\pi r^2 \frac{\exp(-qr)}{r} \times Z_i Z_j [g_{ij}(r) - 1] x_i x_j, \tag{8}$$

where F_{ig} is the free energy of a two-component ideal mixture, F_0 the excess free energy of the two-component reference system (i.e., a mixture of hard spheres of diameters σ_1 and σ_2), and the g_{ij}'s are the appropriate radial distribution functions for the reference system. The objects of interest are again the Laplace transforms

$$G_{ij}(\lambda) = \int_0^\infty dr \exp(-\lambda r) r g_{ij}(r).$$

These quantities have been given for the two-component hard-sphere system in the Percus-Yevick approximation by Lebowitz.[14] We combine these with the form for F_0 quoted by Umar *et al.*,[3] which corresponds to an equation of state derived from the hard-sphere partition function rather than from the compressibility, (or even an interpolation between the two), and is used because it conveniently separates into structure-dependent and structure-independent parts. The free energy can again be given in a closed form depending on two parameters, which it is useful to take as the total packing fraction $\eta = \frac{1}{6}\pi n(\sigma_1^3 + \sigma_2^3)$ and the ratio of hard-sphere diameters, $\alpha = \sigma_2/\sigma_1$, chosen to lie between 0 and 1. For the dielectric function we use Eq (3) with

$$q = (6\pi e^2 Z^* n/E_F)^{1/2},$$

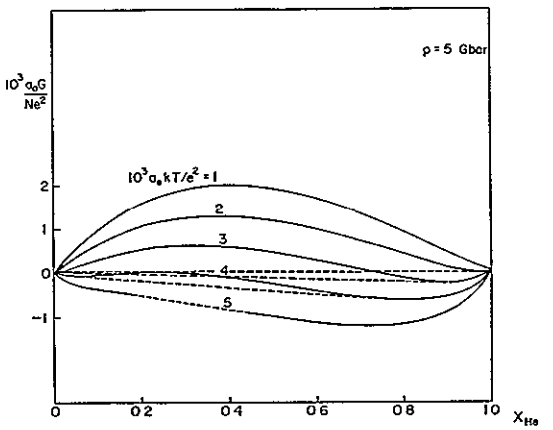


FIG 9 Typical excess Gibbs energies in atomic units for H-He mixtures at 5 Gbar, showing common tangents

with

$$Z^* = \sum_i x_i Z_i.$$

Since we now have two independent parameters, it proves most efficient to do the minimization by a search procedure in (η, α) space, not using derivatives. We make use of Brent's modification of Powell's algorithm[15] for this purpose, which is found to give excellent convergence for the functions in question.[16]

Our calculation has been directed primarily toward the question of phase separation in these fluids. Since physically meaningful calculations of this type must be performed at constant pressure, we compute the Gibbs function $G(T, p, x_1)$ at the specified pressure, using a zero-finding procedure to determine the necessary values of r_0, and then perform a Maxwell common-tangent construction to obtain the phase boundary.

Phase diagrams were computed in this manner for hydrogen-helium mixtures at three pressures: 60 Mbar, 5 Gbar, and 10 Gbar, or, respectively, 0.204, 16 995, and 33.990 a.u The 60 Mbar pressure corresponds to $r_s \simeq 0.84$, which is outside the regime in which the Thomas-Fermi dielectric function is expected to be realistic; it is provided for comparison with the work of Stevenson,[4] who performed a similar calculation using the Hubbard-Geldart-Vosko dielectric function[17] and including in the free-energy terms arising from the next order in the electron-gas response and the leading quantum correction to the ionic structure factor. For the 5 Gbar pressure, $r_s \simeq 0.38$, and at 10 Gbar, $r_s \simeq 0.33$; so for these pressures the Thomas-Fermi form is suitable. We display in Fig. 9 some typical forms of the excess Gibbs potential defined as $G(T, p, x_1) - x_1 G(T, p, 1) - (1 - x_1) G(T, p, 0)$ for a pressure of 5 Gbar, and in Figs. 10 and 11 the phase diagrams themselves, together with

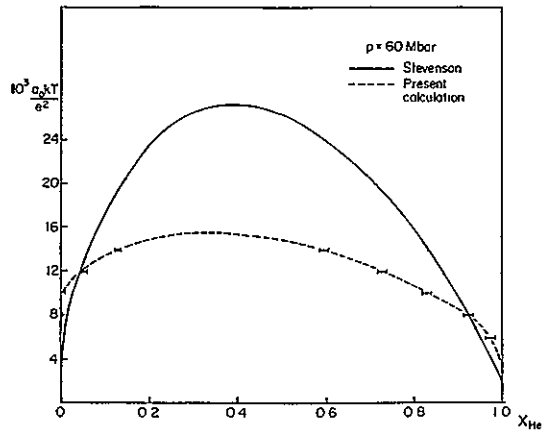


FIG 10 Phase boundaries for H-He mixtures at 60 Mbar, together with results from Ref 4 Error bars show estimated uncertainty in phase boundary due to the choice of interpolation scheme in the common tangent construction.

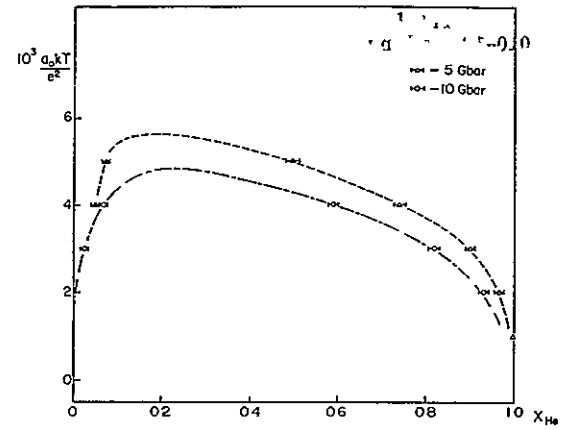


FIG 11 Phase boundaries for H-He mixtures at 5 and 10 Gbar, error bars as in Fig 10

Stevenson's results. It will be noted that the phase boundaries are highly asymmetric—a characteristic they share with those recently calculated by Hansen[18] for the *unscreened* Coulomb system by numerical solution of the hypernetted-chain equations—and are qualitatively similar to Stevenson's, but differ substantially in the temperature scale, a difference which seems to be due to the differing pair potential, in accordance with long-standing belief[19] that the details of phase separation are determined chiefly by the long-range portion of the pair potential rather than the hard core. It is of interest to note that the ratio of hard-sphere diameters is quite insensitive to temperature, pressure, and composition; it stays in the range 0.76–0.78 throughout, a result also typical of Stevenson's calculations. Furthermore, at the two higher pressures considered, the critical point is found to correspond to $\eta = 0.62$ or $\eta = 0.65$ (for 5 and 10 Gbar, respectively). These values of η are high enough to suggest that at the corresponding pressures the mixtures may solidify before phase separation begins in the liquid, a fact which may be of some astrophysical interest.

Finally, the use of hard-sphere structure factors other than Percus-Yevick might be expected to shift the phase boundaries, but should not alter the conclusions concerning either the existence of phase separation or the onset of solidification.

IV DISCUSSION

We have obtained analytic variational estimates for the thermodynamic properties of a particularly simple class of screened Coulomb potentials, which may provide useful comparisons both with Monte Carlo calculations and astrophysical data. Dense, fully ionized systems of the type we consider occur, and may be of observational interest, in cooled white dwarf stars; it is also possible that some pellet-compression schemes for controlled fusion may involve the formation of regions of appropriate density and temperature—i.e., $r_s \simeq 0.3$, pressures of a few Gbar, and temperatures of a few eV. Furthermore, our results may be useful in improving the accuracies of hard-sphere variational calculations for metals under more ordinary conditions by supplying a better analytic approximation to the free energy than the Madelung energy which has hitherto been used.

As we mentioned above, our calculation can also be performed for the case of Debye-Hückel screening. In this case, we have $qr_0 = (6\Gamma/Z)^{1/2}$, where Γ is the plasma parameter defined in Sec. II. (Since the Debye-Hückel system is purely classical, Γ is the sole parameter of interest; i.e., the density and temperature dependences of all thermodynamic quantities are related in a simple scaling fashion.) The approximation is again valid for $qr_0 \ll 1$, or $\Gamma \ll \frac{1}{6}$—provided, of course, that the electron gas is far from degeneracy, that is, $kT \gg E_F$. It is readily found, however, that in this regime the excess free energy is dominated by the structure-independent self-energy terms, to which structure-dependent terms add a correction of only a few percent. Nevertheless, if questions of phase separation in mixtures in this regime prove to be of interest, calculations analogous to those of Sec. III could be performed.

*Work supported in part by NASA Grant No NGR-33-010-188 and in part by NSF Grant No DMR-74-23494 and through the Materials Science Center of Cornell University, Technical Report No 2725 One of the authors (B F) was supported by an NSF Graduate Fellowship.

[1]G A Mansoori and F. B Canfield, J Chem. Phys 51, 4958 (1969)
[2]H D. Jones, J. Chem. Phys 55, 2640 (1971).
[3]I H Umar, A. Meyer, M. Watabe, and W. H. Young, J Phys F 4, 1691 (1974).
[4]D Stevenson, Phys Rev B 12, 3999 (1975)
[5]D Stroud and N W Ashcroft, Phys Rev A 13, 1660 (1976)
[6]J K Percus and G. J. Yevick, Phys Rev. 110, 1 (1958).
[7]M S Wertheim, Phys Rev Lett 10, 321 (1963).
[8]E Thiele, J. Chem. Phys 39, 474 (1961)
[9]N F. Carnahan and K E Starling, J Chem Phys 51, 635 (1969), 53, 600 (1970)
[10]See, for example, R P Feynman, *Statistical Mechanics* (Benjamin, New York, 1972), p. 67
[11]See, for example, A Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, New York, 1971), p 166
[12]M Ross and D Seale, Phys. Rev A 9, 396 (1974).
[13]L Verlet and J. Weis, Phys Rev. A 5, 939 (1972)
[14]J. L Lebowitz, Phys Rev. A 133, 895 (1964)
[15]R P Brent, *Algorithms for Minimization Without Derivatives* (Prentice-Hall, Englewood Cliffs, N.J , 1973)
[16]All calculations were performed using double-precision arithmetic on the Cornell University IBM 370/168 computer.
[17]D. J. W. Geldart and S H Vosko, Can J Phys 44, 2137 (1966); see also D Stroud and N W. Ashcroft, Phys Rev B 5, 1371 (1972).
[18]J P. Hansen and P Vieillefosse, Phys Rev Lett. 37, 391 (1976)
[19]See, for example, R L Henderson and N W Ashcroft, Phys Rev A 13, 859 (1976)

Combined representation method for use in band-structure calculations: Application to highly compressed hydrogen*

Carlos Friedli† and N W Ashcroft
Laboratory of Atomic and Solid State Physics and Materials Science Center, Cornell University, Ithaca, New York 14853
(Received 19 January 1977)

A representation is described whose basis functions combine the important physical aspects of a finite set of plane waves with those of a set of Bloch tight-binding functions. The chosen combination has a particularly simple dependence on the wave vector $\vec{k}$ within the Brillouin zone, and its use in reducing the standard one-electron band-structure problem to the usual secular equation has the advantage that the lattice sums involved in the calculation of the matrix elements are actually independent of $\vec{k}$. For systems with complicated crystal structures, for which the Korringa-Kohn-Rostoker, augmented-plane-wave, and orthogonalized-plane-wave methods are difficult to use, the present method leads to results with satisfactory accuracy and convergence. It is applied here to the case of compressed molecular hydrogen taken in a $Pa3$ (α-nitrogen) structure for various densities but with mean interproton distance held fixed. The bands show a marked free-electron character above 5 to 6 times the normal density, and the overall energy gap is found to vanish at 9.15 times normal density. Within the approximations made, this represents an upper bound for the molecular density in the transition to the metallic state from an α-nitrogen structure.

I. INTRODUCTION

The method described below evolved from an attempt to obtain the band structure of a system such as molecular hydrogen in a relatively complex crystal structure, and over a range of densities. For certain regions of the density it is expected on general grounds that neither the low-density tight-binding approach [with a representation of linear combinations-of-atomic-orbitals (LCAO) Bloch functions] nor the methods using a representation with a basis of simple plane waves (PW) are physically adequate.

For reasons principally connected with the structure, the other familiar methods are also not entirely adequate,[1,2] at least in their standard formulations. The Korringa-Kohn-Rostoker (KKR) and augmented-plane-wave methods not only require a substantial amount of computational effort, but are based on a muffin-tin approximation to the actual one-electron potential.[3-5] This means a "sphericalization" (taking the average over angles) of the potential arising from the contents of a unit cell, a procedure which is difficult to justify when the molecules in the crystal have no obvious spherical symmetry. Although such models yield useful physical information especially at lower densities, it is difficult to estimate their accuracy, particularly at higher densities, where steric effects and the requirements of proper crystal symmetry may become important. The effects of the latter on the resulting band structure may well be important as has been shown by Painter[6] in his treatment of non-muffin-tin corrections to KKR bands by the discrete variational method.[7]

Furthermore, there is often no clear-cut separation between core levels (actually nonexistent for hydrogen) for which tight binding is adequate, and the rest of the band levels (valence and conduction), which would make an orthogonalized-plane-wave method appropriate. Even if one makes an arbitrary separation between valence and conduction levels, and treats the first with tight-binding functions and the second with orthogonalized-plane-wave functions orthogonalized to the valence levels,[8] one still has the possibility of significant overlap of these "core" levels in situations such as the one here, where large variations in density are of physical interest.

For these reasons it is natural to investigate alternative representations whose basis functions combine in some way the advantages of both the LCAO functions (with their physically correct atomic behavior near the nuclei) and the PW, which are more satisfactory in the region between atoms. One such basis set was recently used by Ramaker *et al.*[9] in exact-exchange crystal Hartree-Fock calculations of molecular and metallic hydrogen. Another, based on a more general and flexible approach, is described below. It is a modification of an idea used successfully by Brown and Krumhansl,[10] which was shown to be mathematically equivalent to the orthogonalized-plane-wave method.

In Sec. II, the representation will be developed and its basic properties described. Section III is devoted to a discussion of the application of the representation to the solution of the one-electron problem in crystals. In Sec. IV, we present the results of the applications of the method to molecular hydrogen [assumed to be in α-nitrogen ($Pa3$) crystal structure] over a wide range of densities,

but with interproton distance generally held fixed. The most interesting point to emerge from the resulting band structure is the observation that valence and conduction bands begin to overlap at a lattice constant of $a = 4.78$ bohr, which corresponds to a density equal to 9.15 times its zero-pressure value. If the crystalline phase remains stable at such densities, this represents a metal-insulator transition at a density of approximately 0.83 g/cm^3.

II. REPRESENTATION

The representation we introduce is formally incomplete: It has a finite set of basis wave functions. This set is made up of a finite number of PW and a set of specially constructed Bloch functions. It is constructed in such a way that the whole set is orthonormal, and although the set is finite, linear combinations of them are expected to give variationally good approximations to the eigenfunctions and corresponding eigenvalues. This expectation is based on the physical way the set is constructed, which will be explained in what follows.

Consider first a monatomic (for example, a simple cubic) lattice with lattice constant a and LCAO Bloch function $h_{\vec{k}}(\vec{r})$ defined with atomic orbital $\Phi(\vec{r})$,

$$h_{\vec{k}}(\vec{r}) = \frac{1}{\sqrt{N}} \sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} \Phi(\vec{r} - \vec{R}) , \tag{1}$$

where N is the number of cells in a volume Ω, $\vec{R}$ designates their position vectors, and $\vec{k}$ is the Bloch wave vector. Expressing this Bloch function in its well-known form

$$h_{\vec{k}}(\vec{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\vec{K}} c_{\vec{k}-\vec{K}} e^{i(\vec{k}-\vec{K})\cdot\vec{r}} , \tag{2}$$

where $\vec{K}$ is the set of reciprocal-lattice vectors corresponding to $\vec{R}$, it is easy to see that

$$c_{\vec{q}} = (N/\Omega)^{1/2} \Phi_{\vec{q}} , \tag{3}$$

where $\Phi_{\vec{q}}$ is the Fourier transform of $\Phi(\vec{r})$.

For the purposes of defining a trial function, $\Phi(\vec{r})$ may be *any* localized orbital, and not necessarily an atomic one. This observation will be used to construct a particularly convenient type of Bloch function. But instead of defining it directly (i.e., in $\vec{r}$ space) it is inferred from conditions imposed on $c_{\vec{q}}$. In this way it is easier to enforce (through them) the properties that one would like the Bloch levels to have. First, some general observations.

One expects the eigenfunctions not to change too much very near (and particularly inside, if there is a core) the atoms or molecules forming the solid from the values they assume in corresponding free atoms or molecules. This remains true even at fairly high densities. Thus, one wants to include in the basis set Bloch functions built with atomic or molecular orbitals to obtain a good representation in this region. But it is clear that for this purpose only those components $c_{\vec{k}-\vec{K}}$ with sufficiently large $\vec{K}$ are relevant (here, $\vec{k}$ is assumed to be restricted to the first Brillouin zone B_0). On the other hand, if the itinerant or free-electron character becomes important (as it will at high densities), plane waves with wave vectors (about the origin) not too large in terms of $2\pi/a$ are obviously indicated. We now construct basis functions incorporating these features. The Bloch function is first modified by truncating its Fourier components of low wave vectors, say $\vec{G}$, in some finite subset G of the reciprocal lattice K. In this way, the plane waves with wave vectors $\vec{k} - \vec{G}$ have been set free to be included in the basis set as independent members orthogonal to the Bloch functions. (For simplicity, in some of the algebraic manipulations the subset G may be chosen symmetrically to include both $\vec{G}$ and $-\vec{G}$, although this is not required in general by the method.) For the simple-cubic-lattice case, for example, we may choose G to be the set of all reciprocal-lattice vectors within or on the surface of a cube centered at the origin, and with faces perpendicular to the axes. Further, let T be the complement of G, that is $G \cap T$ is empty and $G \cup T = K$. Next, the Bloch functions of the basis are to be chosen to have *as simple* a form as possible, a requirement for both analytical and computational purposes. In particular, the most simple functional dependence on $\vec{k}$ is essential.

In the case of a Bravais lattice, a set of Bloch functions satisfying these criteria can be taken to have components

$$c_{\vec{q}} = \left(\frac{N}{\Omega}\right)^{1/2} \sum_{\vec{K}} \chi_{B_0}(\vec{q} - \vec{K}) \chi_T(\vec{K}) \Phi_{\vec{K}} , \tag{4}$$

where the characteristic function $\chi_A(\vec{x})$ is given by

$$\chi_A(\vec{x}) = \begin{cases} 1 & \text{if } \vec{x} \in A \\ 0 & \text{otherwise.} \end{cases}$$

Here, $\Phi(\vec{r})$ is a localized orbital. Figure 1 shows a schematic one-dimensional example of the procedure just outlined; there, the dotted curve represents the Fourier transforms Φ_q of a localized orbital and the discontinuous curve the components $(\Omega/N)^{1/2} c_q$ given by Eq. (4); note also that the set G contains *by choice* only the reciprocal-lattice vectors 0 and $\pm 2\pi/a$.

The functions defined by Eq. (4) all have the properties of Bloch functions, and can, of course, be written as

$$h_{\vec{q}}(\vec{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\vec{K}} c_{\vec{q}-\vec{K}}\, e^{i(\vec{q}-\vec{K})\cdot\vec{r}} \,. \tag{5}$$

This reduces, for $\vec{q}=\vec{k}\in B_0$, to the standard form

$$h_{\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}}\left(\frac{\sqrt{N}}{\Omega}\sum_{\vec{K}\in T}\Phi_{\vec{K}}\, e^{i\vec{K}\cdot\vec{r}}\right), \tag{6a}$$

and is equivalent also to

$$h_{\vec{k}}(\vec{r}) = \frac{1}{\sqrt{N}}\, e^{i\vec{k}\cdot\vec{r}}\left[\sum_{\vec{R}}\Phi(\vec{r}-\vec{R}) - \frac{N}{\Omega}\sum_{\vec{G}\in G}\Phi_{\vec{G}}\, e^{i\vec{G}\cdot\vec{r}}\right]. \tag{6b}$$

where the quantity in square brackets clearly has the periodicity of the lattice. The prefactor in the expression for $c_{\vec{q}}$ is not important except to keep track formally, and in a consistent way, of the various constants and factors involved. (It cancels, of course, when normalizing the functions.)

The norm of $h_{\vec{k}}(\vec{r})$, $\|h\|$, is independent of $\vec{k}$ and is given by

$$\|h\|^2 = \frac{N}{\Omega}\sum_{\vec{K}\in T}|\Phi_{\vec{K}}|^2 \,, \tag{7}$$

or equivalently, by

$$\|h\|^2 = \sum_{\vec{R}}\langle\Phi(\vec{r})|\Phi(\vec{r}-\vec{R})\rangle - \frac{N}{\Omega}\sum_{\vec{G}\in G}|\Phi_{\vec{G}}|^2 \,. \tag{8}$$

With the normalized functions $h_{\vec{k}}(\vec{r})/\|h\|$, the corresponding Wannier function $w(\vec{r})$ can be obtained,

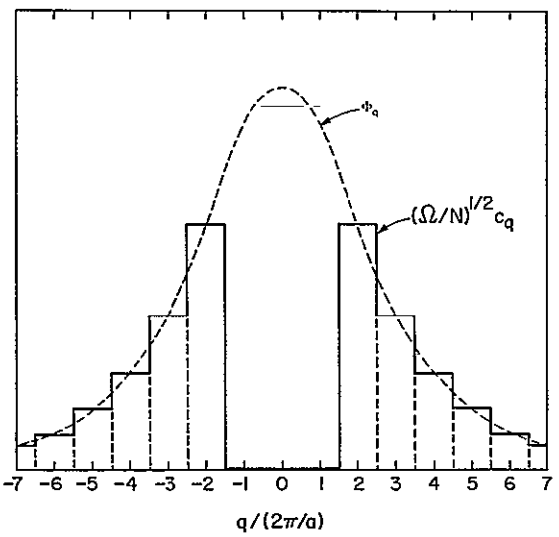


FIG. 1 Schematic one-dimensional example of components $(\Omega/N)^{1/2}c_q$ of a member of the new representation given by Eq. (4) (discontinuous curve) in terms of the Fourier transform Φ_q of a localized orbital (dotted curve). The reciprocal-lattice vectors correspond here to $q/(2\pi/a)$ = integer. Note that c_q is identically zero in the central zones (corresponding to a choice here of a set of reciprocal-lattice vectors $G=\{-2\pi/a, 0, 2\pi/a\}$) and constant within each zone corresponding to the reciprocal-lattice vectors falling outside G (set T)

and is given by

$$w(\vec{r}) = \frac{1}{\|h\|\sqrt{N\Omega}}\sum_{\vec{q}\ (\text{all})} c_{\vec{q}}\, e^{i\vec{q}\cdot\vec{r}} \,, \tag{9}$$

which in this form is automatically normalized. It is, of course, orthogonal to $w(\vec{r}-\vec{R})$ for $\vec{R}\neq 0$. Substituting in Eq. (10) for $c_{\vec{q}}$, one gets

$$w(\vec{r}) = \frac{1}{\|h\|}\left(\frac{N}{\Omega}\right)^{1/2}\left(\sum_{\vec{K}\in T}\Phi_{\vec{K}}\, e^{i\vec{K}\cdot\vec{r}}\right)w_0(\vec{r}) \tag{10a}$$

or

$$w(\vec{r}) = \frac{1}{\|h\|}\left(\frac{\Omega}{N}\right)^{1/2}\left[\sum_{\vec{R}}\Phi(\vec{r}-\vec{R}) - \frac{N}{\Omega}\sum_{\vec{G}\in G}\Phi_{\vec{G}}e^{i\vec{G}\cdot\vec{r}}\right]w_0(\vec{r}) \,, \tag{10b}$$

where for the case of a simple cubic Bravais lattice

$$w_0(\vec{r}) = \frac{1}{\sqrt{N\Omega}}\sum_{\vec{k}\in B_0} e^{i\vec{k}\cdot\vec{r}} = \left(\frac{N}{\Omega}\right)^{1/2}\frac{\sin(\pi x/a)}{\pi x/a}\,\frac{\sin(\pi y/a)}{\pi y/a}\,\frac{\sin(\pi z/a)}{\pi z/a} \tag{11}$$

is the empty lattice lowest-band Wannier function.

It is clear from the form of $h_{\vec{k}}(\vec{r})$ and $w(\vec{r})$ that these functions have the right behavior near and at the lattice sites $\vec{R}$, particularly if the finite set G does *not* contain large wave vectors. And for all $\vec{G}\in G$, $h_{\vec{k}}(\vec{r})$ is automatically orthogonal to the plane waves with wave vector $\vec{k}-\vec{G}$.

In this way, we have an incomplete but orthonormal basis set which would clearly be sufficient for a monatomic lattice if it were not necessary to use more than one localized $\Phi(\vec{r})$.

Except for small $\vec{k}$, the Bloch function $h_{\vec{k}}(\vec{r})$ just defined will *not* in general be a good approximation to the solution $\Psi_{\vec{k}}(\vec{r})$ of the one-electron problem of the crystal if G is empty (i.e., if no PW are included in the basis). The functions $h_{\vec{k}}(\vec{r})$ and $\Psi_{\vec{k}}(\vec{r})$ can differ substantially for larger $\vec{k}$, particularly near the boundaries of the Brillouin zone, simply because the Fourier components of $e^{-i\vec{k}\cdot\vec{r}}\Psi_{\vec{k}}(\vec{r})$ are functions of $\vec{k}$, while those of $e^{-i\vec{k}\cdot\vec{r}}h_{\vec{k}}(\vec{r})$ are not. Nevertheless, considering their expansions in reciprocal space, we find that as $\vec{K}$ increases, the difference in their components *decrease*, since *by construction* both functions have the same form inside the atoms. Therefore, by truncating the components of low $\vec{K}$, and including the corresponding PW with wave-vector $\vec{k}-\vec{K}$ in the basis, we will increasingly improve the approximation as the number of PW increases.

Certainly it would be a better approximation to start by truncating the usual tight-binding Bloch

function $h_{\vec{k}}^{TB}(\vec{r})$ [defined with $\Phi(\vec{r})$] and choosing components

$$c_{\vec{q}} = (N/\Omega)^{1/2}\Phi_{\vec{q}} , \tag{12}$$

so that

$$h_{\vec{k}}^{TB}(\vec{r}) = \frac{\sqrt{N}}{\Omega} \sum_{\vec{K}\in T} \Phi_{\vec{k}-\vec{K}} e^{i(\vec{k}-\vec{K})\cdot\vec{r}} . \tag{13}$$

However, this would not have the immense computational advantages of form (6), which permits all the terms there to be expressed in lattice sums *independent* of $\vec{k}$. Nevertheless, for some cases, higher accuracy requirements together with the necessity to keep the number of PW within reasonable limits might make it mandatory to use better Bloch functions than those defined by Eq. (6). {One way of defining these that would still give lattice sums independent of $\vec{k}$, is to take

$$c_{\vec{k}-\vec{K}} = (N/\Omega)^{1/2}[\Phi_{-\vec{K}} + \vec{k}\cdot(\nabla_{\vec{k}}\Phi_{\vec{k}-\vec{K}})_{\vec{k}=0} + \cdots]_{|} \tag{14}$$

up to some order, but, of course, the higher the order chosen, the more cumbersome and time consuming become the computations.}

For the case where a set of more than one linearly-independent localized orbital must be used, a special Bloch function $h_{\vec{k}}(\vec{r})$ must be included for each. If the cell contains several atoms, say M atoms, with position vectors $\vec{B}_\iota (\iota = 1,2,\ldots,M)$, a set $h_{\vec{k}}(\vec{r}-\vec{B}_\iota)$ $(\iota = 1,2,\ldots,M)$ of linearly-independent Bloch functions, or M-independent linear combinations of them, must be included in the basis set. All the special Bloch functions are assumed constructed with a truncated set of plane waves of wave vectors $\vec{k}-\vec{G}$ with reciprocal-lattice vectors $\vec{G}$ belonging to one and the same subset G. The basis will then contain for the same $\vec{k}$ (other than the truncated set of plane waves) a set of linearly independent Bloch functions orthogonal to them but not in general to each other. An orthogonalization procedure must then be used to get an orthonormal basis set. The use of this orthonormal basis ultimately results in a secular equation with the energy eigenvalues residing *only* on the main diagonal, and has distinct analytical and computational advantages. The selection of one particular linearly independent set of Bloch functions (over other possible equivalent sets) depends on a judicious evaluation (as far as this possible) of how well they represent the true eigenfunctions of the crystal, and how their form may help the orthogonalization procedure in efficiently producing a physically convenient orthogonal set.

Let the initial set of Bloch functions, *before the orthogonalization procedure,* be a set of linearly independent combinations defined by

$$f_{n\vec{k}}(\vec{r}) = \sum_{j=1}^{M} a_{nj} h_{j\vec{k}}(\vec{r}), \quad n = 1,2,\ldots,M \tag{15}$$

where the constants a_{nj} will be determined shortly. Here, the $h_{\iota\vec{k}}(\vec{r})$ are the Bloch functions defined-for simplicity (but without loss of generality) with only one localized orbital in one of the monatomic sublattices of the basis. Hence,

$$h_{j\vec{k}}(\vec{r}) = h_{\vec{k}}(\vec{r}-\vec{B}_j) . \tag{16}$$

Now we use the Gram-Schmidt orthogonalization procedure to get from $\{f_{n\vec{k}}\}$ an orthogonal set $\{g_{n\vec{k}}\}$. The $g_{n\vec{k}}$ have the following recursion relations:

$$|g_{1\vec{k}}\rangle = |f_{1\vec{k}}\rangle , \tag{17}$$

$$|g_{n\vec{k}}\rangle = \frac{|f_{n\vec{k}}\rangle}{\|f_{n\vec{k}}\|} - \sum_{m=0}^{n-1} \frac{|g_{m\vec{k}}\rangle}{\|g_{m\vec{k}}\|} \frac{\langle g_{m\vec{k}}|f_{n\vec{k}}\rangle}{\|g_{m\vec{k}}\|\,\|f_{n\vec{k}}\|} \quad \text{for } n = 2,3,\ldots,M ,$$

and the norms $\|g_{n\vec{k}}\|$ are given by

$$\|g_{n\vec{k}}\|^2 = 1 - \sum_{m=0}^{n-1} \frac{|\langle g_{m\vec{k}}|f_{n\vec{k}}\rangle|^2}{\|g_{m\vec{k}}\|^2\|f_{n\vec{k}}\|^2} . \tag{18}$$

These may be used in slightly modified form which subsequently reduces the numerical work. Let $g_{n\vec{k}}(\vec{r})$ be expressed first as linear combinations of $h_{j\vec{k}}(\vec{r})$.

$$|g_{n\vec{k}}\rangle = \sum_j b_{nj\vec{k}} |h_{j\vec{k}}\rangle \quad \text{for } n = 2,3,\ldots,M . \tag{19}$$

Then,

$$\langle g_{m\vec{k}}|f_{n\vec{k}}\rangle = \sum_\iota \sum_j b^*_{m\iota\vec{k}} a_{nj} \langle h_{\iota\vec{k}}|h_{j\vec{k}}\rangle , \tag{20}$$

and

$$b_{nj\vec{k}} = \frac{a_{nj}}{\|f_{n\vec{k}}\|} - \sum_{m=0}^{n-1} b_{mj\vec{k}} \frac{\langle g_{m\vec{k}}|f_{n\vec{k}}\rangle}{\|g_{m\vec{k}}\|^2\|f_{n\vec{k}}\|} \quad \text{for } n = 2,3,\ldots,M . \tag{21}$$

(Note that, in general, these are functions of $\vec{k}$.) Further,

$$\|f_{n\vec{k}}\|^2 = \sum_\iota \sum_j a^*_{n\iota} a_{nj} \langle h_{\iota\vec{k}}|h_{j\vec{k}}\rangle . \tag{22}$$

Next, let an orthonormal (incomplete) basis set $\{\Psi^{(0)}_{\alpha\vec{k}}(\vec{r}),\ \alpha\in A,\ \vec{k}\in B_0\}$ be defined by

$$\Psi^{(0)}_{\alpha\vec{k}}(\vec{r}) = \begin{cases} \Psi^{(0)}_{\vec{G}\vec{k}}(\vec{r}) = (1/\sqrt{\Omega})e^{i(\vec{k}-\vec{G})\cdot\vec{r}} & \text{for } \alpha = \vec{G}\in G, \\ \Psi^{(0)}_{n\vec{k}}(\vec{r}) = g_{n\vec{k}}(\vec{r})/\|g_{n\vec{k}}\| & \text{for } \alpha = n,\ 1\le n\le M . \end{cases} \tag{23}$$

Then, $A = G\cup\{n, 1\le n\le M\}$. The superscript zero indicates this is a basis *in which to expand the unknown variational approximations to the eigenfunctions* $\Psi_{\vec{k}}(r)$, i.e.,

$$\Psi_{\vec{k}}(\vec{r}) = \sum_{\alpha\in A} x_{\alpha\vec{k}} \Psi^{(0)}_{\alpha\vec{k}}(\vec{r}) . \tag{24}$$

Equation (24), as an expansion of the one-electron function, will be used in Sec. III as a trial function for the one-electron problem in crystals. Note that, although incomplete, the finite basis set (23) is orthonormal and contains by construction localized orbitals appropriate for the cores of the molecules forming the crystal and plane waves adequate for the *intermolecular* region. Therefore, we can expect linear combinations of them to be good approximations for the eigenfunctions of the lower bands, the accuracy improving as the number of PW in G increases, particularly for $\vec{k}$ near the boundaries of the Brillouin zone.

III APPLICATION TO THE SOLUTION OF THE ONE-ELECTRON PROBLEM IN CRYSTALS

Substituting Eq. (24) into the one-particle Schrodinger equation for the crystal, the band-structure problem reduces to

$$\sum_{\beta\in A} H_{\alpha\beta\vec{k}} x_{\beta\vec{k}} = E_{\vec{k}} x_{\alpha\vec{k}} \quad \text{for all } \alpha\in A \ , \tag{25}$$

with

$$H_{\alpha\beta\vec{k}} = \langle \Psi^{(0)}_{\alpha\vec{k}} | \hat{H} | \Psi^{(0)}_{\beta\vec{k}} \rangle \ . \tag{26}$$

Here, $\hat{H}$ is the single-particle crystal Hamiltonian. The reason why only one $\vec{k}$ is involved everywhere is the usual one, that $\hat{H}$ is a linear operator invariant under the translation group of the crystal, for which

$$\langle \Psi^{(0)}_{\alpha\vec{k}} | \hat{H} | \Psi^{(0)}_{\beta\vec{k}'} \rangle = \delta_{\vec{k}\vec{k}'} \langle \Psi^{(0)}_{\alpha\vec{k}} | \hat{H} | \Psi^{(0)}_{\beta\vec{k}} \rangle \ . \tag{27}$$

The matrix elements $H_{\alpha\beta\vec{k}}$ are given by

$$H_{\vec{G}'\vec{G}\vec{k}} = (\hbar^2/2m)(k-G)^2 \delta_{\vec{G}'\vec{G}} + U_{\vec{G}'-\vec{G}} \ , \tag{28}$$

$$H_{\vec{G}n\vec{k}} = \sum_j b_{nj\vec{k}} \langle \Psi^{(0)}_{\vec{G}\vec{k}} | \hat{H} | h_{j\vec{k}} \rangle \| g_{n\vec{k}} \|^{-1} \tag{29}$$

$$H_{n'n\vec{k}} = \sum_i \sum_j b^*_{n'j\vec{k}} b_{mi\vec{k}} \langle h_{j\vec{k}} | \hat{H} | h_{i\vec{k}} \rangle \| g_{n'\vec{k}} \|^{-1} \| g_{n\vec{k}} \|^{-1} \ , \tag{30}$$

where the plane-wave matrix element of the local one-electron crystal potential is given by

$$U_{\vec{K}} = (N/\Omega) V_{\vec{K}} \ , \tag{31}$$

with

$$V_{\vec{K}} = \int d\vec{r} e^{-i\vec{K}\cdot\vec{r}} V(\vec{r}) \tag{32}$$

and

$$U(\vec{r}) = \sum_{\vec{R}} V(\vec{r}-\vec{R}) \ . \tag{33}$$

Because of the special form [Eq. (6)] of $h_{i\vec{k}}(\vec{r})$, the products $\langle h_{i\vec{K}} | h_{j\vec{k}} \rangle$ and the matrix elements $\langle \Psi^{(0)}_{\vec{G}\vec{k}} | \hat{H} | h_{j\vec{k}} \rangle$ and $\langle h_{i\vec{K}} | \hat{H} | h_{j\vec{k}} \rangle$ can be expressed in terms of reciprocal (or reciprocal *and* direct) lattice sums which are independent of the point in the Brillouin zone (all the $\vec{k}$ dependence being factored out). For the case of only one localized orbital but with a basis of several atoms, we have

$$\langle h_{i\vec{k}} | h_{j\vec{k}} \rangle = (N/\Omega) e^{i\vec{k}\cdot(\vec{B}_i - \vec{B}_j)} S_{ij} \ , \tag{34}$$

$$\langle \Psi^{(0)}_{\vec{G}\vec{k}} | \hat{H} | h_{j\vec{k}} \rangle = (N/\Omega)^{1/2} e^{-i\vec{k}\cdot\vec{B}_j} S_{\vec{G}j} \ , \tag{35}$$

and

$$\langle h_{i\vec{k}} | \hat{H} | h_{j\vec{k}} \rangle = (N/\Omega) e^{i\vec{k}\cdot(\vec{B}_i - \vec{B}_j)} \times [(\hbar^2/2m)(S''_{ij} - 2i\vec{k}\cdot\vec{S}'_{ij} + k^2 S_{ij}) + S^D_{ij}] \ , \tag{36}$$

where

$$S_{ij} = \sum_{\vec{K}\in T} |\Phi_{\vec{K}}|^2 e^{i\vec{K}\cdot(\vec{B}_i - \vec{B}_j)} \ , \tag{37}$$

$$\vec{S}'_{ij} = \sum_{\vec{K}\in T} i\vec{K} |\Phi_{\vec{K}}|^2 e^{i\vec{K}\cdot(\vec{B}_i - \vec{B}_j)} \ , \tag{38}$$

$$S''_{ij} = \sum_{\vec{K}\in T} K^2 |\Phi_{\vec{K}}|^2 e^{i\vec{K}\cdot(\vec{B}_i - \vec{B}_j)} \ , \tag{39}$$

$$S_{\vec{G}j} = \sum_{\vec{K}\in T} \Phi_{-\vec{K}} U_{\vec{G}-\vec{K}} e^{i\vec{K}\cdot\vec{B}_j} \ , \tag{40}$$

and

$$S^D_{ij} = \sum_{\vec{K}} \sum_{T} \sum_{\vec{K}'\in T} \Phi^*_{\vec{K}'} \Phi_{\vec{K}} U_{\vec{K}'-\vec{K}} e^{i(\vec{K}'\cdot\vec{B}_i - \vec{K}\cdot\vec{B}_j)} \ . \tag{41}$$

These lattice sums can be expressed in part as *direct* lattice sums, using the convolution theorem or by application of Eq. (6b). For example,

$$S_{ij} = \sum_{\vec{R}} \langle \Phi(\vec{r}) | \Phi(\vec{r} + \vec{B}_i - \vec{B}_j - \vec{R}) \rangle - \frac{N}{\Omega} \sum_{\vec{G}\in G} |\Phi_G|^2 e^{i\vec{G}\cdot(\vec{B}_i - \vec{B}_j)} \ . \tag{42}$$

From this, $\vec{S}_{ij}$ and S''_{ij} can be obtained, respectively, by taking the gradient and the negative of the Laplacian with respect to the spatial variable. A similar result can be obtained with $S_{\vec{G}j}$ and S^D_{ij}, but here it would be of no advantage if only the Fourier transform of the potential is available.

The number of different lattice sums that must be actually computed is greatly reduced by exploiting crystal symmetries. First of all, the sums are invariant under a transposition of indices, except for $\vec{S}'_{ij}$ (which only changes sign) and $S_{\vec{G}j}$. In general a simultaneous change of $\vec{B}_i$, $\vec{B}_j$, and $\vec{G}$ (in the case of $S_{\vec{G}j}$) under the same cubic or other symmetry will also leave S_{ij}, S''_{ij}, $S_{\vec{G}j}$, and S^D_{ij} unaltered, and will take $\vec{S}'_{ij}$ into the corresponding symmetric vector. In this way, for example, the 64 S^D_{ij} sums of the $Pa3$ (or α-N_2) crystal structure[10] are reduced to only four, and the $S_{\vec{G}j}$ sums to only two for each $\vec{G}$, and in both classes of sums this leads to an enormous reduction in computational time.

Once the lattice sums are evaluated, we can proceed to solve the secular eigenvalue problem [Eq. (25)] for a particular $\vec{k}$ by first obtaining the corresponding basis set [Eq. (23)] with the help of Eqs. (19)–(22), then the matrix elements $H_{\alpha\beta\vec{k}}$ with Eqs. (28)–(30), and finally diagonalizing Eq. (25). In this way, we obtain the valence and lowest conduction bands and the coefficients $\chi_{\alpha\vec{k}}$ in the expansion of the corresponding eigenfunctions in terms of the basis set [Eq. (23)].

IV BANDS OF COMPRESSED MOLECULAR HYDROGEN

We turn now to an application of the combined-representations method to the case of solid H_2 in the α-nitrogen phase. It should be mentioned that this structure is not the only candidate for the ground-state configuration of molecular hydrogen.[11–14] We have selected it here because of the various possibilities, it is lowest in symmetry and therefore represents the most complex case numerically. Other structures have higher symmetry and the method is computationally easier to apply.

The α-N_2 structure[15] has the space group $Pa3$. It is simple cubic with a basis of four molecules. In the case of hydrogen, there are eight protons and eight electrons per primitive cell. There are sufficient electrons to fill four valence bands provided there is no overlap with conduction bands. In most of the results discussed below, it is important to note that the interproton distance (0.741 Å) is held fixed at all densities considered. We return to this point in Sec. V.

To apply (25), we need to specify the one-electron potential $U(\vec{r})$ that best represents the interaction of the electrons with the protons and with themselves Since we are mostly interested in the high-density situation we have taken this to result from the bare Coulomb interaction of the protons and screened by a Lindhard-type dielectric function. Unlike other systems, hydrogen has the advantage that the bare interactions are known precisely. The dielectric approach accounts for the bulk of the many particle effects and all residual uncertainty in $U(\vec{r})$ a reflection of exchange and correlation in the choice of the dielectric function itself. For the smallest reciprocal lattice vector that enters in (28), the dielectric function is already close to unity and such corrections are of diminishing concern as the density increases into the primary range of interest ($r_s \gtrsim 1.5$).

The bands have been calculated along the standard simple cubic directions[15,16] ΓX, MR, and $R\Gamma$ (see Fig. 2) for lattice constants of 10, 6, 5, and 4.5 bohrs. (Computational and other details may be found in the Appendix). These bands are shown

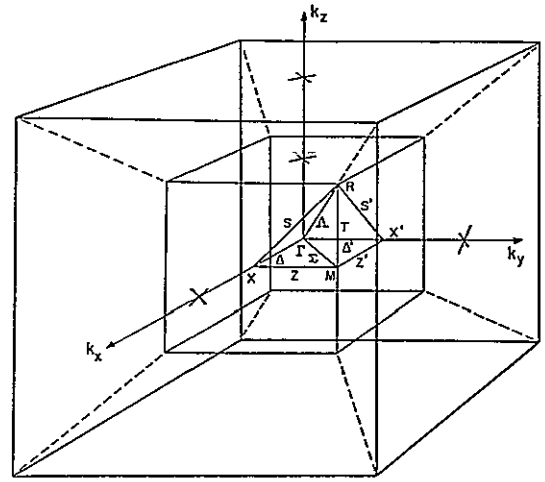


FIG. 2. The inner cube here is the Brillouin zone of the $Pa3$ (α-N_2) crystal structure The letters correspond to high-symmetry points and lines in the basic domain (unprimed) or the larger representation domain (including primes). The outer cube is limited by (100) planes, and is an example of a set G with $l_1=1$, containing, then, 27 reciprocal-lattice vectors

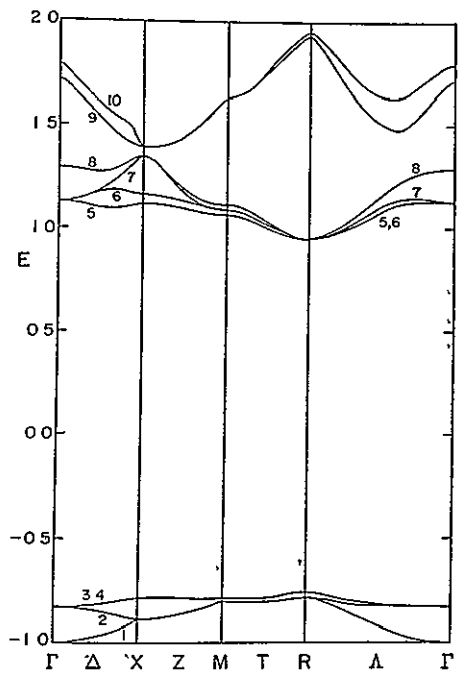


FIG 3 Band structure of the α-N_2 phase of hydrogen, with lattice constant $a=10$ bohrs or equivalently, $r_s=3.102$ (pressure zero) The energy E is normalized to $(\hbar^2/2m)(2\pi/a)^2=0$ 3948 Ry The numbers indicate, in order, the ten lowest bands calculated. Note that in order to display the overall form of the band structure the scale does not permit the resolution of certain bands. For example, in Figs. 4, 5, and 6, bands 2, 3, and 4 along $R\Gamma$ are not all degenerate as can be seen from Table I and also from this figure.

in Figs. 3–6. Figure 7 displays the empty lattice bands to which the bands at lattice constants 4.5, 5, and even 6 bohrs reveal a striking similarity. This nearly-free-electron character (at high density) gives at least *ex post facto* support to the dielectric formulation used in constructing the matrix elements of the potential.

Although the primary interest here is in the bands of highly compressed hydrogen it is worth noting that for the zero-pressure case ($a \sim 10$ bohr) we find an overall band gap of 9.2 eV. This is close to the observed value for the onset of absorption in the optical spectrum[17], it is also close to the value deduced from energy-loss experiments.[18] (Regarding the optical data, it must be said that there is, at present, disagreement in the interpretation of the data.[19,20]) Further, the overall gap agrees well with the value of 10.7 eV obtained by Zunger[20] using a truncated crystal approach, and also with the energy of the lowest-allowed optical transition obtained by the KKR method.[1]

V RESULTS AND CONCLUSIONS

We first comment on the form of the bands of highly compressed hydrogen, and then on the method used to obtain these bands

Referring to Figs. 4–6, perhaps the most inter-

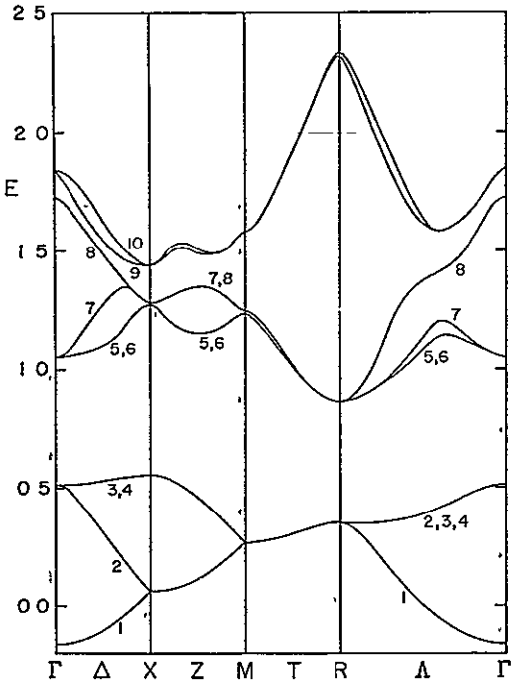


FIG 4 Band structure of the α-N_2 phase of hydrogen with lattice constant $a=6$ bohrs or equivalently, r_s =1 861 The energy E is normalized to $(\hbar^2/2m)(2\pi/a)^2$ =1 0966 Ry The numbers indicate in order the ten lowest bands calculated.

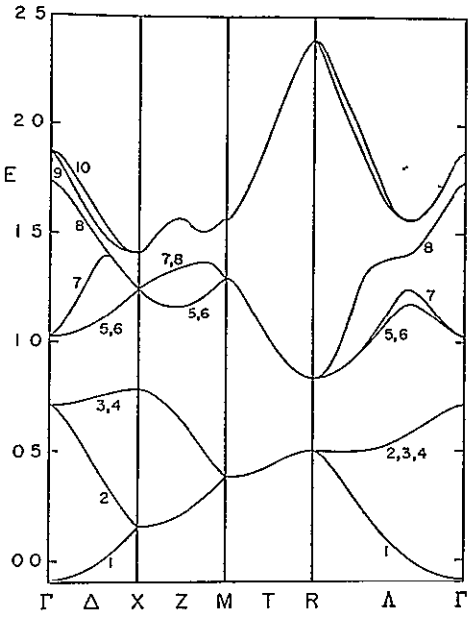


FIG 5 Band structure of the α-N_2 phase of hydrogen, with lattice constant $a=5$ bohrs or equivalently, r_s =1 551 The energy E is normalized to $(\hbar^2/2m)(2\pi/a)^2$ =1 5791 Ry The numbers indicate in order the ten lowest bands calculated

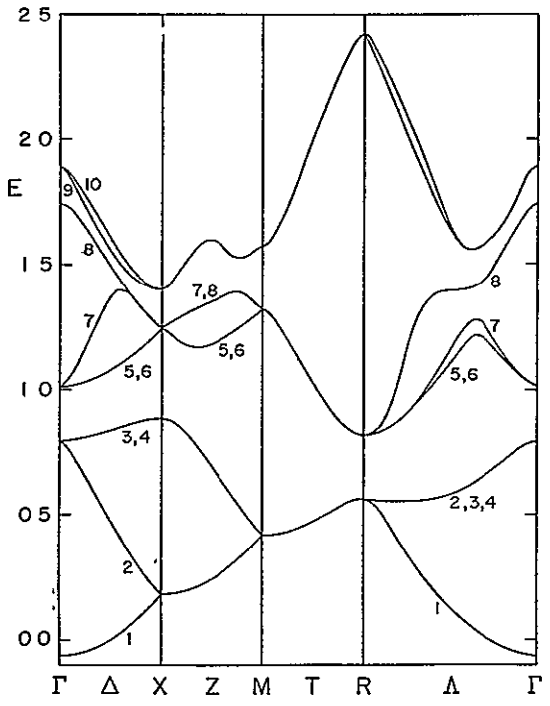


FIG 6. Band structure of the α-N_2 phase of hydrogen, with lattice constant $a=4$ 5 bohrs or equivalently, r_s =1 396 The energy E is normalized to $(\hbar^2/2m)(2\pi/a)^2$ =1 9496 Ry The numbers indicate in order the ten lowest bands calculated Note that the overall band gap in Figs. 3–5 is no longer present in this figure

esting point to emerge is the fact that the overall band gap (which becomes indirect at higher densities) vanishes at a lattice constant of $a = 4.78$ bohrs. The vanishing corresponds to the crossing of the highest valence band at X and lowest conduction-band-at R. In Fig. 8, this-gap-has-been plotted [normalized to $(\hbar^2/2m)(2\pi/a)^2$] as a function of the lattice constant a, and the critical value $a = 4.78$ is determined by linear interpolation between the gap values for $a = 4.5$ and $a = 5$ bohrs. As suggested by the calculated points, the normalized gap varies almost linearly with a. For constant interproton distance, the vanishing of the gap represents a second-order metal-insulator transition, provided, of course, that the crystalline phase of metallic hydrogen remains stable up to this point in density. The point where the molecular phase becomes metallic, i.e., $\rho = 0.83$ g/cm^3, represents a possible upper bound for the molceular density at which, for fixed interproton distance, the transition is made to a metallic state. The situation here therefore parallels somewhat the case of solid iodine in its progression with increasing pressure. As discussed recently by McMahan *et al.*[21] the metallization of iodine is evidently not a first-order transition, at least at lower pressures, and a band-overlap phenomenon preceding total pressure dissociation is therefore possible.

It is important to reemphasize that the results just described are apposite to an approximation in which the protons are both static and held at constant interproton separation within molecules. The inclusion of lattice-dynamical effects, particularly at high density, can be expected to lead to noticeable corrections, as they do for crystalline phases of metallic hydrogen.[22,23] As a decreases, we may expect the intermolecular electron density to increase in value at the expense of the intramolecular density. From a consideration of electrostatic terms alone, we would anticipate that expressed as a fraction of lattice constant, the interproton separation will increase with increasing density. A total energy calculation of the ground-state energy of molecular hydrogen will be required to determine this trend. However, a guide to the size of the effects associated with possible variations in interproton spacing $2D$ is relatively straightforward to obtain, since $2D$ is one of the basic input parameters. We have recomputed the bands of Figs. 5–7 with interproton spacing ranging between about 1.1 and 1.7 bohrs and from these have extracted by interpolation the density, for a given D, at which band overlap begins. The results are summarized in Fig. 9 as a line separating metallic from insulating regions for the $Pa3$ structure. The implication of the apparent linear trend over the limited range of parameters is that once a given band-overlap state has been attained, the interproton spacing is required to fall with unreasonable rapidity if such a state were imagined to pass once again into an insulating phase by imposing an additional increase in density.

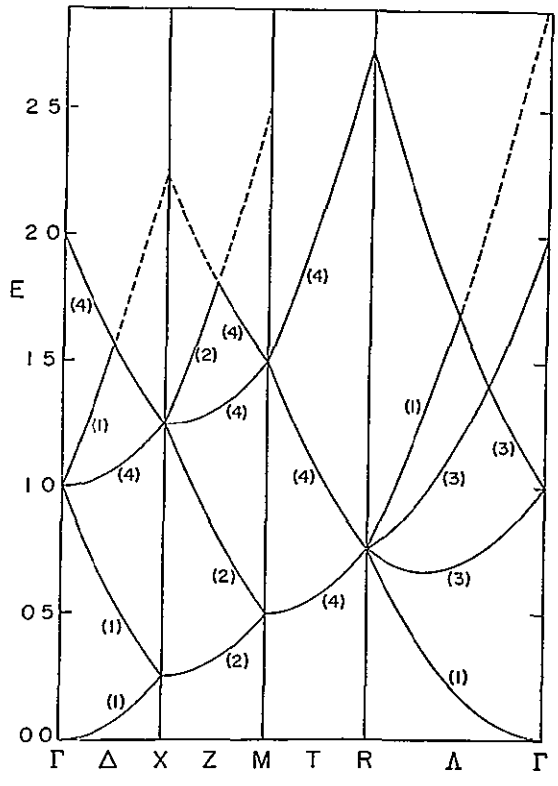


FIG 7 Band structure of the sc empty lattice The energy E is normalized to $(\hbar^2/2m)(2\pi/a)^2$ The numbers indicate the degeneracy of each band The bands drawn with a full line are the limit to which the ten lowest calculated for H_2 tend as lattice constant approaches zero

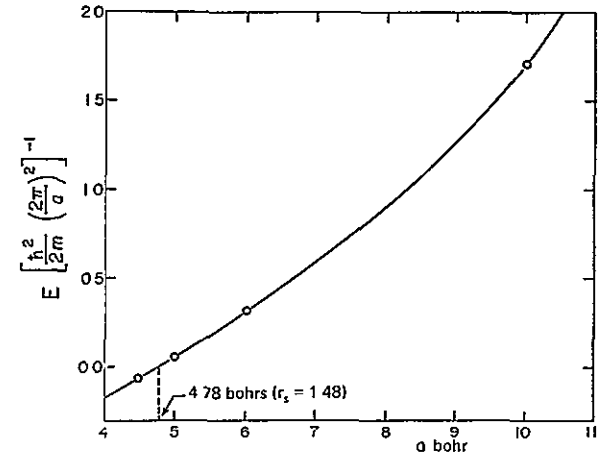


FIG 8 Energy gap normalized to $(\hbar^2/2m)(2m/a)^2$ as a function of the lattice constant a The solid line is an approximate interpolation between the calculated values, which are indicated by circles

Finally, returning to the method itself, we have shown that the subspace spanned by the orthonormal finite basis set of functions [Eq. (23)] can be expected to yield a satisfactory approximation to the one-electron eigenfunctions for electrons moving in a periodic potential. The set is of manageable size and at the same time leads to good convergence by virtue of its construction in terms of orbitals which represent both intra- and intermolecular features. This is accomplished in a rather simple way with a few plane waves and orbitals depending on $\vec{k}$ *only* through a factor $e^{i\vec{k}\cdot\vec{r}}$. It leads, however, to lattice sums independent of $\vec{k}$ when calculating the matrix elements of the secular problem [Eq. (25)], to which the band-structure problem has been reduced. As a consequence, it is necessary to evaluate the sums *only once* for a given lattice parameter and crystal structure. Even for low-symmetry structures, such as the one treated here, it is quite straightforward to obtain the necessary matrix elements in (25) for any $\vec{k}$ in the zone.

The method does not require the muffin-tin approximation to the potential, as do the standard formulations of the KKR or augmented-plane-wave methods. It is readily adaptable to systems where non-muffin-tin corrections are likely to be important, such as molecular systems or systems with complex crystal structures which can be treated, for example, by systematic correction of the KKR bands.[6] The level of analytic complexity and computational difficulty does not exceed that of such methods. When compared specifically with the OPW method, its main advantage appears to be a simpler formulation which makes no specific reference to core levels.

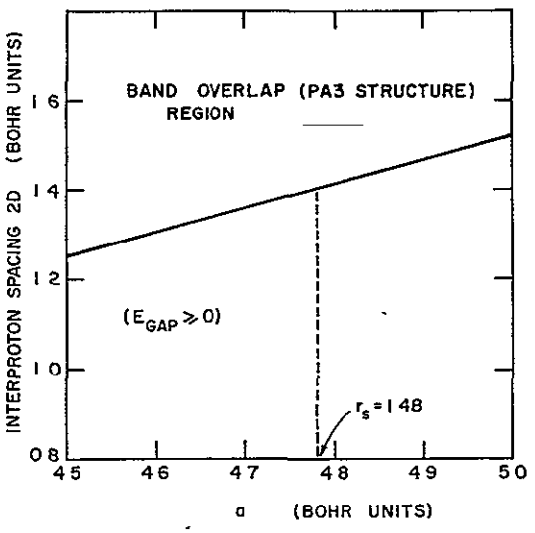


FIG. 9 A plot of the variation of interproton spacing $2D$ required, for a given density (or lattice spacing a) to lead to a vanishing of the overall band gap of H_2 in the $Pa3$ structure. The region above the line represents a ground-state metallic phase, below it the phase is insulating. Plotted vertically at $a = 4.78$ is a line which intersects the boundary at an interproton spacing 1.4 bohrs. This summarizes the band-overlap results of Fig. 4–7 (Note that for a fixed lattice constant a reduction in $2D$ tends to lead in this range of densities to a stronger one-electron potential and hence to larger band gaps.)

TABLE I Four valence bands and the lowest conduction band at selected points of the Brillouin zone and functions of l_1 and l_2 (see Appendix). Here, the lattice constant is $a = 5$ bohrs, and energies are normalized to $(\hbar^2/2m)(2\pi/a)^2 \equiv 1.5791$ Ry.

l_1	l_2	Γ	X	R
−1	4	1.3178	1.5679	2.4526
		0.7384	0.9659	1.1035
		0.7384	0.9519	1.1034
		0.7384	0.5479	0.8786
		−0.0537	0.1961	0.6948
−1	5	1.2930	1.5432	2.4314
		0.7261	0.9530	1.0875
		0.7261	0.9388	1.0875
		0.7260	0.5317	0.8619
		−0.0548	0.1951	0.6936
0	3	1.3739	1.5949	2.5006
		0.7655	0.9942	1.1387
		0.7655	0.9805	1.1386
		0.7655	0.5836	0.9032
		−0.0755	0.1737	0.6679
0	4	1.3176	1.5374	2.4526
		0.7384	0.9659	1.1034
		0.7384	0.9518	1.1033
		0.7384	0.5478	0.8668
		−0.0834	0.1656	0.6580
0	5	1.2927	1.5119	2.4275
		0.7260	0.9529	1.0874
		0.7260	0.9387	1.0873
		0.7260	0.5316	0.8505
		−0.0873	0.1616	0.6529
1	4	1.0318	1.2622	0.8442
		0.7347	0.8121	0.5407
		0.7247	0.8110	0.5407
		0.7247	0.1681	0.5381
		−0.0834	0.1592	0.5323
1	5	1.0283	1.2580	0.8428
		0.7146	0.8010	0.5344
		0.7146	0.8000	0.5344
		0.7146	0.1632	0.5318
		−0.0874	0.1549	0.5256
2	5	1.0246	1.2483	0.8318
		0.7111	0.7803	0.4994
		0.7111	0.7802	0.4990
		0.7111	0.1529	0.4990
		−0.0876	0.1504	0.4986

APPENDIX

In the calculation of the bands shown in Figs. 3–6, some parts of the lattice sums defined in Eqs. (37)–(41) were calculated in direct space and some in reciprocal space. In general, the choice is dictated by the convergence properties of the functions under consideration. For the present case, $\Phi(r)$ can be taken as a 1s orbital

$$\Phi(r) = (\alpha^3/\pi)^{1/2} e^{-\alpha r} \tag{A1}$$

with Fourier transform

$$\Phi_q^+ = (\alpha^3/\pi)^{1/2} 8\pi\alpha/(q^2+\alpha^2)^2 .$$

The direct lattice sum in Eq. (42) requires[24]

$$\langle \Phi(\vec{r}-\vec{r}') | \Phi(\vec{r}') \rangle = e^{-\alpha r}(1+\alpha r+\tfrac{1}{3}\alpha^2 r^2) ,$$

which leads to rapid convergence in direct space for the s_{ij}, S'_{ij}, and S''_{ij}. Since S_{G_i} and S^D_{ij} involve both $\Phi(r)$ (falling exponentially with r) and $V(r)$ (falling roughly as r^{-1}), a similar conclusion can be drawn about their convergence in direct space. But we also observe that in reciprocal space the convergence of the sums in (37)–(41) is also rapid since Φ_K falls as K^{-4} and U_K eventually as K^{-2}.

We turn now to general convergence properties. For the simple cubic system, we select G, on the basis of symmetry, to be all the reciprocal-lattice vectors within or on the surface of a cube centered on the origin, with faces perpendicular to the ones and aside of length $(2\pi/a)2l_1$ (see Fig. 2). Here, l_1 is a positive integer. Lattice sums in reciprocal space were computed by including only those terms with reciprocal-lattice vectors within and on the surface of a cube also centered at the origin and also having its faces normal to the axes. The side of this cube is taken as $(2\pi/a)(2l_2+1)$. [For sums in direct space, we include terms with direct lattice vectors $\vec{R}$ lying within and on the surface of a cube of side $(2l_R+1)a$]. With these definitions the number N^1_{PW} of plane waves in the basis set is $(2l_1+1)^3$: The corresponding number N^2_{PW} of plane waves in the expansions of the ortho-normal Bloch-functions of the basis is $(2l_2+1)^3 - N^1_{PW}$ (provided $l_2 > l_1$). Table I shows convergence of four valence-bands and the lowest conduction-band energies at selected points of the zone lattice constant $a=5$ bohrs. (Note that the absence of any plane waves in the expansion is symbolically designated here by the choice $l_1=-1$.) At these densities sums computed in direct space were found to converge for l_R below 4 or 5. Finally, the maximum matrix order used was 133; symmetries could be used to further reduce this number.

In constructing the Bloch functions for hydrogen, only a simple 1s orbital was used. That this is reasonable is indicated by the following Let G contain reciprocal-lattice vectors with components of magnitude $\leq 2\pi/D$, where $2D$ is the interproton distance (about $1.4a_0$ if the separation is not much affected by pressure). With this range of reciprocal-lattice vectors, the truncated set of plane waves will then represent well the electron distribution in the intermolecular region. The inclusion of 1s orbitals will give a good representation within the molecule for spatial variations in the wave function no more rapid than a change of sign in going from one proton in a molecule to the other. More rapid spatial oscillations imply the existence of higher-energy components in the intermolecular region and can therefore be neglected there altogether. *Within* a molecule, the spatial oscillations lowest in energy can be represented by atomic orbitals, the most important being 1s, 2s, 3s..., etc. To first order, these have the same leading form, i.e., $e^{(-r/a_0)}$.

---

*Work supported in part by NASA (Grant No NGR-33-010-188) and NSF through the facilities of the Materials Science Center (Grant No DMR-72-03029)

†Present address Instituto de Física, Universidad Católica de Chile, San Joaquin, Santiago, Chile

[1]R Monnier, E L. Pollock, and C Friedli, J. Phys C 7, 2467 (1974); for more details and additional considerations, see R Monnier, thesis (Université de Neuchâtel, Switzerland, 1974) (unpublished)

[2]C Friedli, Ph D thesis (Cornell University, 1975) (unpublished)

[3]F S Ham and B. Segall, Phys Rev 124, 1786 (1961)

[4]F S Ham and B. Segall, *Methods in Computational Physics· Energy Bands in Solids* (Academic, New York, 1968), Vol 8, p 251.

[5]T. Loucks, *Augmented Plane Wave Method* (Benjamin, New York, 1967)

[6]G. S Painter, Phys. Rev. B 7, 3520 (1973)

[7]See, for example, G. S Painter and D E. Ellis, Phys Rev B 12, 474 (1970).

[8]G Pastori Parravicini, L. Villa, and M Vittori, Phys. Status Solidi B 67, 345 (1975)

[9]D E Ramaker, L. Kumar, and F. E Harris, Phys Rev Lett 34, 812 (1975)

[10]E Brown and J A Krumhansl, Phys Rev 109, 30 (1958)

[11]H M James and J C Raich, Phys Rev 162, (1967)

[12]H. M. James, Phys Rev B 2, 2213 (1970)

[13]R J Lee and J C. Raich, Phys Rev B 5, 1591 (1972)

[14]H Meyer, F. Weinhaus, B Mararıglia, and R. L Mills, Phys Rev B 6, 1112 (1972).

[15]C. J. Bradley and A P Cracknell, *The Mathematical Theory of Symmetry in Solids* (Clarendon, Oxford, 1972), pp 133, 377, and 416, see also, T A Scott,

Phys Rep. 27, 99 (1976)
[16]In principle, the minimum region in the zone that must be considered is the *representation domain* and not the smaller *basic domain* As defined by Bradley and Cracknell (Ref 15), the former generates regions which will cover the Brillouin zone exactly under the action of all the operations of the point group to which the space group of the structure is isogonal. (It is obtained by taking all the point-group operations present from among the operations of that space group) The basic domain is defined in the same way except that the isogonal point group is replaced by the full holosymmetric point group of the crystal system to which the structure belongs In the present case it is the cubic system. This distinction is necessary (as is this footnote) because of the curiosity that of the 230 space groups $Pa3$ alone is anomalous and there are differences in the irreducible representations along certain ostensibly *equivalent* directions in the zone correspondingly, there should be differences in the energy eigenvalues for electrons (or phonons) for such directions In fact, this was indeed found to be the case for the bands of $Pa3$ molecular hydrogen, but the calculated differences in energy along such directions (along Z^1 and Z^{15}) are sufficiently small that for our purposes, where high densities are of primary concern, only the usual basic domain needs to be considered
[17]G Baldini, Jpn J Appl Phys Suppl 14, 613 (1965).
[18]L Schmidt, Phys Lett A 36, 87 (1971).
[19]See A Gedanken, B Raz, and J Jortner, J Chem Phys. 59, 2752 (1973), and Ref 8
[20]A Zunger, J Phys. Chem Solids, 36, 229 (1975)
[21]A K McMahan, B. L. Hoard, and M Ross, Phys Rev B 15, 726 (1977)
[22]D Straus and N W Ashcroft, Phys Rev Lett 38, 415 (1977).
[23]D Straus, thesis (Cornell University, 1977) (unpublished) (Materials Science Center Report No. 2739)
[24]J C Slater, *Quantum Theory of Molecules and Solids* (McGraw Hill, New York, 1963), Vol 1, pp 23–25

PHYSICAL REVIEW B VOLUME 16, NUMBER 12 15 DECEMBER 1977

Einstein-Kanzaki model of static and dynamic lattice relaxation: Application to vacancies in metallic hydrogen†

J F Dobson* and N. W Ashcroft
Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853
(Received 31 January 1977)

A method is proposed for calculating the formation energy of localized defects in crystalline solids with pair forces of arbitrary range The theory is most useful in the cases of small mass or high temperature for which, in addition to the usual static relaxation, changes in the lattice vibrations make a significant contribution Defect migration is not described however A self-consistent Einstein approach is used, each particle in the crystal oscillating with its *own* frequency about an average position The total free energy is minimized with respect to *all* of these frequencies and positions This minimization is made tractable by the assumption that *large* changes in frequency and position occur only for a finite number of particles near the defect, the changes for all the other particles are treated linearly The result is very similar to Kanzaki's $\vec{k}$-space "lattice statics" formalism However, instead of being 3×3 the lattice Green's function becomes a 4×4 matrix, thereby encompassing changes in Einstein frequencies as well as particle positions The method is applied to calculate the free energy of vacancy formation in metallic hydrogen

I INTRODUCTION

This paper describes a self-consistent Einstein method for calculating formation energies of localized[1] crystal defects within a k-space formalism. Changes in zero-point and/or thermal lattice vibrations are taken into account, together with static lattice relaxation. The analysis, however, is hardly more involved than that required to calculate the static relaxation alone by conventional Kanzaki[2] or Green's-function[3] techniques. One therefore has the chance to handle quite complicated particle interactions. As an example, the case of a vacancy in metallic hydrogen will be computed using a screened proton-proton interaction which is long ranged and oscillatory. The technique is self-consistent and is expected to be valid well into the high-temperature or small-mass regimes where relaxation of lattice vibrations is important; this will be referred to as "dynamic relaxation."

To begin with, a brief account is given of some previous work relevant to this problem and necessary to place the present work in perspective.

A Defects in "classical" crystals

These are crystals in which the thermal and/or zero-point particle vibrations are very small An important phenomenon associated with such a classical defect is the static relaxation of the lattice to accommodate the defect. This affects every particle in the crystal, the displacements typically fall off only as the inverse square of the distance from the defect. Descriptions of this phenomenon based on the "linear lattice statics" method have been discussed by Tewary.[3] In this method, one derives a 3×3 matrix $\underline{G}(\vec{R})$ known as the "static lattice Green's function." [Its Fourier transform $G(\vec{q})$ for $\vec{q}\neq0$ is essentially the inverse of the well-known dynamical matrix $D(\vec{q})$ which governs phonon motion.] The defect exerts a "Kanzaki force" $\vec{F}(\vec{R})$ on the lattice particles, and quantities such as the particle displacements and total strain field energy can be calculated by integrating combinations of $G(\vec{q})$ and $\vec{F}(\vec{q})$ with respect to wave number $\vec{q}$ over a Brillouin zone. In the small $\vec{q}$ limit, this theory reduces to the "elastic-continuum" model in which a handful of elastic constants completely specify the problem The theory as described so far allows only for small relaxations of the lattice, but if one has very-short-ranged forces one can also treat large displacements of a few particles near the defect (as is done, for example, in the work of Benedek and Ho[4]). Here, it is desired to treat forces whose range may be many lattice spacings, so a modified version of Benedek and Ho's method will be given. (This appears to be a new departure, even in the context of "classical" crystals which are not, however, the main concern of this paper.)

A second interesting feature associated with localized defect formation is a change in the phonon spectrum. All modes are shifted slightly in frequency, and spatially localized modes may appear with frequencies discretely separated from the rest. Theories of these effects have been given by Maradudin and co-workers,[5] and independently by Lifshitz and collaborators.[6] At finite temperatures the change in phonon modes will contribute to the defect formation energy, but the effect is small for "classical" crystals (in the sense defined above.) In the "nonclassical" regime of higher temperatures, however, the phonon modes may be

strongly modified in a complicated fashion so that a self-consistent theory is needed. Aksenov[7] has considered such a theory but omitted the static lattice relaxation around the defects, his method is therefore not suitable for examining defect formation energies, since relaxation may contribute a large fraction of the total formation energy.

B Localized defects in quantum crystals

A quantum crystal[8] is one in which particle masses and interparticle forces are small, so that large zero-point excursions occur. Static relaxation of average particle positions and modification of the particle motion are both important here The latter effect is related to changes in the phonon spectrum caused by the presence of the defect Caron[9] has considered an average t-matrix approach for calculation of the phonon spectrum in the presence of such defects taken as randomly distributed, his method does not appear to include the static deformations so important in calculating the formation energy. In an earlier paper,[10] Caron used an Einstein model in calculating defect formation energies in metallic hydrogen at $T=0$ °K. He treated the static relaxation of only a few particles near the defect and omitted the change in Einstein frequencies as negligible. A theory permitting a change in Einstein frequency for one shell of neighbors round a metallic defect was also reported recently.[11] The present work generalizes these ideas and permits relaxation of *all* positions and frequencies in a tractable formalism. More complex theories permitting such universal static and dynamic relaxation have been proposed by Varma[12] and Jacobi and Zmuidzinas[13] in terms of self-consistent phonons.[14] For quantum crystals the defect causes significant changes in all the phonon modes, making perturbation theory invalid. A fully self-consistent phonon scheme is, of course, very difficult to implement here, because the defect breaks the translational symmetry so that the spatial dependence of the phonon modes should be determined variationally along with the frequencies. Varma overcomes this problem by using a trial state in which the spatial variation of the phonon modes is obtained from a classical non self-consistent theory[5,6]; only the *frequencies* are determined self-consistently. While this enormously simplifies the algebra, the method as it stands still requires iteration of some very complicated self-consistent equations, much more involved than the ones used for self-consistent phonons in a perfect crystal.[14] In fact, Varma[12] resorted to a Debye approximation in order to obtain a practical computation procedure (Jacobi and Zmuidzinas did not indicate how one would actually solve their equations). Neither method appears to deal with the difficulty that the static relaxation of the average particle positions should be calculated *self-consistently* with the changes in vibrational motion, the static relaxation is simply added after the dynamic relaxation has already been given. The Einstein theory to be given here is quite explicit and tractable in both these respects, and has been applied to the vacancy problem in metallic hydrogen. For this case, one requires a complicated long-ranged oscillatory proton-proton interaction which would render the self-consistent phonon theories[12,13] quite unworkable without further approximation.

C Defect migration

For sufficiently high temperature or low mass, the defect can diffuse or tunnel from site to site. The tunneling at low temperature in a quantum crystal seems to have been proposed first by Hetherington.[15] Such tunneling states or "defectons" have subsequently received some theoretical attention,[16] though there does not seem to be any firm experimental evidence for them. Indeed, it appears that such tunneling phenomena will be important only for *highly* quantal crystals, if at all. Defecton motion was not considered in Refs 7, 9, 10, 11, 12, or 13, nor will it be considered here (except briefly in Sec VI). The diffusive migration of defects near the melting temperature is probably important, however, and although this phenomenon is not attacked directly here, some suggestions are made for use of the present work as *input* to a better calculation.

Set now in the context of previous work the paper is organized as follows· In Sec. II, the self-consistent Einstein picture is presented for $T=0$ °K, and its validity is discussed. In Sec. III, a generalized "lattice statics" is derived from the $T=0$ °K Einstein model. Relaxation of the zero-point motion around a defect is included on a par with static relaxation, by introducting a 4×4 "lattice Green's function" instead of the usual 3×3 one. In Sec. IV, the generalization to nonmigratory defects at $T\neq0$ °K is shown to be almost trivial if one uses the Gibbs-Bogoliubov inequality. In Sec. V the method is applied to calculate the free energy of vacancy formation in fcc metallic hydrogen for $0.6\leq r_s\leq1.5$ and $0\leq T<5000$ °K. Sec. VI contains further discussion, while Sec. VII gives conclusions.

II SELF-CONSISTENT EINSTEIN MODEL AT $T=0$ °K

The model is a very simple variational one, permitting a description of an imperfect quantum crystal at zero temperature. One minimizes the total energy over a trial N-particle crystal wave function Ψ of the Hartree type,

$$\Psi(\vec{r}_1,\ldots,\vec{r}_N)=\phi_1(\vec{r}_1-\vec{x}_1)\times\phi_2(\vec{r}_2-\vec{x}_2)\cdots\phi_N(\vec{r}_N-\vec{x}_N). \quad (1)$$

Here $\vec{r}_1,\ldots,\vec{r}_N$ are the particle coordinates, and $\vec{x}_1,\ldots,\vec{x}_N$ are the *average* particle positions. For a crystal without defects, the $\{\vec{x}_i\}$ lie on a perfect lattice, while for a crystal with defects they lie on a distorted lattice exhibiting a strain field as discussed in Sec. I. The localized functions $\{\phi_i\}$ represent the zero-point motion of the particles about their average positions $\{\vec{x}_i\}$; in general there will be a different function ϕ_i for each site i, except in the case of a perfect monatomic crystal.

An obvious deficiency of the Einstein trial state (1) is that it fails to correlate the zero-point motion of particles on different sites. Correspondingly, it does not describe any properties relating to the long-wavelength phonon modes. However, these modes contribute least of all to the total energy, so (1) should be a reasonable ansatz for calculating the total energy of defect formation. Indeed, the total energy will be especially well given compared with other quantities, since it is precisely the one which is stationary in the best trial state. (This point has already been noted by Varma,[12] who was concerned with thermal conductivities and spin relaxation rates for which an Einstein theory is less likely to be accurate.) One would seem to be justified in using (1) to obtain the total energy in situations for which a more complicated theory would prove intractable.

For simplicity of exposition in this paper the Hamiltonian operator $\hat{H}$ will be assumed to include only two-body forces

$$\hat{H}=\sum_{i=1}^{N}\frac{\hat{p}_i^2}{2M_i}+\frac{1}{2}\sum_{i\neq j}^{N}V(\hat{r}_i-\hat{r}_j), \quad (2)$$

where $\hat{r}_i$ and $\hat{p}_i$ are position and momentum operators for the ith particle. For metals, it may be necessary to include effective volume-dependent and many-body forces acting between the ions whose coordinates appear explicity in (2). The theory can be generalized in surprisingly compact form to include n-body forces; this work will be described shortly.[17]

The expectation value of the Hamiltonian (2) in the trial state (1) is

$$\langle\hat{H}\rangle=\langle\hat{T}\rangle+\langle\hat{V}\rangle=\sum_{i=1}^{N}t_i+\frac{1}{2}\sum_{i\neq j}^{N}U\{\phi_i,\vec{x}_i\ \phi_j,\vec{x}_j\}, \quad (3)$$

where

$$t_i=\frac{\hbar^2}{2M_i}\int d^3x\,|\nabla\phi_i(\vec{x})|^2 \quad (4)$$

and

$$U\{\phi_1,\vec{x}_1\cdot\phi_2,\vec{x}_2\}=\int d^3y_1 d^3y_2\,|\phi_1(\vec{y}_1)|^2\times|\phi_2(\vec{y}_2)|^2 V(\vec{x}_1+\vec{y}_1-\vec{x}_2-\vec{y}_2). \quad (5)$$

One can regard U as an effective "smeared" pair potential acting between point particles at $\vec{x}_1$ and $\vec{x}_2$. If the $\{U\}$ do not fall off rapidly with particle separation $|\vec{x}_i-\vec{x}_j|$ it may be convenient to convert to a k-space representation Defining Fourier-transformed pair potentials $V(\vec{k})$ and particle-density distributions f by the relations

$$V(\vec{r})=\frac{1}{\Omega}\sum_{\vec{k}}V(\vec{k})e^{i\vec{k}\cdot\vec{r}}, \quad (6)$$

and

$$|\phi_i(\vec{x})|^2=\frac{1}{\Omega}\sum_{\vec{k}}f_i(\vec{k})e^{i\vec{k}\cdot\vec{x}}, \quad (7)$$

one obtains from (5),

$$U\{\phi_i,\vec{x}_i\ ,\phi_j,\vec{x}_j\}=\frac{1}{\Omega}\sum_{\vec{k}}f_i(\vec{k})f_j(\vec{k})V(-\vec{k})\times e^{i\vec{k}\cdot(\vec{x}_j-\vec{x}_i)} \quad (8)$$

Here Ω is the volume and the sum $\sum_{\vec{k}}$ becomes an infinite integral $\Omega(2\pi)^{-3}\int d^3k$ in the thermodynamic limit

For a *perfect monatomic crystal*, the local wave functions ϕ_i are all the same and the average positions $\vec{x}_i$ are the perfect lattice sites $\vec{R}_i$. Thus, using the identity

$$\sum_{i=1}^{N}e^{-i\vec{k}\cdot\vec{R}_i}=N\delta_{\vec{k},\vec{g}}S(\vec{g}), \quad (9)$$

one obtains from (8) the result for the total potential energy in (3),

$$\langle\hat{V}\rangle=\frac{N}{2}\left(\frac{N}{\Omega}\sum_{\vec{g}}f^2(\vec{g})V(\vec{g})S(\vec{g})-\frac{1}{(2\pi)^3}\int d^3k f^2(\vec{k})V(\vec{k})\right). \quad (10)$$

Here the $\{\vec{g}\}$ are the reciprocal-lattice vectors and $S(\vec{g})$ is the structure factor of the unit cell [$S(\vec{g})=1$ for primitive Bravais lattices]

So far nothing has been said about the form of the local functions $\phi_i(\vec{x})$. For classical solids (those with very little particle motion) a good choice for ϕ is a Gaussian In fact, the standard Einstein model of a perfect crystal is obtained by choosing ϕ_i to be the (Gaussian) harmonic oscillator function which solves the one-particle Schrödinger equation in the spherically averaged harmonic potential set up at each site i by the other $(N-1)$ par-

ticles perfectly localized on their lattice sites. The present work is intended for *moderately* nonclassical crystals for which a Gaussian should remain a reasonable trial function[18]; however, in contrast to the classical Einstein model described above, this will be a "*self-consistent*" harmonic Einstein model in which the total energy is minimized with respect to all the harmonic frequencies. The localized trial wave functions are then generalized Gaussians

$$\phi_i(\vec{y}) = \left(\frac{M_i}{\pi\hbar}\right)^{3/4} (\det\underline{\omega}_i)^{1/4} \exp\left(\frac{-M_i\vec{y}\cdot\underline{\omega}_i\cdot\vec{y}}{2\hbar}\right), \quad (11)$$

where M_i is the particle mass. If the 3×3 frequency matrix $\underline{\omega}_i$ is of the form

$$\underline{\omega}_i = \mathrm{diag}(\omega_i, \omega_i, \omega_i), \quad (12)$$

then one has an isotropic Einstein trial state. For anistropic crystals, it may be necessary to choose different frequencies for the zero-point motion along the three Cartesian axes, so that $\underline{\omega}_i$ is of the form

$$\underline{\omega}_i = \mathrm{diag}(\omega_{i1}, \omega_{i2}, \omega_{i3}). \quad (13)$$

Regardless of crystal symmetry, it may be necessary, in the case of very strong lattice distortions, to allow some frequency matrices ω to have principal axes in directions other than the Cartesian axes, and (11) is general enough to cover this case also.

The Fourier-transformed density corresponding to (11) is

$$f_i(\vec{k}) = \exp(-\tfrac{1}{2}\vec{k}\cdot\underline{\gamma}_i\cdot\vec{k}) \quad (14)$$

[see Eq. (7)]. Here,

$$\underline{\gamma}_i = (\hbar/2M_i)\underline{\omega}_i^{-1}, \quad (15)$$

and the trace of the matrix $\underline{\gamma}_i$ is the mean-square displacement of the ith particle about its average position $\vec{x}_i$. For much of the rest of this paper, $\underline{\gamma}$ will be used in place of $\underline{\omega}$ to specify the Einstein states.

III GENERALIZED KANZAKI METHOD AT $T=0$°K

In this section, a modified lattice "statics" is described which allows for changes in the zero-point motion as well as relaxation of particle positions. It is convenient to specify both the average position $\vec{x}_j$ and mean-square Einstein amplitude matrix $\underline{\gamma}_j$ of the jth particle in terms of a single complex[19] column vector $\underline{X}_j$, to be termed the "coordinate" of particle j. Symbolically,

$$\underline{X}_j = (\vec{x}_j, -\tfrac{1}{2}i\underline{\gamma}_j). \quad (16)$$

Thus, the first three components of $\underline{X}_j$ are the Cartesian components of $\vec{x}_j$,

$$X_{j\mu} = x_{j\mu} \quad (\mu = 1, 2, 3). \quad (17)$$

The remaining components of $\underline{X}_j$ are chosen according to the degree of generality that has been built into the trial Einstein function. For example, if *isotropic* Einstein states are expected to give an adequate trial function then the mean-square amplitude matrix is specified by a single number γ_j, $\underline{\gamma}_j = \mathrm{diag}(\gamma_j, \gamma_j, \gamma_j)$; thus, $\underline{X}_j$ has dimension 4 with

$$X_{j4} = -\tfrac{1}{2}i\gamma_j. \quad (18)$$

On the other hand, for an anisotropic crystal one may need to have $\underline{\gamma}_j = \mathrm{diag}(\gamma_{j1}, \gamma_{j2}, \gamma_{j3})$ in which case $\underline{X}_j$ has dimension 6 with

$$X_{j4(5,6)} = -\tfrac{1}{2}i\gamma_{j1(2,3)}. \quad (19)$$

In the most general case, $\underline{X}_j$ can be taken as a nine-component column with the last six components

$$X_{j4,5,6,7,8,9} = -\tfrac{1}{2}i(\gamma_{j11}, \gamma_{j22}, \gamma_{j33}, \gamma_{j23} + \gamma_{j32}, \gamma_{j13} + \gamma_{j31}, \gamma_{j12} + \gamma_{j21}). \quad (20)$$

The total energy can now be written

$$\langle\hat{H}\rangle = E(\underline{X}_1, \ldots, \underline{X}_N) = \sum_{i=1}^{N} \frac{\hbar^2}{8M_i} \mathrm{Tr}\underline{\gamma}_i^{-1} + \frac{1}{2}\sum_{i\neq j}^{N} U(\underline{X}_i, \underline{X}_j), \quad (21)$$

where the *smeared* pair potential U can be found from (5) and (11) but is more compactly expressed in k space by using (8) with (14)

$$U(\underline{X}_i, \underline{X}_j) = \frac{1}{(2\pi)^3}\int d^3k\, V(\vec{k}) \exp[-\tfrac{1}{2}\vec{k}\cdot(\underline{\gamma}_i + \underline{\gamma}_j)\cdot\vec{k} - i\vec{k}\cdot(\vec{x}_i - \vec{x}_j)] = \frac{1}{(2\pi)^3}\int d^3k\, V(\vec{k}) \exp[-i\underline{K}\cdot(\underline{X}_i - \underline{X}_j^*)]. \quad (22)$$

Here a higher-dimensional wave number, symbolically

$$\underline{K} = (\vec{k}, \vec{k}\vec{k}) \quad (23)$$

has been introduced. To be specific, its components are

$$\underline{K} = (k_x, k_y, k_z, k^2)$$

or

$$\underline{K} = (k_x, k_y, k_z, k_x^2, k_y^2, k_z^2)$$

or (24)

$$\underline{K} = (k_x, k_y, k_z, k_x^2, k_y^2, k_z^2, k_yk_z, k_zk_x, k_xk_y),$$

in the three cases previously outlined in defining $\underline{X}$. (A caution: $\vec{k} = \vec{k}_1 + \vec{k}_2$ does not imply $\underline{K} = \underline{K}_1 + \underline{K}_2$)

The essence of the proposed method is that in an inhomogeneous situation one can *explicitly* minimize the energy (21) with respect to *all* the particle "coordinates" $\{\underline{X}_i\}$, provided that the deviations from the perfect-crystal "coordinates" $\underline{X}_i^0 = (\vec{R}_i, -\frac{1}{2}i\underline{\gamma}_0)$ can be treated linearly *except* at a finite number of sites. These few nonlinear sites near the defect constitute the *"core"* (c) of the defect, the remaining sites will be termed the "bulk" sites. The calculation proceeds in several steps:

(a) The Einstein frequency $\underline{\omega}_0 = \hbar\underline{\gamma}_0^{-1}/2M$ is found which minimizes the energy for a perfect crystal.

(b) The core sites are assigned "coordinates" $\{\underline{X}_i : i \in c\}$, which are later treated as explicit variational parameters. The energy cost of creating the core is computed with the bulk "coordinates" $\{\underline{X}_i : i \notin c\}$ held at the perfect crystal values $\{\underline{X}_i^0\}$.

(c) The bulk "coordinates" are given linear increments $\underline{X}_i \to \underline{X}_i^0 + \underline{\xi}_i$, $i \in c$, the $\underline{\xi}_i$ are chosen to minimize the total energy subject to the given core "coordinates." This minimization is achieved explicitly in $\vec{k}$ space by a generalization of the Green's-function method of lattice statics[3] changes in the zero-point motion are computed self-consistently with static relaxation, by making the lattice Green's function a 4×4 (or 6×6, or 9×9) matrix instead of a 3×3 one as in conventional lattice statics.[2,3]

(d) The relaxed crystal energy is now known as a function of the core "coordinates." Finally, these core "coordinates" are chosen to give an overall minimum energy.

These four steps will now be discussed in detail.

Step (a) The perfect crystal

In the perfect crystal all sites have the same Einstein oscillator width $\underline{\gamma}_0$, and the particle coordinates are

$$\underline{X}_j^0 = (\vec{R}_j, -\tfrac{1}{2}i\underline{\gamma}_0). \tag{25}$$

Using (21), (22), and (25) and defining an equilibrium form U^0 of the smeared potential,

$$U^0(\vec{R}) = U(\underline{X}_0, \underline{X}_1), \tag{26}$$

with $\underline{X}_0 = (\vec{0}, -\frac{1}{2}i\underline{\gamma}_0)$ and $\underline{X}_1 = (\vec{R}, -\frac{1}{2}i\underline{\gamma}_0)$, one obtains the total energy per particle as a sum over direct lattice vectors $\vec{R}$,

$$E^{(0)}(\underline{\gamma}_0) = \frac{E(\underline{X}_1^0, \ldots, \underline{X}_N^0)}{N} = \frac{\hbar^2}{8M}\,\mathrm{Tr}(\underline{\gamma}_0^{-1}) + \frac{1}{2}\sum_{\vec{R}\neq\vec{0}} U^0(\vec{R}) \tag{27}$$

With the aid of (9) this can also be expressed in reciprocal space

$$E^{(0)}(\underline{\gamma}_0) = \frac{\hbar^2}{8M}\,\mathrm{Tr}(\underline{\gamma}_0^{-1}) + \frac{1}{2}\left[\frac{N}{\Omega}\sum_{\vec{g}} S(\vec{g})U^0(\vec{g}) - (2\pi)^{-3}\int d^3k\, U^0(\vec{k})\right] \tag{28}$$

Here $U^0(\vec{k})$ is the Fourier transform of the smeared equilibrium pair potential,

$$U^0(\vec{k}) = V(\vec{k})\exp(-\vec{k}\cdot\underline{\gamma}_0\cdot\vec{k}) \tag{29}$$

This step of the calculation is completed by choosing $\underline{\gamma}_0$ to minimize (27) or (28), whichever is more convenient.

Step (b) Formation of the core

The details of this step depend on the type of local defect being considered. In the case of vacancy or interstitial formation at constant[20] particle number N, a particle presumably has to be transferred to or from the surface. To begin with, this process will be considered without any relaxation of the coordinates $\underline{X}_i$ of the other $(N-1)$ particles. There appears to be some ambiguity concerning the energy involved in this process, and it has been the subject of some dispute.[21] This controversy will not be entered into here, since it arises in any calculation involving vacancies or interstitials, and has nothing specifically to do with the new features of the model under consideration. For definiteness, the results of Caron[10] for vacancy and interstitial formation without relaxation will be adopted, they have the advantage of being calculated in the framework of the Einstein model and so are compatible with the present work. The constant-volume method will be adopted. It is certainly more convenient in the case of metals, since the "volume-dependent forces" are not brought into play; at any rate, Caron[10] has shown that the overall results at constant pressure must be the same. For reference, his result for vacancies will be quoted in the notation of the present work

$$\Delta E_{0\,(\text{vacancy})} = -\frac{1}{2}\sum_{\vec{R}\neq\vec{0}}\left(\frac{1}{3}\vec{R}\cdot\frac{\partial U^0(\vec{R})}{\partial\vec{R}} + U^0(\vec{R})\right), \tag{30}$$

in the same notation as (27), where the two terms come from compression of the lattice at constant volume to create new sites, followed by removal of particles from those sites. This result can also be expressed in k space after an integration by parts

$$\Delta E_{0\,(\text{vacancy})} = \frac{1}{6}\,\frac{N}{\Omega}\sum_{\vec{g}} S(\vec{g})\vec{g}\cdot\frac{dU^0(\vec{g})}{d\vec{g}} + \frac{1}{2(2\pi)^3}\int d^3k\, U^0(\vec{k}). \tag{31}$$

The considerations given so far in this step were

special to vacancies and interstitials whose formation involved transfer of a particle to or from the surface. The second half of the present step involves a deformation of the core region ($\underline{X}_i^0 \to \underline{X}_i$, $i \in c$), this applies equally to all kinds of local defects including for example mass defects and vacancy interstitial pairs[22] as well as the above types considered The core deformation costs an energy

$$\Delta E_{\text{core}} = \sum_{i\in c} \frac{\hbar^2}{8M_i} \mathrm{Tr}(\underline{\gamma}_i^{-1} - \underline{\gamma}_0^{-1}) + \sum_{i\in c, j\in c} [U(\underline{X}_i, \underline{X}_j^0) - U(\underline{X}_i^0, \underline{X}_j^0)] + \frac{1}{2} \sum_{i\neq j\in c} [U(\underline{X}_i, \underline{X}_j) - U(\underline{X}_i^0, \underline{X}_j^0)], \tag{32}$$

with U given by (22). The second sum in (32) is an unrestricted sum on j over the direct lattice, with the core sites excluded. With the aid of (9), it can be reduced to a finite direct lattice sum, plus a reciprocal-lattice sum:

$$\sum_{i\in c, j\in c} [U(\underline{X}_i, \underline{X}_j^0) - U(\underline{X}_i^0, \underline{X}_j^0)] = -\sum_{i\in c, j\in c^*} [U(\underline{X}_i, \underline{X}_j^0) - U(\underline{X}_i^0, \underline{X}_j^0)]$$
$$+ \frac{N}{\Omega} \sum_{i\in c} \sum_{\vec{g}} S(\vec{g}) V(\vec{g}) [\exp(-\tfrac{1}{2}\vec{g}\cdot(\underline{\gamma}_0 + \underline{\gamma}_i)\cdot\vec{g} - i\vec{g}\cdot\mathrm{x}_i) - \exp(-\vec{g}\cdot\underline{\gamma}_0\cdot\vec{g} - i\vec{g}\cdot\vec{\mathrm{R}}_i)] \tag{33}$$

(Here the perfect lattice sites in the core are denoted c^*.)

Step (c) Linear relaxation including zero-point motion

The major results of the present work are contained in this step. The bulk particles are now taken to undergo small "coordinate" changes

$$\underline{\xi}_i = \underline{X}_i - \underline{X}_i^0. \tag{34}$$

The first three components of $\underline{\xi}_i$ give the deviations of the average positions $\vec{\mathrm{x}}_i$ from the perfect lattice sites $\vec{\mathrm{R}}_i$ (i.e., they specify the conventional strain field) while the higher components $\{\xi_{i\mu}, \mu > 3\}$ measure the changes in the mean-square displacements $\underline{\gamma}_i$ around the average positions.

If the defect were not present, the energy required to produce the bulk distortions $\{\underline{\xi}_i, i\in c\}$ could be expanded to second order in the $\{\underline{\xi}_i\}$

$$\Delta E_{\{\underline{\xi}_i\}}^{(\text{no defect})} = \frac{1}{2} \sum_{\mu,\nu=1}^{4} \sum_{i,j\in C} D_{\mu\nu}(\vec{\mathrm{R}}_i - \vec{\mathrm{R}}_j) \times \xi_{i\mu}\xi_{j\nu}^* + O(\xi^3). \tag{35}$$

(Summation on μ and ν will henceforth be implicit for repeated indices.) In (35), D is the Taylor-series expansion coefficient

$$D_{\mu\nu}(\vec{\mathrm{R}}_i - \vec{\mathrm{R}}_j) = \left.\frac{\partial^2 E(\underline{X}_1, \cdots, \underline{X}_N)}{\partial X_{i\mu} \partial X_{j\nu}^*}\right|_0 \tag{36}$$

The energy $E(\underline{X}_1, \ldots, \underline{X}_N)$ is defined in (21) and the subscript 0 means that the $\underline{X}_i$ are set to the perfect lattice values $\underline{X}_i^0 = (\vec{\mathrm{R}}_i, -\frac{1}{2}i\underline{\gamma}_0)$ after differentiation No linear term is present in (35) since

$$\left.\frac{\partial E(\underline{X}_1, \cdots, \underline{X}_N)}{\partial X_{i\mu}}\right|_0 = 0 \tag{37}$$

For $\mu = 1, 2, 3$, (37) is just the statement that the perfect crystal is in equilibrium under the pair forces at the chosen volume or pressure, this is automatic for systems with inversion symmetry. For $\mu > 3$, (37) is not automatic but is satisfied because $\underline{\gamma}_0$ has been chosen in step (a) to guarantee precisely this stationarity of the energy

The zone Fourier transform of (36) is defined by the direct lattice sum

$$D_{\mu\nu}(\vec{\mathrm{q}}) = \sum_{\vec{\mathrm{R}}} D_{\mu\nu}(\vec{\mathrm{R}}) e^{-i\vec{\mathrm{q}}\vec{\mathrm{R}}}, \tag{38}$$

with inversion formula

$$D_{\mu\nu}(\vec{\mathrm{R}}) = \frac{1}{N} \sum_{\vec{\mathrm{q}}\in z} D_{\mu\nu}(q) e^{i\vec{\mathrm{q}}\cdot\vec{\mathrm{R}}}, \tag{39}$$

where Z is the Brillouin zone The matrix $D_{\mu\nu}(\vec{\mathrm{q}})$ is a 4×4 (or 6×6, or 9×9) generalization of the ordinary 3×3 dynamical matrix which appears in the classic theories of lattice statics and dynamics.[3,5,6] The upper 3×3 block of $\underline{D}$ is just the ordinary dynamical matrix evaluated using the "smeared" particle interaction U^0 [Eq. (26) or (29)] in place of the pair potential V [Eq. (2) or (6)]. The remaining components of $\underline{D}$ (those with $\mu > 3$ or $\nu > 3$) have no counterpart in the classic theory. they express the response of the Einstein zero-point motion to disturbances in the crystal.[23]

Explicit expressions for the generalized dynamical matrix can be obtained by application of the definition (36) and direct differentiation of the energy formula (21). For simplicity only the isotropic case will be written, so that $\underline{D}$ is 4×4 and $\underline{\gamma}_0 = \mathrm{diag}(\gamma_0, \gamma_0, \gamma_0)$. The result can be written in terms of direct lattice sums on the smeared potential U^0 of Eq. (26),

$$D_{\mu\nu}(\vec{q}) = \sum_{R\neq 0} (1-e^{-i\vec{q}\cdot\vec{R}}) \frac{\partial^2}{\partial R_\mu dR_\nu} U^0(\vec{R}) \quad (\mu,\nu\leq 3),$$

$$D_{\mu 4}(\vec{q}) = D_{4\mu}(\vec{q}) = -2i \sum_{\vec{R}\neq 0} e^{-i\vec{q}\cdot\vec{R}} \frac{\partial^2}{\partial R_\mu \partial\gamma_0} U^0(\vec{R}) \quad (\mu\leq 3) \tag{40}$$

and

$$D_{44}(\vec{q}) = 4\sum_{\vec{R}\neq 0} (1+e^{-i\vec{q}\cdot\vec{R}}) \frac{\partial^2}{\partial\gamma_0^2} U^0(\vec{R}) + \frac{3\hbar^2}{M\gamma_0^3}.$$

With the aid of Eq (9), these results can also be written in k space, with μ, ν running from 1 to 3,

$$D_{\mu\nu}(\vec{q}) = \frac{N}{\Omega}\sum_{\vec{g}} S(\vec{g})[(\vec{g}+\vec{q})_\mu(\vec{g}+\vec{q})_\nu U^0[\vec{g}+\vec{q}] - g_\mu g_\nu U^0(\vec{g})],$$

$$D_{\mu 4}(\vec{q}) = D_{4\mu}(\vec{q}) = \frac{N}{\Omega}\sum_{\vec{g}} S(\vec{g})(\vec{g}+\vec{q})^2 \times (\vec{g}+\vec{q})_\mu U^0(\vec{g}+\vec{q}),$$

$$D_{44}(\vec{q}) = \frac{3\hbar^2}{M\gamma_0^3} + \frac{N}{\Omega}\sum_{\vec{g}} S(\vec{g})[|\vec{g}+\vec{q}|^4 U^0(\vec{g}+\vec{q}) + g^4 U^0(\vec{g})] - 2(2\pi)^{-3}\int d^3k\, k^4 U^0(\vec{k}). \tag{41}$$

The last expression exhibits $\underline{D}(\vec{q})$ as a real symmetric matrix

Equation (35) was derived for small distortions in the bulk of an otherwise perfect crystal In the presence of a defect core, these bulk distortions will cost an extra energy

$$-\sum_{i=1}^{N} F_\mu(\vec{R}_i)\xi^*_{i\mu} + O(\xi^2), \tag{42}$$

where the "generalized Kanzaki force" F_μ is given for $i \in c$ by

$$F_\mu(\vec{R}_i) = -\sum_{j\in c} \frac{\partial U(\underline{X}_j, \underline{X}_i)}{\partial X^*_{i\mu}}\bigg|_{\underline{X}_i=\underline{X}_i^0} + \sum_{j\in c^*} \frac{\partial U(\underline{X}_j^0, \underline{X}_i)}{\partial X^*_{i\mu}}\bigg|_{\underline{X}_i=\underline{X}_i^0}, \tag{43}$$

while

$$F_\mu(\vec{R}_i) = 0 \text{ for } i \in c^*. \tag{44}$$

(c^* again refers to the perfect lattice sites inside the core region: for a vacancy, c^* has one more site than c) The neglect of terms higher than the first order in (42) is a standard approximation of lattice statics known as the "first Kanzaki approximation "[2] The total energy associated with the bulk distortions $\{\underline{\xi}_i: i \in c\}$ is now

$$\frac{1}{2}\sum_{i,j\notin c} D_{\mu\nu}(\vec{R}_i - \vec{R}_j)\xi_{i\mu}\xi^*_{j\mu} - \sum_{i\in c} F_\mu(\vec{R}_i)\xi^*_{i\mu}. \tag{45}$$

This is minimized when the $\{\underline{\xi}_i\}$ satisfy

$$\sum_{j\in c} D_{\mu\nu}(\vec{R}_i - \vec{R}_j)\xi_{j\nu} = F_\mu(\vec{R}_i), \quad i \in c. \tag{46}$$

If the $\{\underline{\xi}_i\}$ satisfy (46) then (45) can be simplified to give the minimum energy

$$\Delta E_{\text{bulk}} = -\frac{1}{2}\Big|\sum_i F_\mu(\vec{R}_i)\xi^*_{i\mu}\Big| = -\frac{1}{2}\sum_i F^*_\mu(\vec{R}_i)\xi_{i\mu} \tag{47}$$

The restriction $i \in c$ has been dropped in the sum (47) since F_μ is defined in (44) to be zero for $i \in c^*$. This is very convenient since (47) can now be directly transcribed into k space as

$$\Delta E_{\text{bulk}} = -\frac{1}{2N}\sum_{\vec{q}\in Z} F_\mu(\vec{q})\xi^*_\mu(\vec{q}). \tag{48}$$

The Fourier-transformed Kanzaki force is obtained from (43) and (44) with the help of (9) and (22);

$$F_\mu(\vec{q}) = \sum_{\vec{R}_i} F_\mu(\vec{R}_i)e^{-i\vec{q}\cdot\vec{R}_i} = \frac{N}{\Omega}\sum_{\vec{g}} iQ_\mu S(\vec{g})V(\vec{g}+\vec{q})\Big[\sum_{j\in c^*}\exp(-i\underline{Q}\cdot\underline{X}'_j) - \sum_{j\in c}\exp(-i\underline{Q}\cdot\underline{X}''_j)\Big] + \sum_{i\in c^*}\Big(\sum_{j\in c}\frac{\partial U(\underline{X}_i^0, \underline{X}_j)}{\partial X^{0\,*}_{i\mu}} - \sum_{j\in c^*}\frac{\partial U(\underline{X}_i^0, \underline{X}_j^0)}{\partial X^{0\,*}_{i\mu}}\Big)e^{-i\vec{q}\cdot\vec{R}_i} \tag{49}$$

[The four-columns $\underline{Q} = (\vec{g}+\vec{q}, (\vec{g}+\vec{q})^2)$, $\underline{X}'_j = (\vec{R}_j, -i\gamma_0)$, and $X''_j = (\vec{x}_j, -\frac{1}{2}i(\gamma_0+\gamma_j))$ are introduced for brevity.]

It remains to find the bulk distortions $\{\underline{\xi}_i\}$, which are the solutions of (46). If it were not for the restriction $j \in c$ on the left-hand side, Eq (46) would be solved trivially by Fourier transformation Although the translational invariance is spoiled by this restriction, an exact k-space solution is still possible at expense of solving a small matrix (of order $4n$, where n is the number of sites in the core) If the pair forces determining $\underline{D}(\vec{R})$ are very short-ranged the solution of (46) can be performed by the matrix partitioning method of Benedek and Ho [1] An alternative approach is given here, since the assumption of short-ranged forces is *not* being made

The solution proceeds by first augmenting (46)

with a set of equations on the core sites

$$\sum_{j\in c} D_{\mu\nu}(\vec{R}_i-\vec{R}_j)\xi_{j\nu}=f_\mu(\vec{R}_i)\,,\quad i\in c^*, \tag{50}$$

where f_μ is to be determined Equations (46) and (50) can next be combined to give a single equation on the entire perfect lattice

$$\sum_{\text{all}\,j} D_{\mu\nu}(\vec{R}_i-\vec{R}_j)\Xi_{j\nu}=\mathfrak{F}_\mu(\vec{R}_i)\quad(\text{all}\,i)\,, \tag{51}$$

where

$$\Xi_{j\nu}=\begin{cases}0, & j\in c^*,\\ \xi_{j\nu}, & j\notin c^*\end{cases}$$

and

$$\mathfrak{F}_\mu(\vec{R}_i)=\begin{cases}f_\mu(\vec{R}_i), & i\in c^*,\\ F_\mu(\vec{R}_i), & i\notin c^*\end{cases} \tag{52}$$

Since (51) has a translationally invariant kernel and is valid on all sites, its solution (with periodic boundary conditions) is trivial in k space;

$$\Xi_\mu(\vec{k})=D^{-1}_{\mu\nu}(\vec{k})\,\mathfrak{F}_\nu(\vec{k})\quad(\vec{k}\neq 0)\,, \tag{53}$$

where $\underline{D}^{-1}$ means the 4×4 reciprocal matrix. If one defines the *generalized lattice Green's function* G by

$$G_{\mu\nu}(\vec{R})=\frac{1}{N}\sum_{k\in Z} D^{-1}_{\mu\nu}(k)e^{i\vec{K}\cdot\vec{r}}\,, \tag{54}$$

then (53) becomes, in real space,

$$\Xi_\mu(\vec{R}_i)=\sum_{\text{all}\,j} G_{\mu\nu}(\vec{R}_i-\vec{R}_j)\mathfrak{F}_\nu(\vec{R}_j)\,. \tag{55}$$

This is more conveniently represented in a $4N\times4N$ matrix notation as

$$\begin{pmatrix}\underline{0}\\ \underline{\xi}\end{pmatrix}=\underline{G}\begin{pmatrix}\underline{f}_1\\ \underline{F}_2\end{pmatrix}=\begin{pmatrix}\underline{g}_{11} & \underline{g}_{12}\\ \underline{g}_{21} & \underline{g}_{22}\end{pmatrix}\begin{pmatrix}\underline{f}_1\\ \underline{F}_2\end{pmatrix}, \tag{56}$$

where the matrices have been partitioned so that, for example, $\underline{g}_{11}$ is a $4n\times4n$ submatrix, it is the restriction of $\underline{G}$ to the core sites, $i\in c^*$. Expansion of the matrix product in (56) gives two equations, the first of which, namely

$$\underline{0}=\underline{g}_{11}\underline{f}_1+\underline{g}_{12}\underline{F}_2\,,$$

gives the unknown "force" $\underline{f}_1$,

$$\underline{f}_1=-\underline{g}^{-1}_{11}\underline{g}_{12}\underline{F}_2$$

The second part of (56) now gives the desired solution

$$\underline{\xi}=\underline{g}_{21}\underline{f}_1+\underline{g}_{22}\underline{F}_2=(-\underline{g}_{21}\underline{g}^{-1}_{11}\underline{g}_{12}+\underline{g}_{22})\underline{F}_2\,. \tag{57}$$

The energy associated with the linear relaxation of all the "bulk" particles is now found by putting (57) into (47)

$$\Delta E_{\text{bulk}}=-\tfrac{1}{2}\underline{F}^*_2(\underline{g}_{22}-\underline{g}_{21}\underline{g}^{-1}_{11}\underline{g}_{12})\underline{F}_2\,. \tag{58}$$

Noting from (44) that $F_\mu(\vec{R}_i)$ vanishes for $i\in c^*$ [and that the Fourier transform $F_\mu(\vec{k})$, Eq. (49), is computed with this in mind] one can extend (58) to a full matrix equation on the whole space

$$\Delta E_{\text{bulk}}=-\tfrac{1}{2}\underline{F}^*(\underline{G}-\underline{G}\underline{g}^{-1}\underline{G})\underline{F}\,, \tag{59}$$

where

$$\underline{g}=\begin{pmatrix}\underline{g}_{11} & \underline{0}\\ \underline{0} & \underline{0}\end{pmatrix}$$

and

$$\underline{F}=\begin{pmatrix}\underline{0}\\ \underline{F}_2\end{pmatrix}$$

This can be transcribed into k space as

$$\Delta E_{\text{bulk}}=-\frac{1}{2}\left(\frac{1}{N}\sum_{\vec{k}\in Z}F^*_\mu(\vec{k})D^{-1}_{\mu\nu}(\vec{k})F_\nu(\vec{k})-\sum_{i\in c^*}f^0_\mu(\vec{R}_i)\xi^{0*}_\mu(R_i)\right)\,, \tag{60}$$

where

$$\xi^0_\mu(\vec{R}_i)=\frac{1}{N}\sum_{\vec{k}\in Z}e^{-i\vec{k}\cdot\vec{R}_i}D^{-1}_{\mu\nu}(\vec{k})F_\nu(\vec{k})\,,$$

and f^0 is the solution of a small equation

$$\sum_{j\in c^*}G_{\mu\nu}(\vec{R}_i-\vec{R}_j)f^0_\nu(\vec{R}_j)=\xi^0_\mu(\vec{R}_i)\quad\text{for}\ i\in c^* \tag{61}$$

Equations (59) or (60) completely solve the problem of minimizing the bulk distortion energy (45). To evaluate (59) or (60), one need only compute the generalized dynamical matrix $\underline{D}[k]$ from (40) or (41), the Green's function $G(\vec{R})$ from (54), and the Kanzaki force F_μ from (43) or (49). Then the problem reduces to solution of the small matrix equation (61), equivalent to finding $\underline{g}^{-1}$. In practice, this solution is often dramatically simplified by point symmetry at the defect site.

The solution (60) becomes especially simple in the case of *completely linear* vacancy relaxation. Here the strongly distorted core c is a null set, so that $\Delta E_{\text{core}}=0$, while (in the case of a vacancy) c^* consists of the single site from which a particle is missing This site can be taken as the origin. It is evident from symmetry that the on-site generalized Green's function $G_{4\nu}(\vec{O})$ is zero when $\nu=1$, 2, 3; this can be verified formally by inspection of (54) and (41). Further, the first three components of the distortion vector $\underline{\xi}^0(\vec{O})$ also vanish because of point symmetry at the defect site. Hence, from (61), $f^0_\mu(\vec{O})=0$ except for $\mu=4$. Specifically,

$$f^0_\mu(\vec{O}) = \begin{cases} \xi^0_4(\vec{O})/G_{44}(\vec{O}), & \mu = 4 , \\ 0 , & \mu = 1, 2, 3 \end{cases} \tag{62}$$

The expression for the Kanzaki force F_μ can also be simplified when there is no strongly distorted core. Equations (43) and (49) become

$$F_\mu(\vec{R}_t) = \begin{cases} \left. \dfrac{\partial U(\underline{X}^0_0, \underline{X}_t)}{\partial X^*_{t\mu}} \right|_{\underline{X}_t = \underline{X}^0_t}, & \vec{R}_t \neq \vec{O} , \\ 0 , & \vec{R}_t = \vec{O} \end{cases} \tag{63}$$

and

$$F_\mu(\vec{q}) = \frac{N}{\Omega} \sum_{\vec{g}} i Q_\mu S(\vec{g}) U^0(\vec{g}+\vec{q}) - \frac{i\delta_{\mu 4}}{(2\pi)^3} \int d^3k\, k^2 U^0(\vec{k}) , \tag{64}$$

with Q_μ defined as in (49).

Now, noting that the operation $(1/N)\sum_{\vec{k}\in Z}$ is just the *Brillouin-zone average* $\langle\ \rangle_{BZ}$, we reduce (60) to the form

$$\Delta E_{bulk} = -\tfrac{1}{2}\left(\langle F^*_\mu(\vec{k}) \xi^0_\mu(\vec{k}) \rangle_{BZ} - \frac{|\langle \xi^0_4(\vec{k}) \rangle_{BZ}|^2}{\langle (\underline{D}^{-1})_{44}(\vec{k}) \rangle_{BZ}} \right), \tag{65}$$

with

$$\underline{\xi}^0(\vec{k}) = \underline{D}^{-1}(\vec{k}) \underline{F}(\vec{k}) . \tag{66}$$

Equation (65) is now the total distortion energy including changes in vibrational energy. Only the undistorted formation energy (30) or (31) need be added to obtain the total vacancy formation energy in this fully linear approximation.

It is also worth noting that in the absence of any relaxation of the Einstein frequencies one would have the usual 3×3 lattice statics formalism. The result for the linear distortion energy would then be

$$\Delta E^{(3)}_{bulk} = -\tfrac{1}{2} \langle F_\alpha(\vec{k}) (D^{(3)-1})_{\alpha\beta}(\vec{k}) F_\beta(\vec{k}) \rangle_{BZ} , \tag{67}$$

where α and β are summed from 1 to 3 and $\underline{D}^{(3)}$ is the usual 3×3 dynamical matrix evaluated with the smeared pair potential U^0. [$\underline{D}^{(3)}$ is the upper 3×3 block of the 4×4 dynamical matrix $\underline{D}$ defined in (40) or (41).]

Step (d) Final minimization

The total energy required to form the defect with a core configuration $\{\vec{X}_t\ t \in c\}$ is

$$\Delta E(\{\underline{X}_t\ t \in c\}) = \Delta E_0 + \Delta E_{core} + \Delta E_{bulk} , \tag{68}$$

where the individual terms are given by (30) and (31) (for the case of a vacancy[24]), (32)–(33), and (59)–(60). If ΔE_{bulk} is a substantial fraction of the formation energy (which it can be even though the bulk distortions $\underline{\xi}_t$ were treated linearly) then it will be necessary to treat ΔE_{core} and ΔE_{bulk} together when searching for the optimal core "coordinates" $\{\underline{X}_t, t \in c\}$. On the other hand, if ΔE_{bulk} is formally regarded as a small quantity then only ΔE_{core} need be varied explicitly, and ΔE_{bulk} can be evaluated afterwards using the core coordinates $\underline{X}_t$ so determined, changes caused by varying the two together are formally of second order. Whether or not the full procedure is necessary can only be decided in specific cases, according to the accuracy required.

In either case, the appropriately computed minimum of (68) is the final answer for the defect formation energy at $T = 0$ °K within the Einstein-Kanzaki model.

IV EXTENSION TO FINITE TEMPERATURE

If the migration of defects between lattice sites is ignored, the generalization of Sec. III to $T \neq 0$ °K is straightforward. The procedure is essentially to minimize the *free* energy F over an Einstein trial state. This imprecise notion can be formalized by using the Gibbs-Bogoliubov inequality[25]

$$F \leq F_{trial} = F^0 + \langle \hat{H} - \hat{H}_0 \rangle_0 . \tag{69}$$

Here $\hat{H}$ is the actual Hamiltonian [i.e., (2)], $\hat{H}_0$ is an exactly soluble trial Hamiltonian, and $\langle\ \rangle_0$ is an exact quantum thermal average over $\hat{H}_0$. In (69), F^0 is the exact free energy for H_0.

The trial Hamiltonian appropriate to an Einstein picture is

$$\hat{H}_0(\hat{\vec{r}}_1, \ldots, \hat{\vec{r}}_N) = \sum_{i=1}^{N} \tfrac{1}{2} M_i (\hat{\vec{r}}_i - \vec{\bar{x}}_i) \cdot \underline{\omega}_i^2 \cdot (\hat{\vec{r}}_i - \vec{\bar{x}}_i) + \sum_{i=1}^{N} \frac{\hat{\vec{p}}_i^2}{2M_i} = \hat{V}_0 + \hat{T} . \tag{70}$$

Here, as in Sec. III, the variational parameters $\{\vec{\bar{x}}_i\}$ and $\underline{\omega}_i$ are average particle positions and Einstein frequency matrices. The idea is to choose these parameters to minimize F_{trial}.

Since the kinetic energy term is common to $\hat{H}$ and $\hat{H}_0$, (69) can be rewritten

$$F \leq F_{trial} = F^0 - \langle \hat{V}_0 \rangle_0 + \langle \hat{V} \rangle_0 \tag{71}$$

The terms of (71) can be evaluated explicitly by using standard harmonic-oscillator results.[26] As before it is convenient to define a "coordinate" $\underline{X}_j = (\vec{\bar{x}}_j, -\tfrac{1}{2} i \underline{\gamma}_j)$ where the mean-square excursion matrix $\underline{\gamma}$ is now evaluated at finite temperature:

$$\underline{\gamma}_i = \langle (\hat{\vec{r}}_i - \vec{\bar{x}}_i)(\hat{\vec{r}}_i - \vec{\bar{x}}_i) \rangle_T = \frac{\hbar}{2M_i} \underline{\omega}_i^{-1} \coth\left(\frac{\hbar}{2k_B T} \underline{\omega}_i \right) . \tag{72}$$

The trial free energy is

$$F_{\text{trial}}(\underline{X}_1,\cdots,\underline{X}_N)$$
$$=\sum_{i=1}^{N}\mathrm{Tr}\left\{k_BT\ln\left[2\sinh\left(\frac{\hbar}{2k_BT}\underline{\omega}_i\right)\right]-\tfrac{1}{4}\hbar\underline{\omega}_i\coth\left(\frac{\hbar}{2k_BT}\underline{\omega}_i\right)\right\}+\frac{1}{2}\sum_{i\neq j}^{N}U(\underline{X}_i,\underline{X}_j), \quad (73)$$

with $U(\underline{X}_i,\underline{X}_j)$ defined, in terms of the $\{\underline{\gamma}_i\}$ and $\{\vec{\mathrm{x}}_i\}$, by (22). Equation (73) is typical of the way in which the theory generalizes to finite temperature. The potential energy terms U depend only on the probability distribution of an Einstein particle, and hence have the same dependence on mean-square displacement $\underline{\gamma}$ as the corresponding $T=0$ °K terms. [Note, however that $\underline{\gamma}$ is now related to the frequency $\underline{\omega}$ by (72)]. On the other hand, the kinetic energy terms *do* change when one goes to finite temperature, as summarized in Table I. The quantity t appearing in the last column of the table is the "kinetic" energy (free energy minus potential energy) of an Einstein oscillator, and is given by

$$t(\underline{\omega})=k_BT\,\mathrm{Tr}[\ln(2\sinh\underline{y})-\tfrac{1}{2}\underline{y}\coth\underline{y}], \quad (74)$$

with

$$\underline{y}=(\hbar/2k_BT)\underline{\omega}. \quad (75)$$

V EXAMPLE VACANCY IN METALLIC HYDROGEN

As an example of the method developed in Secs. I–IV, the free energy of vacancy formation in fcc metallic hydrogen will now be calculated. The problem is of interest because of the possible role of localized defects in the decay of metastable metallic hydrogen. This system may exhibit high-temperature superconductivity[27] (or other forms or electronic or nuclear order) and also has astrophysical significance.[28]

Although pressures in excess of a megabar are apparently required to form the metal,[29] it has been conjectured that it may be metastable relative to the molecular phase when the pressure is decreased to more easily maintained values, perhaps on the order of tens of kilobars or less. Surface-decay of the metal[30] can probably be controlled by a suitable coating, and in the absence of unstable phonon modes down to moderate pressures[31] it appears that the principal decay modes will involve some kind of crystal defect. A likely decay mode is the formation of hydrogen atoms or molecules inside voids or aggregates of vacancies. The prototype of this configuration is the monovacancy, which will be studied here. If this can be understood properly, one can hope to proceed to more complicated defects. A very low or negative vacancy formation energy would be suggestive of an instability; it will be shown here that no such instability towards monovacancies occurs in low-temperature fcc metallic hydrogen.

The zero-temperature vacancy formation energy in fcc metallic hydrogen has already been estimated by Caron,[10] who used an Einstein model for the proton zero-point motion. As noted above he permitted relaxation of the positions of a few protons near the vacancy, but took as negligible any changes in the zero-point motion during defect formation. However, Straus and Ashcroft[31] recently showed that the proton zero-point motion is crucial in determining the structure of a perfect crystal of metallic hydrogen. One might therefore suspect that changes in the zero-point motion, not necessarily localized near the defect, would be important in the vacancy formation process.

The motivation for the present calculation, then is twofold: (a) one would like to know if there are any slight but poorly localized changes in zero-point motion which might significantly affect the free energy of formation, both at zero temperature and above; and (b) such a calculation will demonstrate that the present Einstein-Kanzaki method

TABLE I Modifications for $T\neq 0$ °K. [See Eqs. (74) and (75) for definitions of $t(\underline{\omega})$ and y.] All equations in Sec. III remain unchanged when one goes to finite temperature, except those listed here

$T=0$ quantity	$T\neq 0$ quantity	$T=0$ kinetic term	$T\neq 0$ kinetic term
E, Eq. (21)	F_{trial}	$\frac{\hbar^2}{8M_i}\mathrm{Tr}\underline{\gamma}_i^{-1}$	$t(\underline{\omega}_i)$
$E^{(0)}$, Eqs. (27), (28)	$F^{(0)}$	$\frac{\hbar^2}{8M}\mathrm{Tr}\underline{\gamma}_0^{-1}$	$t(\underline{\omega}_0)$
ΔE_{core}, Eq. (32)	ΔF_{core}	$\frac{\hbar^2}{8M_i}\mathrm{Tr}(\underline{\gamma}_i^{-1}-\underline{\gamma}_0^{-1})$	$t(\underline{\omega}_i)-t(\underline{\omega}_0)$
$D_{44}[\vec{\mathrm{q}}]$, Eqs. (40), (41)	$D_{44}[\vec{\mathrm{q}}]$	$\frac{3\hbar^2}{M\gamma_0^3}$	$24\hbar^{-1}M^2\omega_0^3\left(\coth y_0+\frac{y_0}{\sinh y_0}\right)^{-1}$

can be carried out in practice for a complicated long-ranged oscillatory pair potential.

The model used for metallic hydrogen was an fcc lattice of vibrating protons[31] interacting via an electronically screened pair potential, given in k space by

$$V(k)=\begin{cases}4\pi e^2/k^2\epsilon(k/2k_F), & \vec{k}\neq\vec{0},\\ 0, & \vec{k}=\vec{0}.\end{cases} \tag{76}$$

Here ϵ is the linear dielectric function of the electron gas, and the vanishing of the screened potential for $\vec{k}=\vec{0}$ reflects the overall charge neutrality of the system. A screened pair-potential model of this kind neglects two phenomena:

(i) Even in the linear screening regime the energy depends on the total volume (i.e., there are "volume-dependent forces"). Here the formation energy at *constant volume* will be considered, so that this effect does not enter into the calculation.

(ii) Nonlinear distortions in the electron gas, caused by the protons, will give rise to *many-proton* forces as well as pair forces. While the present formalism can in fact be generalized to cover many-particle potentials,[17] the proton motion can be expected to wash out such three-body and higher effective forces to a large degree. (This phenomenon is discussed by Straus[31] in connection with the perfect metallic hydrogen crystal.) Here only pair potentials were considered, as was the case in Caron's[10] work.

The linear electron-gas dielectric function was taken to be the Hubbard[32] version, as modified by Geldart and Vosko[33] so as to satisfy the compressibility sum rule. Thus

$$\epsilon(x)=1+\Lambda(x)g(x)\alpha r_s/\pi x^2, \tag{77}$$

$$g(x)=\tfrac{1}{2}+\frac{1}{4x}(1-x^2)\ln\left|\frac{1+x}{1-x}\right|, \tag{78}$$

$$\Lambda(x)=\left(1-\frac{g(x)}{r_K+(2\pi/\alpha r_s)x^2}\right)^{-1}. \tag{79}$$

Here r_s is the usual Wigner-Seitz radius measured in Bohr radii, and $\alpha=(4/9\pi)^{1/3}$. In (79), $r_K=K/(K-K_0)$ is determined by the ratio of the true electron gas compressibility K to the compressibility K_0 of the noninteracting electron gas. The value of r_K was taken as that obtained by differentiating the Vashishta-Singwi electron-gas energy formula.[34,35] Thus,

$$r_K^{-1}=1-\frac{K_0}{K}=\frac{\alpha}{\pi r_s}\left(1+\frac{0.0335}{2}\pi\alpha r_s+\frac{0.02\pi\alpha r_s^2}{3}\,\frac{0.1+2r_s}{(0.1+r_s)^3}\right). \tag{80}$$

The above form of the dielectric function has the advantage of being analytic while yielding a good "compressibility limit"[35] as $k\to 0$. It is important to treat ϵ accurately near $k=2k_F$, since the behavior there is responsible for the long-ranged Friedel oscillations of the real-space screened potential. However, for values $k/2k_F\gtrsim 1.5$, which are safely away from the $2k_F$ singularity, it is convenient to know the large-wave-number asymptotic expansion of (76)–(79),

$$\epsilon^{-1}(x)\underset{x\to\infty}{\sim}1-\frac{\alpha r_s}{3\pi}x^{-4}-\frac{\alpha r_s}{15\pi}x^{-6}-\left[\frac{\alpha r_s}{35\pi}-\frac{1}{18}\left(\frac{\alpha r_s}{2}\right)^2\right]x^{-8}-\left[\frac{\alpha r_s}{63\pi}-\frac{1}{45}\left(\frac{\alpha r_s}{\pi}\right)^2-\frac{r_K}{36}\left(\frac{\alpha r_s}{\pi}\right)^3\right]x^{-10}+O(x^{-12}). \tag{81}$$

Since the interest here is principally in any slight but long-ranged disturbance to the proton motion, the completely linear relaxation is sufficient. There is thus no strongly perturbed "core," and the set of sites c^* is just the vacant site at the origin. The free energy ΔF of vacancy formation was found by working through Sec. III step by step, using the electronically screened and motionally smeared proton-proton potential

$$U^0(\vec{k})=\begin{cases}\dfrac{4\pi e^2 e^{-\gamma_0 k^2}}{k^2\epsilon(k/2k_F)}, & \vec{k}\neq\vec{0},\\ 0, & \vec{k}=\vec{0}\end{cases} \tag{82}$$

The necessary steps are now listed for reference, together with some relevant details of numerical methods.

(a) The perfect crystal free energy $F^{(0)}(\gamma_0)$ was found from Eq. (28), modified as in Table I when $T\neq 0$ °K. γ_0 was chosen to minimize $F^{(0)}$.

(b) The free energy $\Delta F_{0,\mathrm{vac}}$ required to form the vacancy without any lattice distortion was found from Eq. (31): The "core distortion" energy ΔF_{core} is, of course, zero.

(c) The total free energy of linear distortion ΔF_{bulk}, including relaxation of lattice vibrations, was found from (65). For comparison, the corresponding result $\Delta F^{(3)}_{\mathrm{bulk}}$ *without* relaxation of lattice vibrations was found from (67). The Brillouin-zone averages specified in (65) and (67) were performed using the ten-term "special point" prescription given for fcc lattices by Chadi and Cohen.[36] The quantities needed in these zone averages were the generalized dynamical matrix $\underline{D}(\vec{k})$ [found from Eqs. (41) with a $T\neq 0$ °K modification as in Table I for D_{44}] and the Kanzaki force vector $\underline{F}(\vec{k})$ [found from Eq. (64)].

(d) The total free energy of formation was found as

$$\Delta F = \Delta F_{0,\mathrm{vac}} + \Delta F_{\mathrm{bulk}}, \tag{83}$$

there being no need for a separate variation of nonlinear core parameters in this purely linear distortion calculation.

Steps (a), (b), and (c) involved numerical evaluation of reciprocal lattice sums of the form

$$\sum_{\vec{g}\neq\vec{0}} |\vec{g}+\vec{q}|^{-n}\epsilon^{-1}\left(\frac{|\vec{g}+\vec{q}|}{2k_F}\right)e^{-\gamma_0(\vec{g}+\vec{q})^2} \tag{84}$$

and integrals of the form

$$\int d^3k\, k^{-n}\epsilon^{-1}\left(\frac{k}{2k_F}\right) e^{-\gamma_0 k^2}, \tag{85}$$

where n is a small positive integer. Since $\epsilon^{-1}\sim 1$ and $\gamma_0>0$, (84) and (85) are formally convergent at large wave number. However, the value of γ_0 is small enough that convergence was too slow for direct numerical evaluation in practice. This difficulty was circumvented by using the five-term asymptotic expansion (81) for $k/2k_F>\lambda$, where $\lambda \sim 1.5$ (The final results were independent of λ over a considerable range, of course.) The advantage of this is that one now has finite sums and integrals, plus infinite sums and integrals of the form

$$\sum_{\vec{g}} |\vec{g}+\vec{g}|^{-p}e^{-\gamma_0(\vec{g}+\vec{q})^2}, \qquad \int d^3k\, k^{-p}e^{-\gamma_0 k^2} \tag{86}$$

for several positive values of p The integrals can be reduced to known special functions and combined with terms which arise when the sums are converted using modified Ewald methods. (See the work of Cohen and Keffer[37] for details of the Ewald methods). The outcome is that one has a number of fairly complicated but rapidly convergent sums. The results of the calculations are shown in Tables II and III and in Fig. 1.

TABLE II Calculated vacancy formation energy, ΔE (rydbergs), in metallic hydrogen at $T=0$ °K and constant volume $\Omega=\frac{4}{3}\pi N(r_s a_0)^3$ The quantities listed are, from left to right, the Wigner-Seitz radius r_s, the rms proton excursion in units of the nearest-neighbor separation, the energy $\Delta E_{0,\mathrm{vac}}$ required to form a vacancy without any lattice relaxation, the linear lattice relaxation energy $\Delta E^{(3)}_{\mathrm{bulk}}$ *ignoring* changes in Einstein frequencies, the linear lattice relaxation energy ΔE_{bulk} *including* changes in the Einstein frequencies, the total vacancy formation energy ΔE in the linear approximation All energies are in rydbergs.

r_s	$(3\gamma_0)^{1/2}/d_{nn}$	$\Delta E_{0,\mathrm{vac}}$	$\Delta E^{(3)}_{\mathrm{bulk}}$	ΔE_{bulk}	ΔE
0.6	$0\,13_7$	$+0\,57_1$	$-0\,27_3$	$-0\,27_5$	$+0\,29_6$
0.7	$0\,13_6$	$+0.42_6$	-0.19_9	$-0\,20_0$	$+0\,22_6$
0 8	$0\,13_5$	$+0\,32_4$	$-0\,14_9$	$-0\,15_0$	$+0\,17_4$
0 9	0.13_5	$+0.25_0$	$-0\,11_4$	$-0\,11_5$	$+0\,13_5$
1.0	$0\,13_5$	$+0\,19_5$	$-0\,08_8$	$-0\,09_0$	$+0\,10_5$
1.1	0.13_6	$+0\,15_3$	$-0\,07_0$	-0.07_1	$+0.08_2$
1 2	$0\,13_7$	$+0\,12_0$	$-0\,05_6$	$-0\,05_7$	$+0\,06_2$
1 3	0.14_0	$+0\,09_4$	$-0\,04_5$	$-0\,04_6$	$+0\,04_8$
1 4	$0\,14_2$	$+0\,07_3$	$-0\,03_7$	$-0\,03_8$	$+0\,03_5$
1 5	0.14_5	$+0.05_7$	$-0\,03_1$	$-0\,03_2$	$+0.02_5$

Table II shows that the vacancy formation energy is *not* significantly altered by relaxation of the proton motion at $T=0$ °K in the range of densities $1.0\lesssim r_s\lesssim 1.5$ relevant to metastable metallic hydrogen. This is seen by comparing columns 4 and 5 of Table II, which give the relaxation energy, first without, then with relaxation of zero-point motion ($\Delta E^{(3)}_{\mathrm{bulk}}$ and ΔE_{bulk}).

Figure I shows that, in the same range of densities, the present results do not differ appreciably from Caron's[10] values. This is actually a valuable check on both calculations, since Caron used a real-space method in which only a few neighbors were relaxed nonlinearly, while the present results came from a linear k-space method which

TABLE III Temperature dependence of free energy of vacancy formation, ΔF (rydbergs), in fcc metallic hydrogen at $r_s=1\,36$ The quantities listed are, from left to right, the temperature T°K, the rms proton excursions as a fraction of nearest-neighbor distance, the free energy $\Delta F_{0,\mathrm{vac}}$ required to form the vacancy without lattice relaxation, the linear lattice relaxation energy $\Delta F^{(3)}_{\mathrm{bulk}}$ *ignoring* changes in proton motion, the linear lattice relaxation energy ΔF_{bulk} *including* changes in proton motion, the total free energy ΔF required to form a vacancy, the concentration $\exp(-\Delta F/k_BT)$ of vacancies in an independent random vacancy model

T(°K)	$(3\gamma_0)^{1/2}/d_{nn}$	$\Delta F_{0,\mathrm{vac}}$	$\Delta F^{(3)}_{\mathrm{bulk}}$	ΔF_{bulk}	ΔF	$e^{-\Delta F/k_BT}$
0	$0\,14_1$	$+0\,080_9$	$-0\,040_3$	$-0\,041_3$	$+0\,038_6$	0
1 000	$0\,15_5$	$+0\,084_8$	$-0\,039_2$	$-0\,041_2$	$+0\,043_6$	$0\,001_0$
2 000	$0\,18_5$	$+0\,094_4$	$-0\,037_8$	$-0\,040_3$	$+0\,054_1$	$0\,01_4$
3 000	$0\,20_1$	$+0\,103_8$	-0.037_5	$-0\,040_5$	$+0\,063_3$	$0\,03_6$
4 000	$0\,22_8$	$+0\,112_1$	$-0\,037_8$	$-0\,041_3$	$+0\,070_8$	$0\,06_1$
5 000	$0\,24_5$	$+0.120_0$	$-0\,038_5$	$-0\,042_5$	$+0\,077_5$	0.08_6
10 000	0.30_9	$+0\,154_0$	$-0\,045_1$	$-0\,051_6$	$+0\,102_4$	0.19_9

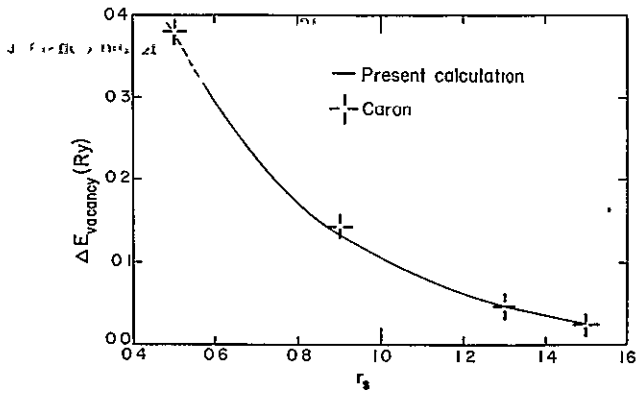


FIG. 1. Vacancy formation energy ΔE in fcc metallic hydrogen at $T=0$ °K. Present calculation is compared with the results of Caron, obtained from Table VI and Figs. 12 of Ref. 10.

included static and dynamic relaxation of every proton in the crystal

Table III shows the effect of raising the temperature The quantities given are now Helmholtz free energies ΔF for the formation of an isolated vacancy, ignoring the entropy of vacancy location In the model of randomly placed noninteracting vacancies, the equilibrium vacancy concentration is then

$$C(T)=\exp[-\Delta F(T)/k_BT], \tag{87}$$

which is tabulated in the last column of Table III Two trends are noticeable in Table III.

(i) The free energy of formation *increases* with temperature, so that the concentration of vacancies does not rise as fast as exponentially when the temperature increases. For example, if the crystal still exists at 5000 °K, the present model gives a concentration $C(T)=9\%$ of vacancies, whereas the usual model involving the $T=0$ °K, formation energy $\Delta F(0)$ would give

$$C_0(T)\doteq\exp[-\Delta F(0)/k_BT]\equiv 29\%$$

of vacancies, a very substantial difference.

(ii) With increased temperature the dynamic relaxation becomes more important, so that at 5000 °K the dynamic relaxation energy is 10% of the total relaxation energy.

Actually it is likely that the crystal has melted by 5000 °K In addition to the 9% vacancy concentration shown in column 7 of Table III, the notion of melting by a few thousand degrees is also supported by column 2 which gives the Lindemann[38] ratio r_L (This is the ratio of rms particle excursion to nearest-neighbor distance. in classical crystals r_L is about $\frac{1}{7}$ at melting) In hydrogen at $r_s=1.36$, r_L is already[31] $\frac{1}{7}$ at $T=0$ °K, and doubles by 5000 °K It should be borne in mind, however, that in computer experiments on quantum crystals with soft-cored pair potentials, Chester[39] *et al.* found values of r_L significantly above $\frac{1}{7}$ at melting.

VI FURTHER POSSIBILITIES

Existence of the "generalized lattice statics" approach suggests that an even simpler theory might be available; the $\vec{q}\to 0$ limit of the present work should yield a "jiggling elastic continuum" model, related to the present microscopic approach in the same way that the usual elastic continuum model is related to the conventional[2,3] lattice statics. This is currently under investigation.

An effect which was not directly considered in Sec. I–IV (and is missing also from Refs 9–13) is the migration of point defects. This will be important in classical crystals near melting,[7] and may occur in quantum crystals with small enough mass to permit significant tunneling [15,16] In the classical case, a crude way to remedy the omission is simply to assume that the total defect free energy (at low defect concentration $C=n/N$) is

$$F/N=C\Delta F+TS/N, \tag{88}$$

where $S\sim k\ln{}^NC_n$ is the configurational entropy associated with the possible sites occupied by n defects, and ΔF is the free energy of defect formation as calculated in Sec. III–IV. Minimization of (88) leads to the equilbrium defect concentration $C(T)$ given in Eq. (87), and tabulated for metallic hydrogen in Table III. A more complete approach would be to use a lattice gas picture of the defect crystal [7] Here ΔF will play the role of a temperature-dependent chemical potential for defects and in this context one could also use the generalized lattice statics to calculate an effective interaction between defects,[5] as mediated by their static *and dynamic* strain fields.

In the case of quantal defect tunneling, the relaxation described in the present work can significally lower the tunneling probability or even cause self-trapping [40] To describe this case one can invoke a tight-binding Hubbard model for defect motion, in which the hopping matrix element t is to be computed from an overlap integral between two of the Einstein states (as used in this paper), one with the defect on a neighboring site relative to the other. The formation energy ΔF computed above will then play the role of a site occupation energy ϵ_i.

Thus, the present model may be useful even near melting or for highly quantal crystals, in the sense that it provides an explicit method of computing the input parameters to more sophisticated theories.

VII CONCLUSIONS

It has been shown in detail how to use the Einstein model to calculate formation energies of crystalline defects, including relaxation of zero-point and thermal lattice motions as well as the usual static lattice deformation. Relaxation of every site in the crystal was explicitly calculated by a generalization of the Kanzaki method; static and dynamic contributions appeared self-consistently in the same 4×4 matrix formalism.

The method is substantially easier to carry out in full than the self-consistent phonon approaches,[12,13] which require specific use of localized phonon modes as well as a separate minimization for static relaxation. On the other hand, the method is more complete than previous Einstein theories of defects[10,11] in which only a few particles are usually relaxed.

Application to metallic hydrogen shows that the method is a practical means of calculating static and dynamic relaxation in the case of complicated long-ranged pair forces. For hydrogen in the density range $0.6 \lesssim r_s \lesssim 1.5$, it was possible to show that dynamical relaxation does *not* upset the stability of the system to vacancy formation as might perhaps have been supposed.

ACKNOWLEDGMENTS

The authors would like to thank Geoffrey Chester, Bob Guyer, David Straus, David Stevenson, and Carl Kukkonen for interesting discussions.

---

†Work supported by NASA under Grant No. NGR-33-010-188.

*Present address: School of Science, Griffith University, Nathan, Queensland 4111, Australia.

[1] A "localized" defect means that a point or small cluster is the focus of the defect, rather than a line or plane of sites. Of course, even a "localized" defect has a long-ranged disturbing influence which is in fact the main concern here.

[2] E. Kanzaki, J. Chem. Phys. Solids 2, 24 (1957).

[3] V. K. Tewary, Adv. Phys. 22, 757 (1973).

[4] R. Benedek and P. S. Ho, J. Phys. F 3, 1285 (1973).

[5] A. A. Maradudin, E. W. Montroll, G. H. Weiss, and I. P. Ipatova, *Theory of Lattice Dynamics in the Harmonic Approximation*, 2nd ed. (Academic, New York, 1971).

[6] I. M. Lifshitz and A. M. Kosevitch, Rep. Prog. Phys. 29, 217 (1966).

[7] V. L. Aksenov, Fiz. Tverd. Tela 14, 1986 (1972) [Sov. Phys.-Solid State 14, 1718 (1973)].

[8] R. A. Guyer, *Solid State Physics* (Academic, New York, 1969), Vol. 23, p. 413.

[9] L. G. Caron, Phys. Rev. B 13, 4545 (1976).

[10] L. G. Caron, Phys. Rev. B 9, 5025 (1974).

[11] C. H. Leung, M. J. Stott, and W. H. Young, APS meeting, Atlanta, Ga., 29 March–1 April, 1976 (unpublished).

[12] C. M. Varma, Phys. Rev. Lett. 23, 778 (1969); Phys. Rev. A 4, 313 (1971).

[13] N. Jacobi and J. S. Zmuidzinas, Report, Jet Propulsion Laboratory, Pasadena, Calif., 1974 (unpublished).

[14] T. R. Koehler, Phys. Rev. 165, 942 (1968).

[15] J. H. Hetherington, Phys. Rev. 176, 231 (1968).

[16] See, for example, Ref. 8. Also, A. F. Andreev and I. M. Lifshitz, Zh. Eksp. Teor. Fiz. 56, 2057 (1969) [Sov. Phys.-JETP 29, 1107 (1969)]; D. I. Pushkarov, *ibid.* 68, 1471 (1975); 41, 735 (1976); J. F. Dobson, Phys. Lett. A 57, 73 (1976).

[17] J. F. Dobson (unpublished).

[18] There is no need to restrict $\phi_t(x)$ to be a Gaussian, in general, and indeed for highly nonclassical crystals one may gain an advantage over the usual self-consistent phonon scheme (Ref. 14) by allowing more freedom in the $\{\phi_t\}$ within an Einstein scheme. This point has been taken up by D. Rosenwald [Phys. Rev. 154, 160 (1967)] who concluded, however, that in comparing the approximations involved the neglect of correlation is far more difficult to justify than the restriction of the trial functions to Gaussian form.

[19] There is no deep significance to the fact that some elements of $\underline{X}$ are imaginary. It turns out to be convenient in $\vec{k}$ space, giving rise to *real* 4×4 lattice Green's functions.

[20] In the case of vacancy or interstitial formation in metals it is highly advantageous to treat the process at constant particle number N, because overall removal or addition of an ion entails removal or addition of a conduction electron also: this is most inconvenient when one is treating the conduction sea as merely a screening agent for the ions.

[21] See conference proceedings *Interatomic Potentials and Simulation of Lattice Defects*, edited by P. C. Gehlen, J. R. Beeler, and R. I. Jaffee (Plenum, New York, 1972), discussion of this point is on p. 456 and elsewhere.

[22] In the case of a vacancy-interstitial pair, one of the displacements $|\vec{x}_t - \vec{R}_t|$ is on the order of a lattice spacing or more.

[23] Terms like D_{41} couple static and dynamic relaxation.

[24] ΔE_0 is zero for mass defects or vacancy-interstitial pairs.

[25] A proof of this result is given in R. P. Feynmann, *Statistical Mechanics* (Benjamin, Reading, Mass., 1972), p. 67.

[26] See A. Messiah, *Quantum Mechanics* (North-Holland, Amsterdam, 1965), Pt. 1.

[27] N. W. Ashcroft, Phys. Rev. Lett. 21, 1748 (1968).

[28] See, for example, W. C. DeMarcus, Astron. J. 63, 2 (1958).

[29] M. Ross and A. K. McMahan, Phys. Rev. B 13, 5154 (1976).

[30] E. E. Salpeter, Phys. Rev. Lett. 28, 560 (1972).

[31] D. Straus and N. W. Ashcroft, Phys. Rev. Lett. 38, 415 (1977), see also D. Straus, Ph.D. thesis (Cornell Uni-

versity, Ithaca, N. Y , 1976) (unpublished)
N. Y , 1976) (unpublished).
[32]J Hubbard, Proc R Soc A 234, 336 (1975).
[33]D J W Geldart and S H. Vosko, Can J. Phys. 44, 2137 (1972)
[34]P Vashishta and K S Singwi, Phys Rev. B 6, 875 (1972).
[35]C A. Kukkonen, Ph D thesis (Cornell University, Ithaca, N.Y., 1975) (unpublished)
[36]D. J. Chadi and M L Cohen, Phys Rev B 12, 5747 (1973)
[37]M H. Cohen and F. Keffer, Phys Rev 99, 1128 (1955).
[38]F. A Lindemann, Z. Phys. 11, 609 (1910).
[39]G. Chester (private communication).
[40]R. Guyer (private communication).

PHYSICAL REVIEW B VOLUME 18, NUMBER 5 1 SEPTEMBER 1978

Analytical treatment of hypernetted-chain and Percus-Yevick equations for bosons*

Sudip Chakravarty and N. W Ashcroft
Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853
(Received 24 April 1978)

We show that for a careful choice of the Jastrow wave function the solution of hypernetted-chain and Percus-Yevick integral equations can be analytically reduced to the solution of a set of coupled algebraic equations These equations are then solved numerically and the ground-state energies of liquid ^{4}He and hard-sphere bosons are obtained.
PACS numbers 1977 67.40.-w 64.30 +t BR1059

I. INTRODUCTION

Various integral equation methods have been used[1,2] to study the ground-state properties of boson fluids. In these methods the analogy between the many-particle Jastrow wave function[1,2] and the Gibbs statistical-probability factor is exploited to carry over the whole machinery of classical theories to the quantum case. However, although there are two specific cases,[3] namely the hard spheres in Percus-Yevick (PY) approximation, and mean spherical model (MSM) for Yukawa closure, where these equations can be solved analytically in the classical statistical mechanics, no such analytical solution exists to date in the quantum case.

In this paper we shall show that, with a suitable choice of the form of the Jastrow wave function, considerable progress can be made toward obtaining analytical solutions of PY and hypernetted-chain (HNC) equations. The method will be applied to quantum hard spheres and ^{4}He interacting via the standard Lennard-Jones potential.

In Sec. II we outline the formulation of the variational problem. The choice of the wave function which allows us to use analytical methods is introduced in Sec. III and the solution of the integral equation is discussed in Sec. IV. Finally in Secs. V and VI we discuss its applications to the helium and hard-sphere problems, respectively.

II. VARIATIONAL PROCEDURE

The Hamiltonian for particles of mass m interacting with a pairwise potential $v(r_{ij})$ is taken to be

$$H \equiv -\frac{\hbar^2}{2m}\sum_{i=1}^{N}\nabla_i^2 + \sum_{i<j} v(r_{ij}) \qquad (2.1)$$

For bosons a variational many-particle Jastrow wave function[1,2]

$$\psi = \prod_{i<j} f(r_{ij}) \qquad (2.2)$$

will be used to obtain an upper bound of the ground state energy Given Eq. (2.2) the energy per particle can be written

$$\frac{E}{N} = \frac{\hbar^2\rho}{8m}\int d\vec{r}\,\vec{\nabla}g(r)\cdot\vec{\nabla}\ln f^2(r) + \frac{1}{2}\rho\int d\vec{r}\,v(r)g(r)\ , \qquad (2.3)$$

where ρ is the average number density and we have used the well-known Jackson-Feenberg identity[1,2] to write down the expectation value of the kinetic energy operator. If $f(r)$ is chosen so that it vanishes at a core distance, say a, but its derivative is discontinuous (the case of hard spheres) at a, the derivation of the Jackson-Feenberg identity requires a careful manipulation of the surface integral at the inner surface.[4] The final result is, however, the same. The pair correlation function $g(r)$ appearing in Eq. (2.3) is defined to be

$$g(r) = \frac{N(N-1)}{\rho^2}\frac{\int|\psi|^2 d\vec{r}_3\cdots d\vec{r}_N}{\int|\psi|^2 d\vec{r}_1\cdots d\vec{r}_N}\ . \qquad (2.4)$$

Note that there are alternative forms of the kinetic energy functional which introduce three particle correlation function. The case is often made[5] that the Jackson-Feenberg identity [Eq. (2.3)] is too sensitive to the short distance behavior of the pair correlation function and hence may be unsuitable for use in conjunction with the correlation functions obtained from approximate integral equations. These in turn are considered inaccurate at short distances and a practical consequence is that such errors are important in the balance between the kinetic and the potential energies. These objections apply to a certain extent in the present calculation. We shall see later on, for example (Sec. V) for the case of ^{4}He, that although the potential energy is calculated with an accuracy of $\sim 2\frac{1}{2}\%$ the corresponding error in kinetic energy is $\sim 7\%$ at the equilibrium ^{4}He density (the comparison is being made with respect to a standard Monte Carlo calculation) A comparison of the pair correlation functions

we calculate will reveal that the inaccuracy is limited in its entirety to short distances (<2.556 Å in the case of ^{4}He). However the important point we want to stress is that the different kinetic energy functionals are all obtained by different integration by parts and should in principle yield exactly the same answer in an *exact* theory. The fact that in practice they do not should be regarded as a shortcoming of the theory and the spread in the results can be considered to be a measure of the accuracy. The alternative forms of the kinetic energy requiring the knowledge of three-body correlation functions are invariably approximated by a Kirkwood superposition approximation.[1,2] It has been stated[5] that this approximation is exact within HNC and therefore a perfectly consistent one to use. There are indeed plausibility arguments in support of this statement but no proof of its validity. If it is not a rigorous result then use of these alternative forms incur additional approximations and are therefore less desirable. Further discussion of this problem is given by Zabolitsky.[6]

Returning to Eq. (2.3) we note that a given wave function (and the corresponding pair correlation function obtained through an integral equation) uniquely defines a variational problem. To proceed further we make use of the Ornstein-Zernike[7] equation which introduces a function $c(r)$ known as the direct correlation function,

$$g(r_{12})-1=c(r_{12})+\rho\int d\bar{r}_3 c(r_{23})[g(r_{13})-1] \ . \tag{2.5}$$

The equation can be regarded as an integral equation for $g(r)$ if a further relation between $c(r)$ and $g(r)$ is prescribed. The PY equation, for example, sets

$$e^{-\beta v(r)}c(r)=g(r)(e^{-\beta v(r)}-1)$$

for a classical fluid. The generalization to the quantum case[1] is given by

$$f^2(r)c(r)=g(r)[f^2(r)-1] \ . \tag{2.6}$$

Similarly the corresponding relation for the hypernetted-chain theory generalized to the quantum case, links $c(r)$ to $g(r)$ by[1]

$$c(r)=g(r)-1-\ln[g(r)/f^2(r)] \ . \tag{2.7}$$

These generalizations to the quantum case are made plausible by noting that $\ln f^2(r)$ plays the role of the classical factor $-\beta v(r)$. For a given $f^2(r)$ Eqs (2.5)–(2.7) can be solved to obtain the pair correlation function which is subsequently used in variational search for the ground state energy.

III. CHOICE OF THE WAVE FUNCTION

We shall make a judicious choice of the wave function to map the problem onto the classical MSM.[7] The wave function is defined in two regions; the physical significance of the separating boundary will become apparent later. For PY we choose

$$f^2(r)=0 \ , \quad r<a$$
$$=g(r)\left[g(r)-\sum_{i=1}^{n}\beta_i\frac{e^{-z_i(r-a)}}{r}\right] \ , \quad r\geq a \tag{3.1}$$

and

$$\lim_{r\to a^+} g(r)=A_m(r-a)^m \ , \quad m=0,1,2,\ldots \ , \tag{3.2}$$

while for HNC we take

$$f^2(r)=\begin{cases}0 \ , & r<a\\ g(r)\exp\left[1-g(r)+\sum_{i=1}^{n}\beta_i\frac{e^{-z_i(r-a)}}{r}\right] \ , & r\geq a\end{cases} \tag{3.3}$$

and

$$\lim_{r\to a^+} g(r)=A_m(r-a)^m \ , \quad m=0,1,2,\ldots \ . \tag{3.4}$$

In either case we can easily verify that

$$c(r)=\sum_{i=1}^{n}\frac{\beta_i}{r}e^{-z_i(r-a)} \ , \quad r>a \ . \tag{3.5}$$

Along with a, which stands for the distance at which we decide to set the wave function to zero, the set of $2n$ coefficients $\{\beta_i,z_i\}$ should be regarded as variational parameters. However, m of these parameters can be eliminated immediately as a consequence of the boundary conditions (3.2) or (3.4). Thus there are, in total, $2n+1-m$ independent variational parameters. The condition (3.2) or (3.4) brings out the quantum nature of the problem and comes from the requirements of continuity of the wave function. To see it more clearly we may characterize the wave function by the requirement that it satisfy

$$g(r)=\frac{dg(r)}{dr}=\cdots=\frac{d^{m-1}}{dr^{m-1}}g(r)=0 \text{ as } r\to a^+ \tag{3.6}$$

which implies that

$$f^2(r)=\frac{d}{dr}f^2(r)=\cdots=\frac{d^{m-1}}{dr^{m-1}}f^2(r)=0 \text{ as } r\to a^+ \ . \tag{3.7}$$

At this point one should note that except for the conditions (3.6) and (3.7), the problem is quite similar to classical mean spherical model for a sum of n Yukawa potentials, which is defined by imposing the conditions

(a) $g(r)=0\,,\quad r<a\,,$ (3.8)

(b) $k_BTc(r)=-\sum_{i=1}^{n}a_i\frac{e^{-b_ir}}{r}\,,\quad r>a\,.$ (3.9)

The similarity is, however, somewhat misleading. In the theory of classical liquids MSM is an approximation. Here on the other hand we are choosing a *wave function*, and this closure which leads to Eq. (3.5) can be regarded as a rigorous procedure in the sense that the wave function is always at our disposal for a *variational calculation*. Furthermore, the boundary conditions discussed above do not arise in the classical context.

IV. SOLUTION OF THE INTEGRAL EQUATION

We now turn to the Wiener-Hopf technique as used by Baxter[8] in solving the classical PY equation for hard spheres and later by Hoye and Blum[9] in the classical mean spherical approximation. We shall outline the main argument and give the modifications necessary for the present problem. The details are given in the Refs. 8 and 9. In Appendix A we shall give another derivation which is more transparent, although less useful from the computational point of view. The details of this method which was first used by Waisman[3] have not yet appeared in print.

The Fourier transform of Ornstein-Zernike equation, leads to

$$1+\rho h(k)=1/[1-\rho c(k)]=S(k) \tag{4.1}$$

where, $S(k)$ is the static structure factor[1] and

$$h(k)=4\pi\int_0^\infty dr\,\cos kr\int_r^\infty t[g(t)-1]\,dt\,. \tag{4.2}$$

Using (3.5) we find

$$1-\rho c(k)=1-4\pi\rho\int_0^a\cos kr\int_r^a dt\,tc(t)\,.$$

$$-\frac{4\pi\rho}{}\sum_{i=1}^{n}\frac{\beta_i}{k^2+z_i^2}\left[az_i\frac{\sin ak}{ak}+\cos ak\right]$$

$$-4\pi\rho\sum_{i=1}^{n}\frac{\beta_i'}{k^2+z_i^2}\left[az_i\frac{\sin ak}{ak}+\cos ak\right] \tag{4.3}$$

Since for the fluid phase $1-\rho c(k)$ does not have zeros on the real axis, is regular within a strip containing the real axis, and tends uniformly to unity at infinity, we follow Baxter[8] and factor $1-\rho c(k)$ as follows

$$1-\rho c(k)=Q(k)Q(-k)\,. \tag{4.4}$$

Note that the contribution $c(r)$ for $r>a$, has been to introduce $2n$ discrete poles on the imaginary axis at $\pm iz_j$. This factorization now leads to a set of two coupled integral equations for $Q(r)$, $c(r)$, and $g(r)$. At this point one might question what has been achieved by the replacement of the original Ornstein-Zernike equation by a set of two coupled equations involving yet another unknown function $Q(r)$. The point is, however that the form of $c(r)$ for $r>a$ as given by Eq. (3.5) fully determines the form of $Q(r)$ for $r>a$. This one can verify by taking a Fourier transform of $Q(r)$ and we find that

$$Q(r)=Q_0(r)+\sum_{i=1}^{n}d_ie^{-z_ir}\,,\quad r>0 \tag{4.5}$$

where d_i's are the contribution due to the residues at the poles $(-iz_j)$ on the imaginary axis, and also that

$$Q_0(r)=0\,,\quad r\geq a\,. \tag{4.6}$$

We shall now see that the choice of a wave function which vanishes for $r<a$ [and correspondingly a $g(r)$ which also vanishes for $r<a$] immediately determines $Q(r)$ for $r<a$ as well. To see how this comes about, we examine the Baxter equations[8] relating g, c, and Q:

$$rg(r)=r-Q'(r)-12\eta\int_0^R(r-t)Q(t)\,dt+12\eta\int_0^R(r-t)g(|r-t|)Q(t)\,dt \tag{4.7}$$

and

$$rc(r)=-Q'(r)+12\eta\int_r^R dt\,Q'(t)Q(t-r)\,, \tag{4.8}$$

where we have introduced

$$\eta=\tfrac{1}{6}\pi\rho a^3 \tag{4.9}$$

and have measured all distances in units of a. (From now on unless otherwise stated, all constants appearing in our calculation which have dimensions of length will be measured in units of a). In (4.7) and (4.8) the upper limit of integration $R=1$ for that part of $Q(r)$ which is $Q_0(r)$ and ∞ for the rest. The conditions $g(r)=0 (r<1)$ and $Q(r)=\sum_{i=1}^{n}d_ie^{-z_ir}(r>1)$ can now be trivially used in Eq. (4.7) to give $Q(r)$ for $r<1$. This completely determines $Q(r)$ everywhere and consequently $c(r)$ everywhere from Eq. (4.8) [We already know $c(r)$ for $r>1$ from the choice of the wave function.] One can easily verify that the form

$$Q_0(r)=\tfrac{1}{2}A(r^2-1)+B(r-1)+\sum_{i=1}^{n}c_i(e^{-z_ir}-e^{-z_i}) \tag{4.10}$$

along with the Eq (4.5), solves Eq (4.7) provided the constants are related by,

$$B(1+2\eta)+\tfrac{3}{2}\eta A=12\eta\sum_{i=1}^{n}\frac{c_i}{z_i^2}[1-(1+z_i+\tfrac{1}{2}z_i^2)e^{-z_i}]+\sum_{i=1}^{n}\frac{d_i}{z_i^2}\,. \tag{4.11}$$

Equation (4.3) should read

$$1-\rho c(k)=1-4\pi\rho\int_0^a\cos kr\int_r^a dt\,tc(t)$$

$$6\eta B-(1-4\eta)A=-1+12\eta\sum_{i=1}^{n}\frac{c_i}{z_i}[1-(1+z_i)e^{-z_i}]$$

$$+\sum_{i=1}^{n}\frac{d_i}{z_i}\ , \tag{4.12}$$

and

$$z_i(c_i+d_i)=12\eta d_i G(z_i)\ , \tag{4.13}$$

where

$$G(z_i)=\int_1^\infty r e^{-rz_i} g(r)\,dr\ . \tag{4.14}$$

The solution is therefore not complete until we obtain an independent equation for $G(z_i)$. This can be obtained by taking a Laplace transform of Eq. (4.7) (see Høye and Blum[9]). We then find

$$G(s)=\frac{s\tau(s)e^{-s}}{1-12\eta q(s)}\ ,$$

and

$$q(s)=\sigma(s)-\tau(s)e^{-s} \tag{4.15}$$

where,

$$\tau(s)=\frac{A}{s^3}(1+s)+\frac{B}{s^2}-\frac{1}{s}\sum_{i=1}^{n}\frac{z_i}{z_i+s}c_i e^{-z_i} \tag{4.16}$$

and

$$\sigma(s)=\frac{1}{s^3}(A+Bs)-\left[\frac{A}{2}+B+\sum_{i=1}^{n}c_i e^{-z_i}\right]\frac{1}{s}$$

$$+\sum_{i=1}^{n}\frac{c_i+d_i}{z_i+s}\ . \tag{4.17}$$

The difficult question now is whether it is possible to solve Eqs. (4.11)–(4.17) and hence obtain all the constants appearing in $Q(r)$. We shall see in Secs V and VI that although it is not possible to obtain an analytic solution in general, the equations can be simplified considerably. The problem can be reduced to a set of simple coupled algebraic equations which in turn can be solved quite simply on a computer. This simplification is due in its entirety to the quantum boundary conditions (3.6) and (3.7). The corresponding classical case is considerably more complex.

For the time being, let us assume that we have solved Eqs. (4.11)–(4.17) and have obtained $Q(r)$ completely. We can now use this $Q(r)$ in Eq. (4.8) and obtain $c(r)$. This calculation (see Høye and Blum[9]) is perfectly straightforward but is extremely tedious. We shall merely give the final results:

$$-rc(r)=A_0 r+B_0 r^2+\frac{1}{2}\eta A_0 r^4+\sum_{i=1}^{n}\frac{u_i}{z_i}(1-e^{-z_i r})$$

$$+\sum_{i=1}^{n}\frac{u_i^2}{2\beta_i e^{z_i}z_i^2}(\cosh z_i r-1)\ ,\quad r<1 \tag{4.18}$$

where the new constants appearing are related to the old constants by

$$A_0=A^2\ , \tag{4.19}$$

$$B_0=-12\eta\left[\frac{1}{2}(A+B)^2+A\sum_{i=1}^{n}c_i e^{-z_i}\right]\ , \tag{4.20}$$

and

$$u_i=24\eta\beta_i e^{z_i}G(z_i)\ . \tag{4.21}$$

We also note that the original parameters β_i which first appeared in the definition of the wave function [Eqs. (3.1) and (3.3)] are connected to these constants by

$$\beta_i e^{z_i}=z_i d_i[1-12\eta q(z_i)]\ . \tag{4.22}$$

where $q(z_i)$ is defined in Eqs. (4.15)–(4.17). Equation (4.18) now completely determines $S(q)$, the static structure factor, and (by Fourier transfromation) the pair correlation function $g(r)$.

We can now return to the variational problem and calculate the ground state energy. However, this is still not easy since the Fourier transformation from $S(q)$ to $g(r)$ cannot be done analytically and hence one cannot obtain the required derivative of $g(r)$ accurately. This is especially true for small r which encompasses the most important region for both kinetic and potential energies for short-range singular potentials of interest here. Thus we need a more accurate method to calculate $g(r)$ at short distances. This will be discussed in the following sections.

V. GROUND-STATE ENERGY OF ^{4}He

By way of application we turn to helium. For $v(r)$ we consider only the most extensively studied Lennard-Jones potential for ^{4}He, defined by

$$v(r)=4\epsilon[(\sigma/r)^{12}-(\sigma/r)^6]\ ,$$

$$\epsilon=10.22\ \text{K}\ , \tag{5.1}$$

$$\sigma=2.556\ \text{Å}\ .$$

We shall present the results for HNC only. For PY we have found as others have found[10] that the computed ground state energies fall *below* the experimental values and is therefore clearly unsuitable for variational purposes. We take $m=4$, i.e., we choose the boundary conditions such that

$$g(r)=\frac{dg(r)}{dr}=\cdots=\frac{d^4g(r)}{dr}=0\ \text{as}\ r\to 1^+\ . \tag{5.2}$$

These correspond to

$$f^2(r)\sim(r-1)^5\ \text{as}\ r\to 1^+\ . \tag{5.3}$$

Our rough estimate of the effect of setting an additional derivative to zero gives a lowering of energy of the order 1%

$$+\sum_{i=1}^{n}\frac{u_i^2}{2\beta e^{z_i}z_i^2}(\cdots)$$

and consequently all our calculations were done with first four derivatives set to zero. Before we carry out the solution of Eqs. (4.11–4.17), we define

$$\lambda_i^* = e^{z_i}[12\eta G(z_i)/z_i] \ , \tag{5.4}$$

$$C_i^* = c_i e^{-z_i} \ , \tag{5.5}$$

$$d_i^* = d_i e^{-z_i} \ . \tag{5.6}$$

Equation (4.7) can also be rewritten for $r > 1$ as

$$rg(r) = Ar + B - \sum_{i=1}^{n} z_i c_i e^{-z_i r} + 12\eta \int_0^r (r-t) g(|r-t|) Q(t)\, dt \ . \tag{5.7}$$

Then from Eq. (5.7) one can easily show that the conditions expressed in (5.2) lead to the following equations:

$$A + B = \sum_{i=1}^{5} z_i C_i^* \ , \tag{5.8}$$

$$A = -\sum_{i=1}^{5} z_i^2 C_i^* \ , \tag{5.9}$$

and

$$\sum_{i=1}^{5} z_i^3 C_i^* = \sum_{i=1}^{5} z_i^4 C_i^* = \sum_{i=1}^{5} z_i^5 C_i^* = 0 \ . \tag{5.10}$$

It is worth summarizing what has been done. Equations (5.8)–(5.10) imply relations between variational parameters (z_i's) and the constants (C_i^*'s). We would now keep z_i's as free variational parameters and choose C_i^*'s such that Eqs. (5.8) and (5.9) are satisfied; these in turn fix the variational parameters (β_i's) through Eq. (4.22). Thus Eqs. (5.8)–(5.10) can easily be solved to express C_i^*'s in terms of variational parameters z_i's and two unknown constants A and B. These can be substituted in Eqs. (4.11)–(4.13) to eliminate the constants A and B. We would then have C_i^*'s expressed in terms of only the constants λ_i^*'s and we may finally use Eqs. (4.15)–(4.17) to reduce the problem to coupled algebraic equations involving the λ_i^*'s. These can be solved iteratively. This reduction is carried out in detail in Appendix B. Thus given a set of variational parameters $\{z_1, z_2, z_3, z_4, z_5\}$ and η, we start with a set of guesses for $\{\lambda_1^*, \lambda_2^*, \lambda_3^*, \lambda_4^*, \lambda_5^*\}$ and analytically solve for $\{C_1^*, C_2^*, C_3^*, C_4^* C_5^*\}, A, B$ and hence $\{\beta_1^*, \beta_2^*, \beta_3^*, \beta_4^*, \beta_5^*\}$ such that the boundary conditions (5.2) are satisfied. This knowledge is then used to calculate a new set of $\{\lambda_1^*, \lambda_2^*, \lambda_3^*, \lambda_4^*, \lambda_5^*\}$. The procedure is iterated until the two successive sets of λ_i^*'s do not differ by more than 10^{-14}. For this purpose it is useful to rewrite Eqs. (4.15)–(4.17) in the form

$$\frac{\lambda^*(s)}{1-\lambda(s)} = 12\eta\left(Az_1z_2z_3z_4z_5 + S\sum_{i=1}^{5} C_i^* z_i^6\right)\Big/\prod_{i=1}^{5}(s+z_i)\left[s^3 - 12\eta\left((A + BS - \Gamma s^2) - s^3\sum_{i=1}^{5}\frac{C_i^*}{s+z_i}\frac{\lambda^*(z_i)}{1-\lambda(z_i)}\right)\right] \ , \tag{5.11}$$

where

$$\lambda(s) = e^{-z_i}\lambda_i^*(s) \tag{5.12}$$

and

$$\Gamma = \frac{A}{2} + B + \sum_{i=1}^{5} C_i^* \tag{5.13}$$

The derivation of Eq. (5.10) is lengthy and requires extensive use of Eqs. (5.8)–(5.10).

A. Pair correlation function

We remarked in Sec. IV that a more reliable method is required to calculate the pair correlation function, especially at short distances. This will be discussed here. We have seen that all constants appearing in Eq. (4.15) can be determined explicitly, and this enables us to invert the Laplace transform of $rg(r)$, which is expressed by Eqs. (4.15)–(4.17). This can be done in the manner indicated by Wertheim[2] and is based on strip-wise Laplace inversion. It can be shown that[9]

$$rg(r) = \sum_{m=1}^{\infty} (12\eta)^{m-1}(-1)^{m-1} \times \frac{1}{2\pi i}\int_{\delta-i\infty}^{\delta+i\infty}\left[\frac{N(s)}{P(s)}\right]^m e^{s(r-m)} s\, ds \ , \tag{5.14}$$

where δ is to be chosen such that the contour lies to the right of the real zero of $P(s)$ and

$$N(s) = s^3\prod_{i=1}^{5}(s+z_i)\tau(s) \ , \tag{5.15}$$

and

$$P(s) = s^3\prod_{i=1}^{5}(s+z_i)[1 - 12\eta\sigma(s)] \tag{5.16}$$

Furthermore it can be verified that $N(s)$ is only a first order polynomial; this simplification comes about, once again, as a consequence of the conditions (5.2). On the other hand $P(s)$ is an eighth order polynomial. We shall see below that $g(r)$ can be obtained once we know the roots of the polynomial $P(s)$. (All details are to be found in this Appendix C.) We have,

$$\omega_1 = \sum_{i=1}^{8} c_0(i) e^{\mu_i (r-1)} \quad , \tag{5.17}$$

$$\omega_2 = \sum_{i=1}^{8} d_0(i) e^{\mu_i (r-2)} [d_1(i) + (r-2) d_2(i)] \quad , \tag{5.18}$$

$$\omega_3 = \sum_{i=1}^{8} e_0(i) e^{\mu_i (r-4)} [e_1(i) + (r-3) e_2(i) + (r-3)^3 e_3(i)] \quad , \tag{5.19}$$

and

$$\omega_4 = \sum_{i=1}^{8} f_0(i) e^{\mu_i (r-4)} [f_1(i) + (r-4) f_2(i) + (r-4)^2 f_3(i) + (r-4)^3 f_4(i)] \quad , \tag{5.20}$$

The coefficients are to be found in Appendix C. The μ_i's are the eight real or complex roots of the polynomial $P(s)$. It is also easy to prove that ω_i's are real, as they should be.

B. Results

Given the expression for the pair correlation function we calculate the ground state energy from Eq. (2.3): in kelvin per atom

$$\frac{E}{N} = 1.8550 \delta^2 T_0 + 490.56 \eta \int_1^\infty x^2 g(x) \left[\left(\frac{\delta}{x} \right)^{12} - \left(\frac{\delta}{x} \right)^6 \right] dx \quad , \tag{5.21}$$

where,

$$T_0 = -\frac{1}{4} \sum_{i=1}^{5} \lambda_i^* \beta_i z_i^3 + 3\eta \int_1^\infty dx\, x^2 \left[\frac{1 - g(x)}{g(x)} \right] \left[\frac{dg(x)}{dx} \right]^2 \quad , \tag{5.22}$$

and

$$\delta = (\sigma/a) = (\pi \rho \sigma^3 / 6\eta)^{1/3} \quad . \tag{5.23}$$

This expression is minimized for a given density (or a given $\rho\sigma^3$) as function of $\{z_1, z_2, z_3, z_4, z_5\}$ and a (or η). The results are tabulated as a function of $\rho\sigma^3$ in Table I, and compared in Fig. 1, both with the molecular dynamics calculation of Schiff and Verlet[11] and with a very recent conventional numerical solution of HNC by Miller.[12] The results are virtually identical (−0.5%) molecular dynamics results. The pair correlation function is in very good agreement for distances greater than the Lennard-Jones diameter $\sigma = 2.556$ Å. For $r < \sigma$ there is apparently only a slight disagreement but it is enough to cause a discrepancy of the order of −1 K in comparisons of the total energy. It is interesting to look at the separation of the energy at the equilibrium density of ^{4}He. We obtain a kinetic energy of 14.75 K as opposed to 13.73 obtained in the molecular dynamics calculation and −19.11 K for the potential energy instead of −19.46 K. This validates our earlier statement (Sec. II) that most of the error lies in the kinetic energy and the substantial cancellation of the two energies lead finally to a large discrepancy in their sum.

Finally we would like to comment on the numerical aspect of the problem. The iterative solution of the constants is rather trivial and very fast. Potential problems arise only in computing the roots of the polynomial $P(s)$, and hence a very high accuracy should be maintained (we use the Laguerre iteration technique[13]). The reason is that from time to time for some choice of the parameters the polynomial can become ill conditioned[13] and the error may propagate rapidly to the determination of coefficients appearing in the expression for the pair correlation function [Eq. (5.16)] This happens especially for large distances and at low densities and is due to the cancellation of a large number of unwieldy terms as should be clear

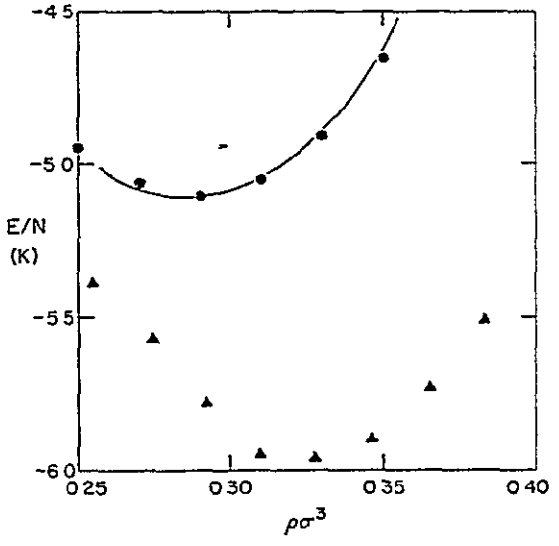


FIG. 1 Calculated ground-state energy of liquid ^{4}He. Solid curve, the present calculation; Δ, molecular-dynamics results obtained by Schiff and Verlet (Ref. 11), •, HNC calculation of Miller (Ref. 12).

[NOTE ADDED IN PROOF: Though the HNC procedure leads to a g(s) in good overall agreement with the simulation results, the thermodynamic functions for the Lennard-Jones potential involve very high moments of g and therefore impose stringent requirements on accuracy, particularly in the neighborhood of the first peak. In this respect a significant improvement has been obtained by solving a modified HNC equation which retains the bridge function (normally set to zero in standard HNC) but in a form appropriate to hard spheres: the hard sphere diameter required in the function is determined by the demands of thermodynamic consistency. When the resulting classical g(r) is used in the calculation

TABLE I Ground-state energy of liquid ^{4}He.

$\rho\sigma^3$	η	Z_1	Z_2	Z_3	Z_4	Z_5	KE/N (K)	PE/N (K)	E/N (K)
0.26	0.05	10 5	8 1	6 7	7.3	8 0	8.424	−13.458	−5 034
0 28	0 053	11 5	7 8	6 4	7 3	8 0	9 358	−14.459	−5 101
0 30	0 057	12.0	7.5	6 2	7 3	8 0	10 587	−15.670	−5 083
0 32	0 0605	13.0	8 1	6.5	7 0	8 0	11 672	−16 652	−4 980
0.34	0 064	12.5	8 4	6 8	7.0	8.0	12.926	−17.700	−4.774
0.3648	0 069	13 5	8 7	6 8	7.0	7 7	14 752	−19 107	−4 344

−4.355 not −4.344

from Appendix C. However, for $1 < r < 2$, the region which contributes most to the energy it is very reliable. For larger distances ($r \sim 4$) it is more reliable to obtain the pair correlation function from the Fourier transform of the structure factor. In any case with care it is not difficult to keep the total numerical error in the energy less than $\frac{1}{2}\%$. As is always the case,[13] for a variational problem with a large number of parameters, one cannot guarantee anything more than a local minimum and we do not claim to have obtained global minima although we have determined that at the quoted minima all partial derivatives are zero.

VI. HARD SPHERES

The calculation proceeds in exactly the manner described in the earlier section but is considerably simpler. We shall therefore omit the details and content ourselves with a simple two-parameter variational search instead of a six-parameter search as described above. The results could of course be improved by the introduction of more parameters but this is not necessary in displaying the method In this case we have to keep in mind that $f(r)$ must be continuous at the core boundary but one must allow for the discontinuity of the derivative $f'(r)$ at this point. This requires

$$g(r) = 0\ , \quad r \to 1^+ \tag{6.1}$$

and

$$\frac{dg(r)}{dr} = 0\ , \quad r \to 1^+ \tag{6.2}$$

Insert period

The kinetic energy, which is also the total energy, can be written

$$\frac{E}{N(\hbar^2/ma^2)} = 3\eta \int_1^\infty dx\, x^2 \left[\frac{dg(x)}{dx}\right]^2 \frac{1-g(x)}{g(x)} - \frac{1}{4}\sum_{i=1}^{2} \lambda_i^* \beta_i z_i^3\ , \tag{6.3}$$

where the notations of the previous sections have been used. The results are tabulated in Table II, and compared with the variational Monte Carlo results of Hansen, Levesque and Schiff.[14] The results are in reasonable agreement throughout and especially in the fluid phase which they found to exist for values $\rho a^3 < 0.244$ (see Fig. 4).

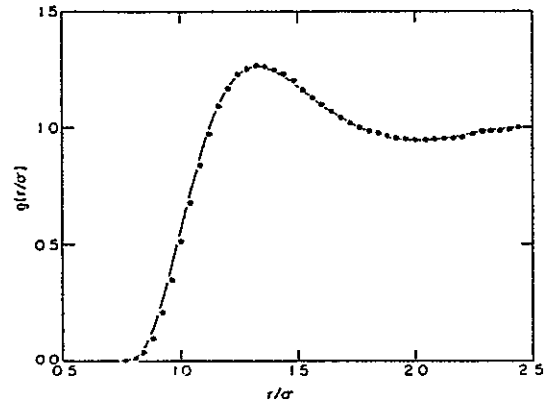


FIG 2. Comparison of the pair correlation function $g(r)$ at the equilibrium density of liquid ^{4}He Solid curve, the present calculation, •, molecular-dynamics results of Schiff and Verlet (Ref 11)

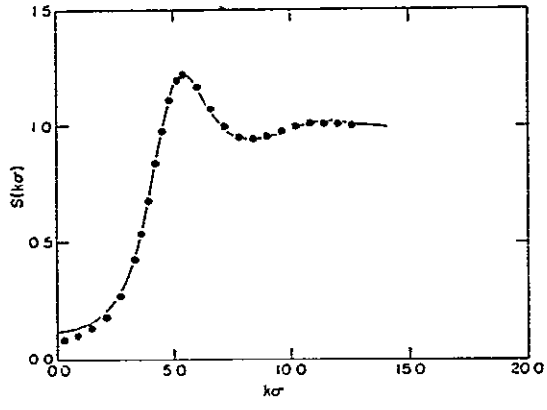


FIG. 3. Comparison of the static structure factor $S(k)$ at the equilibrium density of liquid ^{4}He. Solid curve, the present calculation. •, molecular-dynamics results of Schiff and Verlet (Ref 11).

Continued from previous page

NOTE ADDED IN PROOF: (cont)

→ accounted for (Y. Rosenfeld and N.W Ashcroft, unpublished).

0.0529
0.0659
0.1047
0.1277
0.1413
0.1571

Please truncate to 4 figures, as shown

TABLE II. Ground-state energy of hard-sphere bosons

ρa^3	η	Z_1	Z_2	$(E_{HS}/N)(\hbar^2/ma^2)$
0.1	0.05236	2.4	2.55	1.963
0.166	0.08691	3.4	3.7	4.663
0.2	0.10472	4.4	3.7	6.625
0.244	0.12775	5.4	4.3	9.886
0.27	0.14137	5.4	5.0	12.263
0.3	0.15708	5.4	6.0	15.501

VII. CONCLUSION

We have been able to show that for a general class of wave functions both the PY and HNC integral equations used widely to calculate variational upper bound to the ground state energies for a variety of Bose liquids can be reduced to a set of coupled *algebraic* equations. Although simple in nature, these equations were solved numerically and the results were applied to the cases of liquid ^{4}He and quantum hard spheres. The results for liquid ^{4}He turned out to be almost identical to the more conventional numerical solution. For hard spheres the results are in reasonable agreement with the previous variational Monte Carlo calculations. One of the advantages of this method is the extreme accuracy with which the pair correlation function can be calculated. This is because the pair correlation function can be obtained analytically from a Laplace inversion technique once the solutions of the algebraic equations mentioned above are known.

We consider this present calculation to be only a first step toward obtaining a complete analytical solution and a considerable amount of mathematical simplification is still required to make this technique more efficient than a more conventional numerical solution of the integral equations. We hope to look into it in the near future.

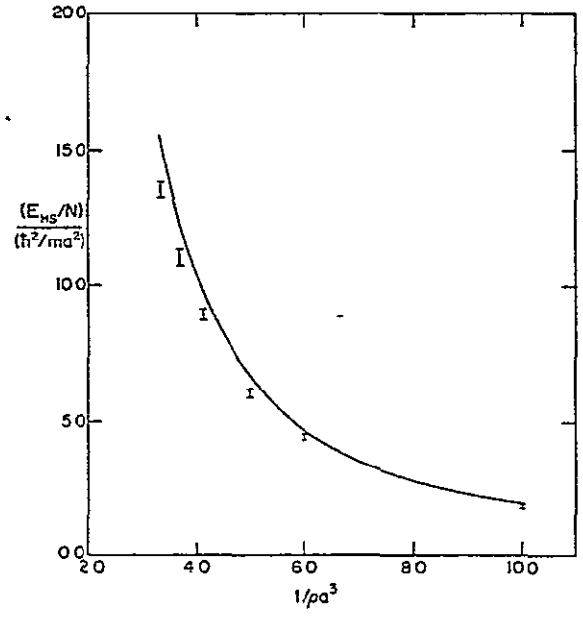


FIG 4 Calculated ground-state energies of quantum hard spheres (bosons) Solid curve, the present calculation, error bars, Monte Carlo calculation of Hansen, Levesque, and Schiff (Ref. 14)

ACKNOWLEDGMENTS

Useful conversations with P Bhattacharya and G. V. Chester are gratefully acknowledged. This work was supported by NASA under Grant No. NGR 33-010-188.

APPENDIX A

We present an alternative solution of the HNC integral equation, analogous to the method indicated by Waisman[13] For the sake of demonstration, consider only a one-Yukawa correction to $c(r)$, i e ,

$$c(r) = \beta e^{-\alpha r}/r \text{ for } r > 1 \ .$$

Following Wertheim[3] one can write down the Laplace transform of the Ornstein-Zernike equation,

$$G(t) = \left\{ \frac{K}{t^2} - F(t) - \frac{\beta e^{-t}}{z+t} + \frac{24\eta\beta\lambda}{z^2 - t^2} \right\} \times \left[1 + \frac{12\eta}{t} \left[F(-t) + \frac{\beta e^{t}}{z-t} - F(t) - \frac{\beta e^{-t}}{z+t} \right] \right]^{-1} ,$$

where

$$K = 1 - 24\eta \int_0^\infty x^2 c(x)\, dx \ ,$$

$$\lambda = e^z \int_1^\infty e^{-\alpha x} x g(x)\, dx \ ,$$

$$G(t) = \int_0^\infty e^{-\alpha x} x g(x)\, dx \ ,$$

and

$$F(t) = -\int_0^1 x c(x) e^{-\alpha x} dx \ .$$

We observe: (i) Both $F(t)$ and $F(-t)$ (being Laplace transforms over finite range) are analytic in the entire complex plane (ii) $G(t)$ being a Laplace transform over an infinite range can have singularities in the left half-plane but is analytic in the right half-plane. (iii) The only singularities of $G(t)$ are therefore the double pole at $t=0$. Note that $t = \pm z$ are not singular points.

One can also show that the function $H(t)$ defined by

$$H(t) = t^4(z^2 - t^2) G(t) N(-t) = t^4(z^2 - t^2) \frac{N(t) N(-t)}{D(t)}$$

$$N(t)=\left[\frac{\eta}{t^2}-F(t)-\frac{\beta e^{t}}{z+t}-\frac{\beta e^{-t}}{z^2-t^2}\right]$$

and

$$D(t)=1+\frac{12\eta}{t}\left[F(t)+\frac{\beta e^{t}}{z-t}-F(t)-\frac{\beta e^{-t}}{z+t}\right]$$

goes as a polynomial, specifically $H(t)\sim O(t^4)$ as $t\to\pm\infty$. Therefore by Liouville's theorem

$$H(t)=\mu+\beta t^2+\gamma t^4 ,$$

where α, β, γ are constants to be determined. Following Wertheim this leads to a form for $c(r)$ for $r<1$ and gives the result obtained by Waisman. The generalization to the case where $c(r)$ is a linear combination of n Yukawas for $r>1$ is trivial.

APPENDIX B

Equations (5.8) and (5.9) can be solved to obtain C_3^*, C_4^*, and C_5^* in terms only of z_i's and C_1^*, C_2^*, i.e.,

$$C_3^*=\alpha_1 C_1^*+\alpha_2 C_2^* ,$$

$$C_4^*=\delta_1 C_1^*+\delta_2 C_2^* ,$$

$$C_5^*=\gamma_1 C_1^*+\gamma_2 C_2^* ,$$

where

$$\alpha_1=\left[\frac{z_1}{z_3}\right]^3\frac{(z_5-z_1)(z_4-z_1)}{(z_3-z_5)(z_4-z_3)} ,$$

$$\alpha_2=\left[\frac{z_2}{z_3}\right]^3\frac{(z_5-z_2)(z_4-z_2)}{(z_3-z_5)(z_4-z_3)} ,$$

$$\delta_1=\left[\frac{z_1}{z_4}\right]^3\frac{(z_5-z_1)(z_3-z_1)}{(z_4-z_5)(z_3-z_4)} ,$$

$$\delta_2=\left[\frac{z_2}{z_4}\right]^3\frac{(z_5-z_2)(z_3-z_2)}{(z_4-z_5)(z_3-z_4)} ,$$

$$\gamma_1=-\left[\left[\frac{z_1}{z_5}\right]^3+\alpha_1\left[\frac{z_3}{z_5}\right]^3+\delta_1\left[\frac{z_4}{z_5}\right]^3\right] ,$$

and

$$\gamma_2=-\left[\left[\frac{z_2}{z_5}\right]^3+\alpha_2\left[\frac{z_3}{z_5}\right]^3+\delta_2\left[\frac{z_4}{z_5}\right]^3\right] .$$

These can be used to obtain C_1^* and C_2^* in terms of z_i's and A, B. The results can be substituted in Eqs (4.11)–(4.12) to obtain

$$C_1^*=[a_{22}(a_0+B_0)-a_{12}a_0]/\Delta$$

$$C_2^*=\qquad\qquad/\Delta$$

where

$$\Delta=a_{11}a_{22}-a_{21}a_{12}$$

and

$$a_{11}=(z_1+\alpha_1 z_3+\delta_1 z_4+\gamma_1 z_5)$$
$$-a_0(\psi_1-\chi_1)+(a_0+b_0)(\phi_1-\epsilon_1) ,$$

$$a_{12}=(z_2+\alpha_2 z_3+\delta_2 z_4+\gamma_2 z_5)$$
$$-a_0(\psi_2-\chi_2)+(a_0+b_0)(\phi_2-\epsilon_2) ,$$

$$a_{21}=4b_0(\psi_1-\chi_1)+a_0(\phi_1-\epsilon_1)$$
$$-(z_1^2+\alpha_1 z_3^2+\delta_1 z_4^2+\gamma_1 z_5^2) ,$$

$$a_{22}=4b_0(\psi_2-\chi_2)+a_0(\phi_2-\epsilon_2)$$
$$-(z_2^2+\alpha_2 z_3^2+\delta_2 z_4^2+\gamma_2 z_5^2) ,$$

$$a_0=\frac{1+2\eta}{(1-\eta)^2} , \quad a_0+b_0=\frac{1+\frac{1}{2}\eta}{(1-\eta)^2} .$$

We also have

$$\psi_i=12\eta\left[\frac{\lambda_i^*}{z_i^2(\lambda_i-1)}+\frac{\alpha_i\lambda_3^*}{z_3^2(\lambda_3-1)}+\frac{\delta_i\lambda_4^*}{z_4^2(\lambda_4-1)}+\frac{\gamma_i\lambda_5^*}{z_5^2(\lambda_5-1)}\right] ,$$

$$\phi_i=12\eta\left[\frac{\lambda_i^*}{z_i(\lambda_i-1)}+\frac{\alpha_i\lambda_3^*}{z_3(\lambda_3-1)}+\frac{\delta_i\lambda_4^*}{z_4(\lambda_4-1)}+\frac{\gamma_i\lambda_5}{z_5(\lambda_5-1)}\right] ,$$

$$\chi_i=12\eta\left[\frac{H(z_i)}{z_i^2}+\alpha_i\frac{H(z_3)}{z_3^2}+\delta_i\frac{H(z_4)}{z_4^2}+\gamma_i\frac{H(z_5)}{z_5^2}\right] ,$$

$$\epsilon_i=12\eta\left[\frac{E(z_i)}{z_i}+\delta_i\frac{E(z_3)}{z_3}+\delta_i\frac{E(z_4)}{z_4}+\gamma_i\frac{E(z_5)}{z_5}\right] ,$$

$$H(z_i)=1+z_i+\tfrac{1}{2}z_i^2 ,$$

and

$$E(z_i)=1+z_i .$$

BR1059/9

Should read —

Equations (5.8) and (5.9) can be solved

This completely determines C_1^*, C_2^* (and hence C_3^*, C_4^*, C_5^* which are linear combination of C_1^* and C_2^*) in terms of z_i's and λ_i^*'s. Also note that we determine A and B which are also linear combination of C_i^*'s. Thus the right-hand side of Eq. (5.10) can be expressed in terms of λ_i^*'s and the variational parameters, i.e., z_i's and a (or $\eta = \frac{1}{6}\pi p a^3$).

APPENDIX C

In this Appendix, we give the formulas to determine $g(r)$. $N(s)$ and $P(s)$ referred to in Eqs. (5.14) and (5.15) are given by

$$N(s) = Az_1z_2z_3z_4z_5 + s\sum_{i=1}^{5} C_i^* z_i^6 \ ,$$

is

$$\begin{aligned}
P(s) = {} & s^8 + s^7\left[X_1 + 12\eta\Gamma - 12\eta\sum_{i=1}^{5}\lambda_i^* d_i^*\right] + s^6\left[X_2 - 12\eta B + 12\eta\Gamma X_1 - 12\eta\sum_{i=1}^{5}\lambda_i^* d_i^*(X_1 - z_i)\right] \\
& + s^5\left[X_3 - 12\eta A - 12\eta B X_1 + 12\eta\Gamma X_2 - 12\eta\sum_{i=1}^{5}\lambda_i^* d_i^*(X_2 - X_1 z_i + z_i^2)\right] \\
& + s^4\left[X_4 - 12\eta A X_1 - 12\eta B x_2 + 12\eta\Gamma X_3 - 12\eta\sum_{i=1}^{5}\lambda_i^* d_i^*(X_3 - X_2 z_i + X_1 z_i^2 - z_i^3)\right] \\
& + s^3\left[X_5 - 12\eta A X_2 - 12\eta B X_3 + 12\eta\Gamma X_4 - 12\eta\sum_{i=1}^{5}\frac{\lambda_i^* d_i^* X_5}{z_i}\right] \\
& + s^2\left[-12\eta A X_3 - 12\eta B X_4 + 12\eta\Gamma X_5\right] + s\left[-12\eta A X_4 - 12\eta B X_5\right] - 12\eta A X_5
\end{aligned}$$

where

$$X_1 = z_1 + z_2 + z_3 + z_4 + z_5 \ ,$$

$$X_2 = z_1z_2 + z_1z_3 + z_1z_4 + z_1z_5 + z_2z_3 + z_2z_4 + z_2z_5 + z_3z_4 + z_3z_5 + z_4z_5 \ ,$$

$$X_3 = z_1z_2z_3 + z_1z_2z_4 + z_1z_2z_5 + z_1z_3z_4 + z_1z_3z_5 + z_1z_4z_5 + z_2z_3z_4 + z_2z_3z_5 + z_2z_4z_5 + z_3z_4z_5 \ ,$$

$$X_4 = z_1z_2z_3z_4 + z_1z_2z_3z_5 + z_1z_2z_4z_5 + z_1z_3z_4z_5 + z_2z_3z_4z_5 \ ,$$

$$X_5 = z_1z_2z_3z_4z_5 \ ,$$

and

$$\Gamma = \frac{A}{2} + B + \sum_{i=1}^{5} C_i^* \ .$$

To express the coefficients $c_0, d_0 \cdots$, etc. [Eqs. (5.17)–(5.20), let us define the roots of the polynomial $P(s) = 0$ by μ_i and use the following abbreviations

$$N \equiv N(\mu_i) \ , \quad P \equiv P(\mu_i) \ , \quad N' \equiv \frac{d}{ds}N(s)\Big|_{s=\mu_i} \ , \quad P' \equiv \frac{d}{ds}P(s)\Big|_{s=\mu_i}$$

and similarly for higher derivatives. Then the contour integration of Eq. (5.13) yields:

$$c_0(\mu_i) = \mu_i N/P' \ ,$$

$$d_0(\mu_i) = \frac{N}{P'^2} \ , \quad d_1(\mu_i) = \left[1 - \frac{\mu_i P''}{P'}\right]N + 2\mu_i N' \ , \quad d_2(\mu_i) = \mu_i N$$

$$e_0(\mu_i) = N/P'^3 \ ,$$

$$e_1(\mu_i) = \left[\frac{N}{P'}\right]^2 \mu_i(3P''^2 - P'P''') + 6N'(N + \mu_i N') - 3N\frac{P''}{P'}(N + 3\mu_i N')$$

$$e_2(\mu_i) = N\left[6N'\mu_i + N\left[2N - 3\mu_i\frac{P''}{P'}\right]\right] \ , \quad e_3(\mu_i) = \mu_i N^2 \ ,$$

$$f_0(\mu_i) = N/P'^4 \ ,$$

$$f_1(\mu_i) = 5\mu_i \left[\frac{N}{p'}\right]^3 P''(2P'P''' - 3P''^2) + T_1T_2T_3 - 24N'T_4T_5$$

$$+ 12N'^2(3N + 2\mu_i N') - \mu_i N^3 \frac{P''''}{p'} \ ,$$

$$f_2(\mu_i) = [\mu_i T_1 T_2 - 12(T_3T_4 - N'T_5)]N \ , \quad f_3(\mu_i) = (3T_3 - 6\mu_i T_4)N^2$$

and

$$f_4(\mu_i) = \mu_i N^3 \ .$$

Finally

$$T_1 = (N/p')^2 \ , \quad T_2 = 15P''^2 - 4P'P''' \ , \quad T_3 = N + 4\mu_i N' \ , \quad T_4 = NP''/P' \ ,$$

and

$$T_5 = 2N + 3\mu_i N' \ .$$

---

[1]E. Feenberg, *Theory of Quantum Fluids* (Academic, New York, 1969).

[2]C. -W. Woo in *Physics of Liquid and Solid Helium*, edited by K H. Benneman and J B Ketterson (Wiley Interscience, New York, 1976)

[3]M. S. Wertheim, Phys Rev. Lett 10, 321 (1963); E Waisman, Mol Phys 25, 45 (1973)

[4]J B. Aviles, Ann Phys (Leipz) 5, 251 (1958).

[5]V R Pandharipande and H. A. Bethe, Phys Rev C 7, 1312 (1973).

[6]G. Zabolitsky, Phys Rev A 16, 1258 (1977)

[7]J. P. Hansen and I R. McDonald, *Theory of Simple Liquids* (Academic, New York, 1976).

[8]R. J. Baxter, Aust J Phys 21, 563 (1968)

[9]J S Høye and L Blum, J. Stat. Phys. 16, 399 (1977)

[10]R. D Murphy and R. O. Watts, J. Low Temp Phys. 2, 507 (1970).

[11]D. Schiff and L Verlet, Phys. Rev 160, 208 (1967).

[12]M D Miller, Phys Rev B 14, 3937 (1976)

[13]G. Dahlquist and A. Sjorck, *Numerical Methods* (Prentice Hall, Englewood Cliffs, N. J., 1974)

[14]J. P. Hansen, D Levesque and D. Schiff, Phys Rev. A 3, 770 (1971).

A. Björck

G Dahlquist, A. Björck and N. Anderson, ---

ON THE GROUND STATE OF METALLIC HYDROGEN

Sudip Chakravarty and N.W. Ashcroft

Laboratory of Atomic and Solid State Physics

Cornell University, Ithaca, N.Y. 14853

ABSTRACT

A proposed liquid ground state of metallic hydrogen at zero temperature is explored and a variational upper bound to the ground state energy is calculated. It is shown that the possibility that the metallic hydrogen is a liquid around the metastable point ($r_s = 1.64$) cannot be ruled out. This conclusion crucially hinges on the contribution to the energy arising from the third order in the electron-proton interaction which is shown here to be more significant in the liquid phase than in crystals.

1. INTRODUCTION

An interesting possibility of a zero temperature liquid ground state of metallic hydrogen has been recently explored in a calculation[1] which makes use of a Jastrow-Slater many particle variational wavefunction[2,3] to calculate the ground state energies of both solid and liquid phases. The symmetric part of the wavefunction is treated by the Monte-Carlo technique; exchange is neglected in the solid and approximated in the liquid by the Wu-Feenberg expansion[2,3]. It is found that the differences in the energies of the liquid and the solid phases varies from 0.1% at $r_s = 1.6$ to about 3% at $r_s = 0.8$, (here $4\pi/3(r_s a_o)^3 = 1/n$ and n is proton or electron density). The solid phase seems to be energetically more favorable throughout the entire range of densities considered. However, the calculation is based on a model of pair-interactions between protons and therefore contains only terms generated to second order in the electron-proton interaction. The contribution coming from the third order in the electron-proton interaction is known to be significant in the calculation of the band-structure energy[4,5] in the solid. In view of the small energy difference between the solid and the liquid phases it is therefore necessary to estimate the third order term for the liquid as well. Furthermore, since in the liquid certain configurations will permit three protons to come closer together than they would in a solid, we might also expect that the contribution from the term third order in the electron - proton interaction may be relatively more important in the liquid phase.

In this paper we shall first show that a simple one-parameter variational wavefunction when combined with the Hypernetted Chain (HNC) integral equation[2] can reproduce the energies calculated in Ref. 1 with a 6-parameter variational wavefunction and the Monte-Carlo technique to within 0.025 - 4.2% and therefore provides a very reasonable upperbound. However, precise agreement is not necessary in order to provide variational answers to the following questions

(a) How much does the third order term contribute to the ground state energy of the liquid? (b) What are the corrections in the liquid state attributable to long wavelength phonons? (c) Is it possible to lower the energy of the liquid by permitting partial alignment of the spins of the protons?

The calculation described below is a judicious combination of variational and perturbative methods and is intended to suggest that for certain densities the possibility of a liquid metallic phase of hydrogen at zero temperature cannot be ruled out. The conclusion hinges on the fact that the third order term is significant and is perhaps more so in the liquid.

2. FORMULATION

In a sense hydrogen is the simplest metal; its Hamiltonian is known exactly: For N protons, N electrons and volume Ω we write

$$H = H_e + H_p + H_{ep}$$

$$= \left(-\frac{\hbar^2}{2m_e}\sum_{i=1}^{N}\nabla^2_{\vec{r}_i} + \sum_{i<j}\frac{e^2}{|\vec{r}_i-\vec{r}_j|}\right) + \left(-\frac{\hbar^2}{2m_p}\sum_{i=1}^{N}\nabla^2_{\vec{R}_i} + \sum_{i<j}\frac{e^2}{|\vec{R}_i-\vec{R}_j|}\right)$$

$$- \sum_{i,j}\frac{e^2}{|\vec{r}_i-\vec{R}_j|} \; . \qquad (2.1)$$

Here we have denoted the proton coordinates by $\{\vec{R}_i\}$ and the electron coordinates by $\{\vec{r}_i\}$. A major simplification takes place[6] when we realize that there are two widely different time scales involved in the problem, allowing us to remove electronic degrees of freedom by assuming that at any instant we can consider the electrons to be in the ground state corresponding to the instantaneous proton configuration. This Born-Oppenheimer adiabatic approximation reformulates the problem in terms of an effective Hamiltonian of protons. The price we pay is that the indirect interaction between the protons, now mediated by the electrons, is no longer a simple Coulombic pair interaction but contains

many body forces[7]. With electron coordinates now integrated out the total Hamiltonian for the protons becomes[8]

$$H_p = E_{eg} + T_p + V_{pp} + E_b^{(2)}(\{\vec{R}_\ell\}) + E_b^{(3)}(\{\vec{R}_\ell\}) + \ldots \tag{2.2}$$

where E_{eg}, which is the exact ground state energy of the interacting electrons in a uniform positive background appears as a constant energy, and simply drops out of the calculation. In Eq. (2.2) T_p and V_{pp} are the parts of the original Hamiltonian of the protons and $E_b^{(n)}(\{\vec{R}_\ell\})$ which are functions of the proton coordinates are the electron mediated interactions between protons which are generated by adiabatic perturbation theory. Provided Eq. (2.2) converges, the procedure is exact within the adiabatic approximation. Most importantly, note that to this point we have not made any assumptions regarding the positions of the ions; the discussion holds for liquids and crystals whether static or dynamic. The precise form of $E_b^{(n)}(\{R_\ell\})$ can easily be written down[8]

$$E_b^{(2)}(\{R_\ell\}) = \tfrac{1}{2}\Omega \sum_{\vec{k}_1}{}' V(\vec{k}_1) V(-\vec{k}_1)\ \chi^{(1)}(\vec{k}_1)\ , \tag{2.3}$$

$$E_b^{(3)}(\{R_\ell\}) = \Omega \sum_{\vec{k}_1,\vec{k}_2,\vec{k}_3}{}' V(\vec{k}_1)V(\vec{k}_2)V(\vec{k}_3)\ \chi^{(2)}(\vec{k}_1,\vec{k}_2,\vec{k}_3)\ \delta_{\vec{k}_1+\vec{k}_2+\vec{k}_3,0}, \tag{2.4}$$

and similarly for the nth order term. Here,

$$V(\vec{k}) = -\sum_{\ell} e^{i\vec{k}\cdot\vec{R}_\ell}\ \frac{4\pi e^2}{k^2}\ , \tag{2.4}$$

and

$$\chi^{(1)}(\vec{k}) = \left(\frac{k^2}{4\pi e^2}\right)\left[\frac{1}{\varepsilon(k)} - 1\right] \tag{2.5}$$

is the exact first order static response of the interacting electron gas to an external potential. Similarly $\chi^{(n)}(\vec{k}_1,\vec{k}_2,\ldots\vec{k}_{n+1})$ is the exact nth order response. In otherwords if we know the nth order response function of the interacting electron gas exactly, we would also know exactly these extra many body interactions between protons, and we can proceed to diagonalize the proton Hamiltonian.

The interesting point to note is that the rewriting of the original Hamiltonian in the form given in Eq. (2.2) splits off a large volume dependent term (order 1 Ry) which does not depend on whether the protons form a liquid or a solid and therefore simply drops out of the difference in energies between the liquid and the solid phases which is the interesting quantity in examining the phase transitions between the two. The uncertainties in the electron gas response functions $\chi^{(n)}(\vec{k}_1,\vec{k}_2,\ldots\vec{k}_{n+1})$ will surely affect each of the terms $E_b^{(n)}(\{R_\ell\})$ but, once again, they will not influence too greatly the difference in energies. Thus this particular reformulation, Eq. (2.2), should be a reliable starting point to calculate the energy difference between liquid and solid phases.

For $\chi^{(1)}(k)$ we shall choose the Hubbard-Geldart-Vosko[9] (HGV) form for the dielectric function $\epsilon(k)$ which is known to be of reasonable accuracy at least for $r_s < 2$. For $\chi_2(\vec{k}_1,\vec{k}_2,\vec{k}_3)$ we shall make use of the form used by Brovman, Kagan and Holas[5] in which the one body interactions are screened by the HGV dielectric function. This approximation for $\chi^{(2)}(\vec{k}_1,\vec{k}_2,\vec{k}_3)$ has been used extensively and is believed to be reasonably accurate. The Hamiltonian can now explicitly be written down[10] if we neglect $E_B^{(n)}(\{\vec{R}_\ell\})$ for $n \geq 4$:

$$H = E_\Omega - \frac{\hbar^2}{2m_p}\sum_{i=1}^{N}\nabla^2_{\vec{R}_i} + \sum_{i<j}\phi^{(2)}(R_{ij}) + \sum_{i<j<k}\phi^{(3)}(R_{ij},R_{jk},R_{ik}) \qquad (2.6)$$

where,

$$E_\Omega = E_{eg} - \frac{N}{2n\kappa} + \frac{N}{2(2\pi)^3}\int_{(\text{all }\vec{k})} d\vec{k}\ \frac{4\pi e^2}{k^2}\left(\frac{1}{\epsilon(k)} - 1\right) \tag{2.7}$$

is a large volume dependent term, which is convenient to separate out. In (2.7) n is the number density (N/Ω) and κ is the compressibility of the uniform interacting electron gas neutralized by a uniform positive background at the same density. Note that the terms $E_b^{(2)}(\{\vec{R}_\ell\})$ and V_{pp} have been combined to give

$$\phi^{(2)}(R_{ij}) = \frac{1}{(2\pi)^3}\int d\vec{k}\ \frac{4\pi e^2}{k^2}\ \frac{1}{\epsilon(k)}\ e^{i\vec{k}\cdot(\vec{R}_i-\vec{R}_j)}. \tag{2.8}$$

an effective <u>linear-response</u> pair potential. Finally the third order term[11] is given by,

$$\phi^{(3)}(R_{ij},R_{j\ell},R_{i\ell}) = -\frac{1}{(2\pi)^6}\int dk_1\int dk_2 e^{i\vec{k}_1\cdot\vec{R}_i+\vec{k}_2\cdot\vec{R}_j-i(\vec{k}_1+\vec{k}_2)\cdot\vec{R}_\ell}\ \tilde{\Lambda}(\vec{k}_1,\vec{k}_2,-\vec{k}_1-\vec{k}_2) \tag{2.9}$$

Here $\tilde{\Lambda}$ is:

$$\tilde{\Lambda}(\vec{k}_1,\vec{k}_2,\vec{k}_3) = \frac{(4\pi e^2)^3}{k_1^2k_2^2k_3^2\epsilon(k_1)\epsilon(k_2)\epsilon(k_3)}\ \Lambda(\vec{k}_1,\vec{k}_2,\vec{k}_3)\ , \tag{2.10}$$

$$\Lambda(\vec{k}_1,\vec{k}_2,\vec{k}_3) = \left(\frac{2m^2}{3\pi^2\hbar^4}\right)\left(\frac{k_R^2}{k_1k_2k_3}\right)\Bigg[\sum_{i=1}^{3}\cos\theta_i\,\ell n\left|\frac{2k_F+k_i}{2k_F-k_i}\right| - 2\theta(k_F-k_R)\tan^{-1}\Delta A - \left\{1-\theta(k_F-k_R)\right\}\ell n\left|\frac{1-\Delta A}{1+\Delta A}\right|\Bigg]\ , \tag{2.11}$$

where $\theta(x) = 1$ for $x \geq 0$ and zero for $x < 0$. The remaining parameters are given below,

$$A = \frac{k_1k_2k_3}{(2k_F)^3}\left[1 - \tfrac{1}{2}\frac{k_1^2+k_2^2+k_3^2}{(2k_F)^2}\right]^{-1}\ , \tag{2.12}$$

$$\Delta = \left|\frac{k_F^2 - q_R^2}{q_R^2}\right|^{\frac{1}{2}} ,$$

$$q_R = \frac{k_1 k_2 k_3}{2\left[k_1^2 k_2^2 - (\vec{k}_1 \cdot \vec{k}_2)^2\right]^{\frac{1}{2}}} , \tag{2.13}$$

$$\cos\theta_1 = -\frac{\vec{k}_2 \cdot \vec{k}_3}{k_2 k_3} , \tag{2.14}$$

$$\cos\theta_2 = -\frac{\vec{k}_3 \cdot \vec{k}_1}{k_3 k_1} , \tag{2.15}$$

and

$$\cos\theta_3 = -\frac{\vec{k}_2 \cdot \vec{k}_1}{k_2 k_1} . \tag{2.16}$$

If we take $\epsilon(k)$ to be the RPA dielectric function then $\tilde{\Lambda}$ would precisely be the RPA approximation for the three tailed diagram.

As mentioned earlier the dielectric function $\epsilon(k)$ is taken to be of the HGV form and is explicitly given as,

$$\epsilon(\eta) = 1 + \frac{\alpha F(\eta)/\eta^2}{1 - \alpha F(\eta)/(2\eta^2 + g)} , \tag{2.17}$$

where

$$F(\eta) = 1 + \frac{(1-\eta^2)}{2\eta} \ln\left|\frac{1+\eta}{1-\eta}\right| , \tag{2.18}$$

$$\alpha = (r_s/2\pi)(4/9\pi)^{1/3} , \tag{2.19}$$

$$g = \frac{1}{\left(1 + 0.031\left(\frac{4}{9\pi}\right)^{1/3}\frac{\pi r_s}{2}\right)} \tag{2.20}$$

and $\eta = k/2k_F$.

Finally, we obtain

$$H = E_\Omega + H^{(2)} + \sum_{i<j<k} \phi^{(3)}(R_{ij}, R_{jk}, R_{ik}) \tag{2.21}$$

where E_Ω is a constant volume dependent term and we have split off the $\phi^{(3)}$ term from $H^{(2)}$ given by

$$H^{(2)} = -\frac{\hbar^2}{2m_p}\sum_{i=1}^{N}\nabla_i^2 + \sum_{i<j}\phi^{(2)}(R_{ij}) \tag{2.22}$$

In Ref. 1, H was approximated by $E_\Omega + H^{(2)}$. We proceed from this point and shall first attempt to diagonalize $H^{(2)}$ as well as possible with a one parameter variational function which, as we shall see, will give an error of no more than 4% when compared to the calculation of Ref. 1 employing 6 variational parameters. An optimum wavefunction obtained in this way will be used to calculate the variational bound for the contribution from $\phi^{(3)}$.

3. CALCULATIONAL TECHNIQUE

In this section we shall outline the method used in calculating the ground state energy of the Fermi liquid corresponding to the Hamiltonian given in Eq. (2.6). A Jastrow-Slater variational wavefunction[2,3]

$$\underline{\bar{\Psi}}\,(1,2,\ldots N) = D\,\underline{\bar{\Psi}}_o^B \tag{3.1}$$

will be used to calculate an upperbound to the ground state energy. In Eq. (3.1) D is a Slater determinant made out of plane waves and $\underline{\Psi}_o^B$ is a symmetric correlating factor designed to take care of the strong inter-particle interactions. It is responsible for a large part of the energy. A subsequent Wu-Feenberg expansion[2,3] then uses an <u>exact</u> transformation to recast the problem into the calculation of two distinct parts: Thus we shall set

$$E = E_B + E_{ex} \tag{3.2}$$

where E_{ex} is the exchange contribution and E_B is the eigenvalue of a symmetric ground state corresponding to the Hamiltonian. Then

$$H(\{R_\ell\})\,\underline{\bar{\Psi}}_o^B = E_B\underline{\bar{\Psi}}_o^B \tag{3.3}$$

where $\underline{\psi}_o^B$ in Eq. (3.1) is chosen to be the eigenfunction of (3.3). The calculation of E_B therefore does not involve the antisymmetric factor and results in a considerably simplified problem. A knowledge of this $\underline{\psi}_o^B$ is then utilized to calculate,

$$E_{ex} = \frac{\hbar^2}{2m} \sum_{\ell=1}^{N} \frac{\int \underline{\psi}_o^{B^2} \nabla_\ell D^* \cdot \nabla_\ell D \, d\vec{r}_1 \ldots d\vec{r}_N}{\int \underline{\psi}_o^{B^2} \, d\vec{r}_1 \ldots d\vec{r}_N} \tag{3.4}$$

which may be calculated by a statistical cluster expansion of the type

$$E_{ex} = E_F^{01} + E_F^{02} + E_F^{03} + \ldots \tag{3.5}$$

where $E_F^{(0n)}$ involves n-particle exchange. These terms are easily calculated (at least up to the 3rd order) as we shall see below. The entire procedure is meaningful when E_B is much greater than E_{ex} and the series in E_{ex} converges rapidly. We shall see later that the first condition is very well satisfied, E_B being several orders of magnitude larger than E_{ex}. However, the second is only moderately well satisfied, each term dropping by a factor of 1/3 to 1/5 of the previous term.

So far we have implicitly assumed a paramagnetic ground state, each level being doubly occupied in the Slater determinant. However, it is easy to extend the result to a departure from double occupancy[2,3,12]. The resulting form for $E_{ex}(x)$ is then

$$E_{ex}(x) = E_F^{01}(x) + E^{02}(x) + E^{03}(x) + \ldots \tag{3.6}$$

where x is the spin imbalance order parameter defined by,

$$x = \frac{N_+ - N_-}{N} \tag{3.7}$$

Here $N_+(N_-)$ are the numbers of up (down) spins and N is the total number of spins. A non zero value of x will signify a magnetically ordered phase Clearly x = 1 will represent a ferromagnetically ordered phase. Notice that E_o^B does not depend on x. We shall try to determine whether $E_{ex}(x)$ possesses a minimum $E_{ex}^m(xm)$ at a non-zero value of x. It will turn out that the energy difference $\Delta E(x) = E_{ex}(x=0) - E_{ex}^m(xm)$ per particle is small, only $\sim 2 \times 10^{-5}$Ry. (It is worth noting that this is not small on the scale of a superconducting pairing energy.)

4. VARIATIONAL METHOD

From the variational point of view E_B in Eq. (3.2) is conveniently split into three parts

$$E_B = E_B^{(2)} + E_B^{(3)} + \Delta E_B^{Ph} \ . \tag{4.1}$$

The first term, $E_B^{(2)}$, is calculated by variationally optimizing the Hamiltonian $H^{(2)}(\{R_\ell\})$ with the many-body Jastrow wavefunction given by,

$$\underline{\psi}_o^B = \prod_{i<j} e^{-\frac{1}{2}u(r_{ij})} \tag{4.2}$$

where,

$$u(r) = (\frac{b}{r})^3 \ e^{-(r/b)^3} \tag{4.3}$$

This wave function is a simplified one-parameter form for that used in Ref. 1. The energy functional is minimized with respect to the parameter b at every value of r_s, the resulting wavefunction is then used to calculate the expectation value of $\phi^{(3)}(\{\vec{R}_\ell\})$. The $E_B^{(3)}$ obtained in this first order perturbation is also a variational bound. The u(r) expressed in Eq. (4.3) is short ranged and does

not include the contribution due to the long wavelength phonons. This is done perturbatively with the help of Chester-Reatto wavefunction[13]. The relevant formulae are summarized below:

$$E_B^{(2)}/N = \frac{1}{\pi}\left(\frac{m_e}{m_p}\right)\left(\frac{9\pi}{4}\right)^{2/3} \frac{1}{r_s^2} \int_o^\infty dx g_B^o(x) \left(\frac{b_F}{x}\right)^3 e^{-(x/b_F)^3} \left\{3\left(\frac{x}{b_F}\right)^6 + 2\left(\frac{x}{b_F}\right)^3 + 2\right\}$$

$$+ \frac{8}{(3\pi^2)} \left(\frac{9\pi}{4}\right)^{1/3} \frac{1}{r_s} \int_o^\infty dx\, x^2 v_o(x) g_B^o(x) \quad . \tag{4.4}$$

$$= T_B^{(2)}/N + P_B^{(2)}/N$$

where all distances are scaled with respect to the inverse Fermi wavevector, $1/k_F$, including the variational parameter b ($b = b_F/k_F$). In Eq. (4.4), r_s denotes the average interparticle distance scaled by the Bohr radius and $g_B^o(x) \equiv g_B^o(r)$, ($r = x/k_F$) is the pair correlation function defined as:[2,3]

$$g_B^o(r_{12}) = \frac{N(N-1)}{n^2} \quad \frac{\int (\psi_o^B)^2 d\vec{r}_3 \ldots d\vec{r}_N}{\int \psi_o^{B^2} d\vec{r}_1 \ldots d\vec{r}_N} \quad . \tag{4.5}$$

Note that ψ_o^B is defined in Eqs.(4.2) and (4.3). The corresponding static structure factor $S_o^B(k)$ is defined by the Fourier transform:[2,3]

$$S_B^o(k) = 1 + n\int d\vec{r}\, e^{i\vec{k}\cdot\vec{r}} \, [g_B^o(r) - 1] \tag{4.6}$$

Finally with the distance and the wavevector scaled,

$$v_o(x) = \int_o^\infty dy \, \frac{\sin xy}{xy} \quad \frac{1}{\epsilon(y)} \quad = \quad \frac{\pi}{2e^2 k_F} \, \phi^{(2)}(x) \tag{4.7}$$

is the screened interaction and $\epsilon(y)$ is the HGV dielectric function. Once again all wavevectors are scaled by k_F ($|k| = yk_F$). For $g_B^o(r)$ we shall use the

Hypernetted Chain Approximation[2,3] which is known to be satisfactory for Bose fluids and has been tested for a variety of interaction potentials.[2,3] In this approximation $g_B^o(r)$ is the solution of the non-linear integral equation relating the direct correlation function $c(r)$ to $g_B^o(r)$:

$$g_B^o(r) - 1 = c(r) + n\int d\vec{r}' c(|\vec{r}-\vec{r}'|)\ [g_B^o(r')-1], \tag{4.8}$$

$$c(r) = g_B^o(r) - 1 - \log g_B^o(r) + u(r) \tag{4.9}$$

The procedure is to solve Eqs.(4.8) and (4.9) for a given value of the variational parameter b by a standard numerical procedure and to use the resulting $g_B^o(r)$ in Eq. (4.4) to calculate the energy. This process is repeated for a number of different values of b to find the optimum $g_B^o(r)$, $u(r)$ and the minimum in energy at a given density or r_s. We then proceed to calculate the contribution due to $\phi^{(3)}(\{\vec{R}_\ell\})$. Thus

$$E_B^{(3)} = \frac{\langle \underline{\psi}_o^B |\ H^3(\{R_\ell\})\ |\underline{\psi}_o^B\rangle}{\langle \underline{\psi}_o^B | \underline{\psi}_o^B\rangle}$$

$$= -\frac{e^6}{\pi^3}\int d\vec{k}\int d\vec{q}\ \frac{1}{q^2\epsilon(q)}\ \frac{1}{k^2\epsilon(k)}\ \frac{1}{(\vec{q}+\vec{k})^2\epsilon(\vec{q}+\vec{k})}\ S_B^o(\vec{k},\vec{q},-\vec{k}-\vec{q})\Lambda(\vec{k},\vec{q},-\vec{k}-\vec{q}) \tag{4.10}$$

where,

$$S_B^o(\vec{k},\vec{q},-\vec{k}-\vec{q}) = \frac{\langle \underline{\psi}_o^B | \rho_{\vec{k}}\rho_{\vec{q}}\rho_{-\vec{k}-\vec{q}} | \underline{\psi}_o^B\rangle}{\langle \underline{\psi}_o^B | \underline{\psi}_o^B\rangle} \tag{4.11}$$

and

$$\rho_{\vec{k}} = \sum_{i=1}^{N} e^{-i\vec{k}\cdot\vec{r}_i}\ , \quad \vec{k} \neq 0\ . \tag{4.12}$$

A distinct feature[11] of the response function of $\Lambda(\vec{k},\vec{q},-\vec{k}-\vec{q})$ is its singular behavior when $\vec{k} + \vec{q} = 0$: i.e.,

$$\Lambda(\vec{k},-\vec{k},0) \sim \ell n \left|k_F^2 - k^2/4\right| . \tag{4.13}$$

This singularity is stronger here than in the second order response where only the derivative has a logarithmic singularity. This amplification is due to the confluence of the usual second order Kohn anomaly which is always present in the third order response and the intrinsic singularity of the third order response. It is clear that the integral in Eq. (4.10) can only be defined if this singularity is cancelled by other terms present in the integrand. To this effect we prove rigorously in the Appendix the following result: $\lim_{k\to 0} S_B(\vec{k},\vec{\ell},-\vec{k}-\vec{\ell}) \to \alpha k$ if $\lim_{k\to 0} S_B(k) \to \alpha k$. Similar results hold when $\ell \to 0$ and $|\vec{k}+\vec{\ell}| \to 0$.

Thus it is necessary that S(k) vanish at least linearly with k in the limit of small k. Furthermore, any approximation for the three particle structure factor must be such as to preserve this property. One such approximation is the convolution approximation[2,3] for the three particle structure factor, an approximation that has been extensively tested for soft core potentials[14] and in many other situations.[14] Thus we set

$$S_B(\vec{k},\vec{q},-\vec{k}-\vec{q}) \sim S_B(k) S_B(q) \; S_B(\vec{k}+\vec{q}) \tag{4.14}$$

which clearly has the required property that it vanishes when any of the three arguments vanaishes. As is made clear in the appendix this is simply because of the fact that the convolution approximation satisfies all the normalization conditions to be required of the probability distribution functions. However, as is well known[2,3], the short range wavefunction written down in Eq. (4.3) does <u>not</u> lead to a $S_B(k)$ which vanishes as $k \to 0$. This needs to be corrected

for the presence, expected physically, of long range phonons before we can evaluate the third order energy given by Eq. (4.10) and (4.14). The procedure is almost standard[15]. The Chester and Reatto wavefunction is long ranged and has the form

$$e^{-\frac{1}{2}U_{Lr}(r)} = e^{-\frac{3m_p c}{\hbar k_F}\frac{1}{(x^2+x_o^2)}} \tag{4.15}$$

where we have scaled the distance by k_F i.e. $r - x/k_F$ and x_o is a variational cutoff parameter. Here c is the velocity of sound in this hypothetical Boson system and can be obtained from the energy, $E_B^{(2)}/N$:

$$c(r_s) = \frac{c_{BS}}{\sqrt{3}}\left\{r_s\left(\frac{4}{9\pi}\right)^{1/3}\left(\frac{r_s}{2}\frac{d^2E_B^{(2)}}{dr_s^2} - r_s\frac{dE_B^{(2)}}{dr_s}\right)^{\frac{1}{2}}\right\} \tag{4.16}$$

where, $c_{BS} = \frac{v_F}{\sqrt{3}}\left(\frac{m_e}{m_p}\right)^{\frac{1}{2}}$ and $v_F = (\hbar k_F/m_e)$. The choice of such a long range wavefunction leads to a sequence of changes given next. The structure factor $S_B^o(k)$ calculated with the short ranged wavefunction gets modified to $S_B(k)$ given by

$$S_B(k) = \frac{S_B^o(k)}{1 + n\cdot S_B^o(k)\ U_{LR}(k)} \tag{4.17}$$

and the corresponding correction in the pair correlation function is

$$\delta g(r) = g_B^o(r)\ (e^{-\Gamma(r)} - 1)\ , \tag{4.18}$$

where

$$g_B(r) = g_B^o(r) + \delta g(r), \tag{4.19}$$

and $U_{LR}(k)$ is the Fourier transform of $U_{LR}(r)$. Finally,

$$\Gamma(r) = \frac{1}{(2\pi)^3}\int e^{i\vec{k}\cdot\vec{r}}\ \frac{S_B^0(k)^2\ U_{LR}(R)}{1+\rho U_L(k)\ S_B^0(k)}\ d\vec{k} \tag{4.20}$$

The correction to the energy is then

$$\frac{\Delta E_B^{Ph}}{N} = \frac{\hbar^2\rho}{8m}\int d\vec{r}\ g_B^0(r)\ \nabla^2 U_{LR}(r) + \frac{\hbar^2\rho}{8m}\int d\vec{r}\,\delta g(r)\ \nabla^2[U(r)+U_{LR}(r)]$$

$$+ \tfrac{1}{2}\rho\int v(r)\ \delta g(r)d\vec{r} \tag{4.21}$$

Finally, Eq. (4.10) can be rewritten to obtain the third order contribution to the energy,

$$\frac{E_B^{(3)}}{N} = -\frac{8e^6}{\pi}\int_0^\infty dk\ \frac{S_B(k)}{\epsilon(k)}\int_0^\infty dq\ \frac{S_B(q)}{\epsilon(q)}\int_0^\pi \sin\theta d\theta \frac{1}{(q+\vec{k})^2}\ \frac{1}{\epsilon(\vec{k}+\vec{q})}\ S_B(\vec{k}+\vec{q})\Lambda(\vec{k},\vec{q},-\vec{k}-\vec{q}) \tag{4.22}$$

where θ is the angle between the vectors $\vec{k}$ and $\vec{q}$. Thus $E_B^{(3)}/N$ can now be calculated numerically if $S_B(q)$ is known.

5. EXCHANGE CONTRIBUTIONS

As mentioned earlier the Wu-Feenberg expansion is used to obtain the exchange contributions to the energy. The total energy per particle is

$$E(x)/N = E_B/N + E_{ex}/N$$

$$= (E_B^{(2)} + E_B^{(3)} + \Delta E_B^{Ph})/N + E_{ex}(x)/N \tag{5.1}$$

where, $E_{ex}(x)/N$ is the exchange energy of the Fermions (protons in this case). In Eq. (5.1) the energy up to third order in exchange is given by:

$$E_{ex}/N = E_{01}^F(n,x)/N + E_{02}^F(n,x)/N + E_{03}^F(n,x)/N + \ldots \tag{5.2}$$

where

$$E_{01}^{F}(n,x)/N = \frac{3}{10}\, e_F\left[(1+x)^{5/3}+(1-x)^{5/3}\right] \tag{5.3}$$

$$E_{02}^{F}(n,x)/N = 12e_F\Big\{(1+x)^{8/3}\int_0^1 (y^4 - \tfrac{3}{2}y^5 + \tfrac{1}{2}y^7)\,[S(2k_F^{+}y)-1]dy$$

$$+ (1-x)^{8/3}\int_0^1 (y^4 - \tfrac{3}{2}y^5 + \tfrac{1}{2}y^7)\,[S(2k_F^{-}y)-1]dy \tag{5.4}$$

and

$$E_{03}^{F}(n,x)/N = -\frac{e_F}{2}\left(\frac{3}{8\pi}\right)^3\Big\{(1+x)^{11/3}\int_{y_i<1} y_{12}^2 S(k_F^{+}y_{12})\,[S(k_F^{+}y_{23})-1]\,[S(k_F^{+}y_{13})-1]d\vec{y}_1 d\vec{y}_2 d\vec{y}_3$$

$$+ (1-x)^{11/3}\int_{y_i<1} y_{12}^2 S(k_F^{-}y_{12})\,[S(k_F^{-}y_{23})-1]\,[S(u_F^{-}y_{13})-1]d\vec{y}_1 d\vec{y}_2 d\vec{y}_3\Big\} \tag{5.5}$$

Note that $e_F = \dfrac{\hbar^2 k_F^2}{2m_p}$, $k_F^{\pm} = k_F(1 \pm x)^{1/3}$ and $x = (N_+ - N_-)/N$. As mentioned earlier our intention is to compute the ground state energy as a function of x. The term E_{03}^{F} is calculated by making the quadratic approximation described in Refs. 2 and 12.

6. RESULTS

In Fig. 1 we show the dimensionless potential function $v_o(x)$, Eq. (4.7), for some typical values of r_s. In Fig. 2 we show the corresponding pair correlation functions $g_B(r)$. The actual Fermion pair correlation function can be obtained from these by the Wu-Feenberg expansion[2,3], Fermion corrections being small in this case. The reason why we have not displayed them is because they are not explicitly required in the method of calculating the Wu-Feenberg series used here. The structure factor $S_B(k)$ corresponding to $g_B(r)$ is shown in Fig. 3 for few typical values of r_s. It is clear from these plots that there is a considerable amount of short range order in liquid metallic hydrogen as compared to say liquid helium. One should also note that the interaction potential exhibits a strong density dependence.

Table 1 compares our results for $E_B^{(2)}$, Eq. (4.4), with the calculation in Ref. 1. It is clear that our one parameter variational wavefunction gives

a reasonably good upperbound. Also shown in the table is the detailed decomposition of $E_B^{(2)}$ into kinetic and potential energies. We should emphasize that precise agreement between our 1-parameter variational results with the 6-parameter Monte Carlo results, Ref. 1, is not necessary since we are simply interested in an upperbound for the contribution arising from the three body forces. These are given in Table 1 along with the volume dependent terms. In calculating E_Ω and E_{eg} we have made use of the Nozieres and Pines interpolation[16] formula for the correlation energy of electron gas which is consistent with our choice of HGV dielectric function. From Table 1 one can also see that $\Delta E_B^{Ph}/N$, Eq. (4.21), makes a negligible contribution to the total energy. The main effect of the long range phonons is to produce an $S_B(k)$ which vanishes in the limit of small k which, in turn, allows us to calculate $E_B^{(3)}/N$, Eq. 4.22. As noted above the integral is ill conditioned if $S_B(k)$ approaches a non zero value as k goes to zero.

In Table 2 we have shown the exchange corrections. It is seen that a partially spin aligned state of protons is in fact favored throughout the entire range of densities considered. As mentioned earlier we should be cautious about this conclusion since E_{03}^F has been calculated with the help of the conventional[2,10,12] quadratic approximation, and thus may be quite inaccurate especially for larger values of the order parameter x. In view of the fact that this term is considerably smaller than the rest and that one needs a complicated numerical procedure to calculate accurately we have not examined it using a more elaborate computational method. We do not believe that the results will change qualitatively. Since the quadratic approximation is good in the neighborhood of x = o, the fact that the energy is lowered for non zero values of x can be established although the exact value of x may be inaccurate. It is also worth remembering that the convergence of Wu-Feenberg series is not rigorously established.

The total energy for the liquid is compared, Table 3, with the static energies for the solid phase obtained by Hammerberg and Ashcroft[4]. Note that the static

hydrogen[16] could easily be of the order of 0.01Ry. The contribution of the third order term in the liquid is more significant than in the solid. For example at $r_s = 1.6$, the third order energy in the liquid is -0.0372Ry as opposed to -0.0322 calculated by Hammerberg and Ashcroft. The corresponding comparison at $r_s = 1.36$, yields -0.0326Ry for liquid as opposed to -0.0281 for the solid[17]. Finally, the liquid state energies calculated in this paper are a variational upperbound and the exact energy is expected to be lower. Thus one cannot in principle exclude the existence of a liquid ground state of metallic hydrogen though it is certainly not established as a preferred ground state.

7. CONCLUSION

We have investigated the possibility for a liquid ground state of metallic hydrogen at zero temperature. We conclude that the possibility of a liquid phase near the metastable zero pressure point cannot be ruled out. We have found out that the third order terms in the liquid are significantly lower than the corresponding ones in the solid and a careful estimate of these terms in the solid phase which also incorporates the dynamics of the protons is essential to determine the liquid-solid transition (if any). We have also found that the contribution to the ground state energy due to the long range phonons is negligible though their presence is necessary. An interesting part of our calculation is the fact that the energy of this proton-electron liquid can be lowered by a partial spin alignment of the protons.

We would like to thank Dr. P. Bhattacharya and Professor G.V. Chester for interesting discussions. This work was supported by NASA, NGR 33-010-188.

Appendix

We shall prove that the limiting value of $S_B(\vec{k},\vec{q},-\vec{k}-\vec{q})$ as any one of the wave vector approaches zero from above vanishes provided the static structure factor $S_B(k)$ vanishes in the same limit. Strictly speaking this result should be considered as a limiting value, defining the function by continuity at the origin and true in the thermodynamic limit.

First note that[2],

$$S_B(\vec{k},\vec{q},-\vec{k}-\vec{q}) = \frac{\langle \underline{\psi}_o^B | \rho_{\vec{k}}\rho_{\vec{q}}\rho_{-\vec{k}-\vec{q}} | \underline{\psi}_o^B \rangle}{N\langle \underline{\psi}_o^B | \underline{\psi}_o^B \rangle}$$

$$= -2+S(k)+S(q)+S(|\vec{k}+\vec{q}|)+ \frac{1}{N}\int e^{i\vec{k}\cdot\vec{r}_1+i\vec{q}\cdot\vec{r}_2-i(\vec{k}+\vec{q})\cdot\vec{r}_3} P(\vec{r}_1,\vec{r}_2,\vec{r}_3) \, d\vec{r}_1 d\vec{r}_2 d\vec{r}_3 \qquad (A1)$$

where the three particle distribution function $P(\vec{r}_1,\vec{r}_2,\vec{r}_3)$ is,

$$P(\vec{r}_1,\vec{r}_2,\vec{r}_3) = \frac{N(N-1)(N-2)}{n^3} \frac{\int \underline{\psi}_o^{B^2} d\vec{r}_4 \ldots d\vec{r}_N}{\int \underline{\psi}_o^{B^2} d\vec{r}_1 \ldots d\vec{r}_N} \qquad (A2)$$

Since $S_B(\vec{k},\vec{q},-\vec{k}-\vec{q})$ is invariant with respect to the interchange of its arguments it is sufficient to prove the result when any one of the wavevectors tend to zero, say $k \to 0^+$. The following cluster decomposition[2] of $P(\vec{r}_1,\vec{r}_2,\vec{r}_3)$ is exact as long as one does not specify $\delta P(\vec{r}_1,\vec{r}_2,\vec{r}_3)$:

$$P(\vec{r}_1,\vec{r}_2,\vec{r}_3) = n^3[1+h(r_{12})+h(r_{13})+h(r_{23})+h(r_{12})h(r_{23})+h(r_{23})h(r_{31}) + h(r_{31})h(r_{32})] + \delta P(\vec{r}_1,\vec{r}_2,\vec{r}_3) \qquad (A3)$$

where, $h(r) = g_B(r) - 1$.

Then one can easily prove from the normalization of the probability distribution functions that[2]

$$\int \delta P(\vec{r}_1, \vec{r}_2, \vec{r}_3) d\vec{r}_3 = -n^3 \int h(r_{13}) h(r_{23}) \, d\vec{r}_3 \tag{A4}$$

Now one can easily evaluate the right hand side of Eq. (A1) for $k \to 0^+$ and obtain the stated result.

FIGURE CAPTIONS

Figure 1 $v_o(r)$ for some typical values of r_s

Figure 2 $g_B(r)$ for some typical values of r_s

Figure 3 $S_B(k)$ for some typical values of r_s

TABLE CAPTIONS

Table 1 Boson part, E_B, of the ground state energy. $E_B^{(2)}$ (MC) is the Monte-Carlo results of Ref. 1. All energies are expressed in units of Rydbergs.

Table 2 Exchange contribution to the ground state energy. All energies are expressed in units of Rydbergs.

Table 3 Comparisons of the ground state energies of the liquid (E(x)/N) and the solid phases (E^S(HA)/N: Hammerberg and Ashcroft, Ref. 4). All energies are expressed in units of Rydbergs. SC: Simple cubic; BCC: Body centered cubic; FCC: Face centered cubic.

REFERENCES

1. K.K. Mon, G.V. Chester, and N.W. Ashcroft, to be published.

2. E.Feenberg, Theory of Quantum Fluids (Academic, New York, 1969).

3. C.W. Woo in Physics of Liquid and Solid Helium, edited by K.H. Benneman and J.B. Ketterson (Wiley Interscience, New York, 1976).

4. J. Hammerberg and N.W. Ashcroft, Phys. Rev. B 9, 409 (1974).

5. E.G. Brovman, Yu. Kagan and A. Holas, Zh. Eksp. Teor. Fiz. 61, 2429 (1971) [Sov. Phys. - JETP 34, 1300 (1972)]; Yu. Kagan, V.V. Pushkarev and A. Holas, Zh. Eksp. Teor. Fiz. 73, 967 (1977).

6. The expansion parameter for the Born-Oppenheimer approximation, $(m_e/m_p)^{\frac{1}{4}}$, is in the case of hydrogen somewhat larger than other common metals. However, we feel that the approximation will not affect the difference in energies between the liquid and the solid phases.

7. It is worthwhile to expand on what we mean by many body forces. For example $E_b^{(3)}(\{\vec{R}_\ell\})$ will contain a pair as well as a three-body interaction, similarly for the higher order terms. For computational purpose there is no need to make this decomposition.

8. E.G. Brovman and Yu. Kagan, Usp. Fiz. Nauk 112, 369 (1974) [Sov. Phys. Usp. 17, 125 (1975)].

9. D.J.W. Geldart and J.H. Vosko, Can. J. Phys. 44, 2137 (1966).

10. These terms are known to be very small in a static crystal. See, for example Refs. 4 and 5. We are assuming that such terms would be equally small in the liquid phase.

11. E.G. Brovman and Yu. Kagan, Zh. Eksp. Teor. Fiz. 63, 1937 (1972) [Sov. Phys. JETP 36, 1025 (1972)].

12. F.Y. Wu and E. Feenberg. Phys. Rev. 128, 943 (1962); C.W. Woo, Phys. Rev. 151, 138 (1966); G. Kaiser and F.Y. Wu, Phys. Rev. 6, 2369 (1972); M.D. Miller and R. Guyer, Phys. Rev. (To be published)

13. G.V. Chester and L. Reatto, Phys. Letters 22, 276 (1966).

14. Sudip Chakravarty and C.W. Woo, Phys. Rev. B 13, 4815 (1976). Also see extensive discussions and numerous applications in Ref. 2.

15. M.H. Kalos, D. Levesque and L. Verlet, Phys. Rev. A 9, 2178 (1974).

16. D.M. Straus and N.W. Ashcroft, Phys. Rev. Lett 38, 415 (1977).

17. D.M. Straus , Thesis, Cornell University, Materials Science Laboratory Report no. 2739 (unpublished).

TABLE 1

r_s	b_F	$T_B^{(2)}/N$	$P_B^{(2)}/N$	$E_B^{(2)}/N$	$E_B^{(2)}(MC)/N$	$\Delta E_B^{Ph}/N$	$E_B^{(3)}/N$	E_Ω/N
0.50	5.35	0.07406	2.76268	2.83674		-0.00158	-0.01442	0.54062
0.80	5.55	0.03195	0.76254	0.79449	0.7943	-0.00054	-0.02120	-0.86188
1.20	5.50	0.01386	0.19986	0.21372	0.2079	-0.00021	-0.02944	-1.10353
1.30	5.435	0.01143	0.14616	0.15759				
1.36	5.40	0.01026	0.12104	0.13130	0.1262	-0.00016	-0.03258	-1.10050
1.40	5.37	0.00954	0.10665	0.11619				
1.45	5.315	0.00865	0.09095	0.09960				
1.488					0.0847			
1.50	5.28	0.00794	0.07726	0.08520		-0.00012	-0.03528	-1.08394
1.55	5.225	0.00723	0.06543	0.07266				
1.60	5.175	0.00661	0.05510	0.06171	0.0592	-0.00011	-0.03718	-1.06790
1.70	5.05	0.00549	0.03824	0.04373		-0.00009	-0.03908	-1.04988
1.80	4.9	0.00452	0.02531	0.02983		-0.00008	-0.04100	-1.03074

TABLE 2

r_s	X	$E_{ex}(x)/N$
0.50	0.589	0.00263
0.80	0.579	0.00102
1.20	0.582	0.00045
1.30	0.585	0.00039
1.36	0.587	0.00035
1.40	0.588	0.00033
1.45	0.591	0.00031
1.50	0.593	0.00029
1.55	0.595	0.00027
1.60	0.598	0.00026
1.70	0.603	0.00023
1.80	0.607	0.00021

TABLE 3

r_s	$E^S(HA)/N$			$E(x)/N$
	SC	FCC	BCC	
0.50				3.36399
0.80				-0.08811
1.00	-0.71188	-0.71929	-0.71819	
1.20	-0.93796	-0.94019	-0.93902	-0.91901
1.25	-0.96842	-0.96961	-0.96843	
1.30	-0.99217	-0.99242	-0.99122	
1.36				-1.00159
1.50	-1.04104	-1.03818	-1.03693	-1.03385
1.60	-1.04759	-1.04345	-1.04222	-1.04322
1.65	-1.04803	-1.04338	-1.04209	
1.70				-1.04509
1.80				-1.04178

THE ASTROPHYSICAL JOURNAL SUPPLEMENT SERIES, 35 221–237, 1977 October

THE PHASE DIAGRAM AND TRANSPORT PROPERTIES FOR HYDROGEN-HELIUM FLUID PLANETS

D. J. STEVENSON AND E E SALPETER
Center for Radiophysics and Space Research and Physics Department, Cornell University
Received 1976 June 23, accepted 1977 April 13

ABSTRACT

Hydrogen and helium are the major constituents of Jupiter and Saturn, and phase transitions can have important effects on the planetary structure. In this paper, the relevant phase diagrams and microscopic transport properties are analyzed in detail. The following paper (Paper II) applies these results to the evolution and present dynamic structure of the Jovian planets

Pure hydrogen is first discussed, especially the nature of the molecular-metallic transition and the melting curves for the two phases It is concluded that at the temperatures and pressures of interest ($T \approx 10^4$ K, $P \approx 1$–10 Mbar), both phases are fluid, but the transition between them might nevertheless be first-order The insulator-metal transition in helium occurs at a much higher pressure ($\sim$70 Mbars) and is not of interest.

The phase diagrams for both molecular and metallic hydrogen-helium mixtures are discussed. In the metallic mixture, calculations indicate a miscibility gap for $T \leqslant 10^4$ K. Immiscibility in the molecular mixture is more difficult to predict but almost certainly occurs at much lower temperatures. A fluid-state model is constructed which predicts the likely topology of the three-dimensional phase diagram The greater solubility of helium in the molecular phase leads to the prediction that the He/H mass ratio is typically twice as large in the molecular phase as in the coexisting metallic phase Under these circumstances a "density inversion" is possible in which the molecular phase becomes more dense than the metallic phase

The partitioning of minor constituents is also considered: The deuterium/hydrogen mass ratio is essentially the same for all coexisting hydrogen-helium phases, at least for $T \geqslant 5000$ K The partitioning of H_2O, CH_4, and NH_3 probably favors the molecular (or helium-rich) phase. Substances with high conduction electron density (e g , Al) may partition into the metallic phase

Electronic and thermal conductivities, viscosity, helium diffusivity, and Soret coefficient are evaluated for the fluid molecular and metallic phases, all to at least order-of-magnitude accuracy The properties of the metallic phase are typical of a liquid alkali metal, and those of the molecular phase are typical of a dense neutral fluid (except that the conductivities may be almost metallic at the transition pressure) The opacities of molecular hydrogen and solar-composition mixtures are discussed for $T \approx 500$ K, where molecular hydrogen alone may be insufficiently opaque to ensure convection in the Jovian planets. Sufficient opacity to initiate convection is probably supplied by the minor constituents. Current uncertainties are assessed

Subject headings: equation of state — planets: interiors

I. INTRODUCTION

Hydrogen and helium comprise roughly 85% of the total planetary mass in our solar system, and are the major constituents of Jupiter and Saturn. They are also the simplest atomic species, so their thermodynamic and transport properties should be amenable to first-principles calculation at those pressures which are presently unattainable by experiment

There has been recent intensive modeling of the interior of Jupiter by several groups (Podolak and Cameron 1975; Zharkov *et al* 1975, Hubbard and Slattery 1976; Stevenson and Salpeter 1976; Podolak 1977), and much attention has been given to the equation of state and other thermodynamic derivatives for hydrogen and hydrogen-helium mixtures. However, all these models assume a *homogeneous* mixture of hydrogen and helium This assumption may be fundamentally incompatible with the phase diagram of hydrogen-helium mixtures

The present paper and the following paper (Stevenson and Salpeter 1977, hereafter Paper II) consider in detail the phase diagram for hydrogen-helium mixtures, and its implications for the interiors of the Jovian planets. Since these implications depend on details of the transport (including fluid-dynamical) processes, the present paper also contains a survey of the current knowledge of the microscopic transport properties of dense hydrogen-helium mixtures

The present paper concentrates on the condensed-matter physics of such mixtures, with emphasis given to the pressure-temperature domain appropriate to Jupiter and Saturn The emphasis is on the fluid state, which is almost certainly applicable to the

present interiors of Jupiter and Saturn, but there is also a discussion of melting curves for the hydrogen-helium phases. Since the Jovian planets contain constituents other than hydrogen and helium, the effects of these are considered briefly. The equation of state and other thermodynamic derivatives are not discussed in detail here, but an extensive review is to be found elsewhere (Stevenson and Salpeter 1976).

In § II, we discuss the properties of pure hydrogen and helium, especially the melting curves and insulator-metal transitions. The nature of the molecular-metallic hydrogen phase transition is not yet well understood, but is expected to occur at 2 Mbar $\lesssim P \lesssim$ 4 Mbar and to be first-order at least until $T \approx 10^3$ K and quite possibly even for $T \gtrsim 10^4$ K. At 10^4 K, the two phases are certainly both fluid. The insulator-metal transition in helium occurs at $P \approx 70$ Mbar, which is too high to be of interest for the Jovian planets.

In § III, calculations (Stevenson 1975) for the phase diagram of metallic hydrogen-helium mixtures are reviewed. A miscibility gap is predicted for a solar composition mixture at megabar pressures and temperatures less than 10^4 K

In § IV, the phase diagram of molecular hydrogen-helium mixtures is discussed Unlike the metallic phase, where an essentially first-principles calculation can be made, calculations for the molecular phase must rely on semiempirical intermolecular potentials, and are necessarily suspect However, the prediction that helium is more soluble in molecular hydrogen than in metallic hydrogen is reliable

In § V, the conclusions of the previous sections are used to model a total phase diagram which simultaneously accounts for the first-order character of the molecular-metallic hydrogen transition, the limited solubility of helium, and the thermodynamic preference for helium to be dissolved in the molecular hydrogen rather than metallic hydrogen phase This model may be numerically imprecise, but is expected to predict the correct topology of the (three-dimensional) phase diagram. The predicted phase diagrams are similar to those suggested by Smoluchowski (1973). This model contains two other useful features: First, it predicts the circumstances for which a "density inversion" occurs (i e., when a helium-poor metallic phase is less dense than a coexisting helium-rich molecular phase) Second, it predicts the limited range of metastability for the molecular phase in the metallic region, and vice versa.

In § VI, minor constituents are discussed. Immiscibilities appear unlikely, but the partitioning of minor constituents among the various hydrogen-helium phases is undoubtedly nonuniform A special case is deuterium, for which calculations indicate that the deuterium/hydrogen mass ratio in each phase is essentially uniform, at least for $T \gtrsim 5000$ K. A model is proposed for other minor constituents, in which partitioning is in favor of the phase with the most similar electron density at the Wigner-Seitz cell boundary. This model predicts that H_2O, NH_3, and CH_4 prefer molecular or helium-rich phases, but the degree of nonuniform partitioning is probably less than an order of magnitude

Section VII is a summary of the microscopic transport properties of the metallic phase. Electronic and thermal conductivities, viscosity, and helium diffusivity are given particular attention.

In § VIII, the corresponding transport properties of the molecular phase are considered. In addition, the opacities of dense molecular hydrogen and solar-composition mixtures are discussed, especially for temperatures of order 500 K.

Section IX concludes with an assessment of current uncertainties. In the following paper (Paper II), specific thermal and compositional evolutions of a hydrogen-helium planet like Jupiter are discussed semiquantitatively

II THE PURE PHASES

a) Hydrogen

Even at $T = 0$ K, there must be some sufficiently high density for which the Pauli exclusion principle precludes the existence of molecules or localized states and dense hydrogen becomes a Coulomb plasma protons immersed in an almost uniform, degenerate sea of electrons. Wigner and Huntington (1935) pointed out that this atomic state would be analogous to the conventional alkali metals and therefore metallic This atomic state is referred to as "metallic hydrogen" to indicate that its high conductivity is a consequence of itinerant electronic states in a monovalent metal, rather than being a consequence of temperature

If the density is reduced sufficiently and the temperature is low enough, then it becomes thermodynamically favorable to pair the protons in the form of H_2 molecules This is the experimentally accessible molecular phase The transition between the molecular and metallic phases occurs at a pressure given approximately by the dissociation energy per molecule divided by the volume per molecule: a few megabars The molecular phase exists in both solid and liquid forms, and the metallic phase is expected to behave likewise. Additional low-temperature phases that cannot be categorized as either metallic or molecular are not yet rigorously excluded, but neither are they indicated experimentally or theoretically. We discuss below the metallic phase, the molecular phase, and the metallic-molecular transition

1) *Metallic Hydrogen*

The evaluation of the thermodynamics of the alkali metals from first principles is well established for both the solid and fluid phases (see, for example, Stroud and Ashcroft 1972), and the properties of metallic hydrogen can be evaluated in a similar fashion. There are two important respects in which metallic hydrogen is unlike the conventional alkalis: the effective electron-ion interaction is stronger (because there are no core states) and quantum effects for the ions (i.e., protons) are significant (because of the larger electron-ion mass ratio) The former is particularly important at low densities whereas the latter is most important at

high densities and low temperatures Hubbard and Smoluchowski (1973) have an excellent review of earlier work on metallic hydrogen and we comment here on more recent work, with a particular emphasis on the solid-fluid transition.

The most recent calculations for a static metallic hydrogen lattice by a variety of perturbative and nonperturbative techniques are in excellent agreement (Ross and McMahan 1976) The most favored lattice structure has not been established, but this is unimportant for most purposes since the energy difference between structures is so small. It has been suggested that the lowest energy structure is highly anisotropic (Brovman, Kagan, and Kholas 1972), but this conclusion is premature (Hämmerberg and Ashcroft 1974; Ross and McMahan 1976). The finite temperature and zero-point motion corrections are not as well understood (Brovman, Kagan, and Kholas 1972; Caron 1974, Straus and Ashcroft 1977) but appear to be describable by a Debye model in which *two* Debye temperatures are defined—one for the longitudinal modes and one for the transverse modes. Most of these calculations indicate that the transverse modes are "soft," and in some instances the stability of the lattice is in doubt

Recent fluid-state calculations have been made by Hubbard and Slattery (1971), Stevenson (1975), Hansen and Vieillefosse (1976), and Hubbard and DeWitt (1976) As with all simple metals, the thermodynamic derivatives with respect to volume or pressure (e g , the equation of state) are very similar to the solid. Thermodynamic derivatives with respect to temperature (e.g , entropy) are, of course, substantially different from the solid, but the various methods used are substantially in agreement. The results are summarized in Stevenson and Salpeter (1976).

The only rigorous way to calculate the melting temperature of a substance (assuming, of course, that the solid state exists) is by equating the Gibbs free energies for the two phases This is a very difficult procedure since, although the energy of each phase is very accurately known, most of the energy is structure-independent, and the energy *difference* between the phases is very small at all temperatures Pollock and Hansen (1973) used their Monte Carlo results for each phase to deduce a melting temperature T_M for metallic hydrogen and found

$$T_M \approx 1500\rho^{1/3}\ \mathrm{K} \qquad (1)$$

by equating Gibbs energies, where ρ is the density in g cm^{-3} This is probably an upper bound since it does not include the effects of screening on the ion-ion interaction A similar calculation, including screening, has been attempted by Stevenson and Straus (unpublished) using the solid-state free energies of Straus, Ashcroft, and Beck (1977) and the fluid-state free energies of Stevenson (1975) The fluid state appeared to always have lower energy, but the energy difference was found to be comparable to the errors inherent in the calculations The conclusion reached is that equation (1) is indeed an upper bound

Several other methods have been tried for estimating T_M One common method is Lindemann's rule, but this method is unreliable for a substance such as metallic hydrogen, where T_M is *less* than the Debye temperature (Stevenson and Ashcroft 1974). Another method is based on the solidification of the classical hard sphere liquid at 45% packing (Wainwright and Alder 1958), but this method predicts $T_M \approx 1100$ K at $\rho = 1$ g cm^{-3}, a value that may be too low for the classical theory to be applicable (Stevenson 1975)

At sufficiently high densities, where screening is unimportant, the large zero-point motion of the protons precludes a solid at $T = 0$ K. The density above which there is no solid is about 10^4–10^5 g cm^{-3} (Glyde *et al* 1976; Van Horn 1967) This is too high to be of interest in the giant planets Whether screening precludes a solid phase at much lower densities has not yet been established.

If the solid exists at $\rho \approx 1$ g cm^{-3}, then it is most likely a superconductor below about 100 K (Ashcroft 1968, Caron 1974). If no solid exists, then an anisotropic superfluid may be possible However, these low-temperature effects are not relevant to the giant planets where $T \gtrsim 10^4$ K is implied (see Paper II), and the fluid state is ensured without invoking quantum effects Subsequent discussion of the metallic state in this paper is mainly for the fluid

ii) *Molecular Hydrogen*

At $P \lesssim 0.1$ Mbar this phase is quite well understood experimentally, but the experimental uncertainty increases as the pressure increases (Ross 1974). Past theoretical calculations are no more accurate than experiment at the highest pressures because of the failure of the pair potential approximation (Ree and Bender 1974), but recent band structure calculations (Ramaker, Kumar, and Harris 1975; Friedli and Ashcroft 1976) are potentially capable of greater accuracy Nevertheless, it is still necessary for most purposes to resort to semiempirical pair potentials that are compatible with the experimental shock data (Ross 1974) yet are also plausible modifications of first-principles calculations (McMahan, Beck, and Krumhansl 1974) The most recent first-principles calculations of the effective pair potential are by Etters, Danilowicz, and England (1975) and include detailed consideration of the anisotropy of the interaction. They found that the energy associated with molecular orientation becomes larger than the zero-point energy as the pressure increases, so that the molecules become "frozen" into a particular configuration at $T = 0$ K and $P \gtrsim 0.3$ Mbar The preferred lattice configuration appears to be the tetragonal γ-nitrogen structure rather than the essentially cubic α-nitrogen structure At megabar pressures, the energy required to rotate a molecule is equivalent to a temperature of order 2000 K.

The excited states of molecular hydrogen are even less well understood than the ground state The characteristic temperature for intramolecular vibration appears to be only weakly dependent on density and may actually decrease at the highest pressures (Silver

and Stevens 1973). Electronic excitation and molecular dissociation at the highest pressures are not understood quantitatively, but are expected to be important The thermodynamic uncertainties are discussed in Stevenson and Salpeter (1976)

Recent fluid-state calculations have been made by Ross (1974) and Stevenson and Salpeter (1976), assuming a sphericalized potential As usual, the solid and fluid equations of state at high pressure are very similar, provided the same potential is used for each. These fluid-state calculations suggest a melting temperature T_M, according to the criterion that the packing fraction in the equivalent hard sphere liquid not exceed 45% (Wainwright and Alder 1958). For $\rho \geqslant 0.4$ g cm^{-3}, Stevenson (1976*a*) finds

$$T_M \approx 2800\, \rho^2 \text{ K}, \qquad (2)$$

and Ross (1974) has obtained similar results. This result is uncertain by perhaps 50%, because of the uncertainty in the effective potential, and also assumes that the potential can be approximated by a spherical average This may be valid for the fluid phase, but if the solid has an ordered configuration of molecular orientations, then the hard sphere criterion may be invalid. However, similar values for T_M are suggested by the Lindemann criterion (Neece, Rogers, and Hoover 1971)

In summary, the thermodynamics of molecular hydrogen at $P \geqslant 0.1$ Mbar are not well understood, and the best constraint on the equation of state is the experimental shock data. The melting temperature is known to about a factor of 2, but is nevertheless almost certainly too low for the solid phase to exist in the present giant planets (see Paper II). Unlike metallic hydrogen, the molecular phase is increasingly classical as the pressure increases (Krumhansl and Wu 1968) Despite the uncertainties, we shall find that useful quantitative calculations can be made

iii) *The Molecular-metallic Transition*

There has not yet been a convincing experimental verification of this transition, although two claims (Grigoryev *et al.* 1972; Vereschchagin, Yakovlev, and Timofeev 1975*a*) have been made. The transition pressure is therefore estimated by theoretical calculations for the energies of the two phases and the usual common tangent construction The most recent and most accurate calculations for $T = 0$ K (Ross 1974) predict a transition pressure of between 2 and 4 Megabars. The factor of 2 uncertainty reflects the uncertainty in the molecular equation of state It has been suggested that there is a comparable uncertainty arising from the possibly incorrect usage of the free electron correlation energy in the metallic-state calculation (Monkhurst and Oddershede 1973; Ross and McMahan 1976). Since the correlation energy is very weakly density-dependent, this would represent an uncertainty in the energy scale and *not* in the equation of state (Computation of the correlation energy in the molecular state from first principles would be even more difficult. This problem does not arise in most calculations at present, which rely on the *experimental* properties of molecular hydrogen) In conclusion, it seems almost certain that the transition pressure exceeds 1 Mbar An upper limit cannot be established with the same certainty, but is probably about 5 Mbar For the "most likely" transition pressure of ~ 3 Mbar, the densities at transition are roughly 0.9 g cm^{-3} for the molecular phase and 1 1 g cm^{-3} for the metallic phase

It is likely that the transition is first-order at zero temperature because of the apparent dissimilarity of the two phases (for example, the large predicted density change at the transition) The nature of the transition is directly related to the sign of the microscopic "surface energy" between the phases In a simple model to be described below, this sign is found to be positive

As the temperature increases, entropy considerations ensure some "mixing" of the phases, and some temperature must exist beyond which the transition ceases to be first-order It is possible that the upper limit of the first-order character is coincident with the melting curve, i e , there exists a triple point at which metallic solid, molecular solid, and a "mixed" fluid phase are in mutual equilibrium (cf Trubitsyn 1972). On the other hand, Landau and Zel'dovich (1943) favor at least one critical point in the fluid region, in which case distinct metallic fluid and molecular fluid phases could coexist The solid-fluid transition is a rather subtle one, from an energetic standpoint, with the main change being the absence of long-range order in the fluid phase. Indeed, the volume change upon melting for either phase is very small (less than 3%), whereas the volume change that accompanies the molecular-metallic transition is comparatively large (20–30%). In other words, the *electronic* structures of the fluid and the solid are very similar whether one considers the molecular or the metallic state, but the electronic structure for molecular hydrogen differs substantially from that for metallic hydrogen

Nevertheless, two calculations (Kerley 1972, Aviram *et al* 1976) suggest that the transition is continuous in the fluid state Neither calculation can be regarded as satisfactory, since neither treats the two extremes (pure molecular and pure metallic) with a comparable degree of sophistication Calculation of the phase diagram requires a very careful calculation of the Gibbs energy for an arbitrary mixture of the two phases We shall not attempt this, but the relevant energies in such a calculation may be indicated by the following model

We first note that it is *not* meaningful to think of the electrons as being "localized" in very dense molecular hydrogen With the exception of small regions centered on each proton (in which the electron density is highly nonuniform in *both* molecular and metallic phases), the electron density is quite uniform In the language of band theory, dense molecular hydrogen is insulating because it is divalent, with a nonvanishing indirect band gap In fact, this band gap is much less than the band *width* at megabar pressures (Friedli and Ashcroft 1976)

Our model rests on three hypotheses:

1 A hydrogen molecule exists as a bound, metastable state when surrounded by metallic hydrogen at $P = P_t$, the transition pressure. This hypothesis is crucial to the model, but difficult to verify.

2. The volume per electron in a mixture of the metallic and molecular phases is approximately independent of position, i.e , the electron density does not fluctuate greatly according to whether one is near a molecule or near an unbound proton. This is reasonable, since the Thomas-Fermi screening length is comparable to typical interproton distances.

3 The energy of a neutral entity (i e , a "molecule," or an unbound proton together with a screening cloud of one electronic charge) is a function only of the volume it occupies This is the Wigner-Seitz hypothesis, and is expected to be quite accurate

Figure 1 shows the $T = 0$ K internal energies of the two pure phases (Ross 1974). Consider the formation of a molecule in the metallic state at the transition pressure $P_t \approx 3$ Mbar According to hypothesis 2, this occurs with essentially no volume change. According to hypothesis 3, the cost in energy per proton is just the difference ΔE_1 shown in Figure 1. Similarly, ΔE_2 is the energy cost per proton for breaking up a molecule in the molecular phase. Since these energies are both positive, we have established from very simple considerations that the microscopic surface energy, between the two phases, is positive The transition will be first-order until a temperature T_c such that the entropy of mixing, roughly $k_B T_c \ln 2$ (where k_B is Boltzmann's constant), is comparable to ΔE_1 or ΔE_2. This predicts that T_c is a few thousand kelvins.

This model has been quantified (Stevenson 1976*a*) by expressing the Gibbs free energy per proton as a function $G(x, P)$ of pressure P and of the fraction x of the protons which are bound in molecules. The transition pressure, critical temperature, and critical concentration are found from simultaneous solution of the equations

$$\frac{\partial G}{\partial x} = \frac{\partial^2 G}{\partial x^2} = \frac{\partial^3 G}{\partial x^3} = 0 , \qquad (3)$$

where the derivatives are at constant pressure and temperature. The results are $P_t \approx 3$ Mbar, $T_c \approx 3500$ K, and $x_c \approx 0.4$

The significance of this model is not in the numerical results, but rather in the identification of the relevant energies. According to this model, the relevant energy characterizing the transition is an order of magnitude smaller than the dissociation energy of an isolated

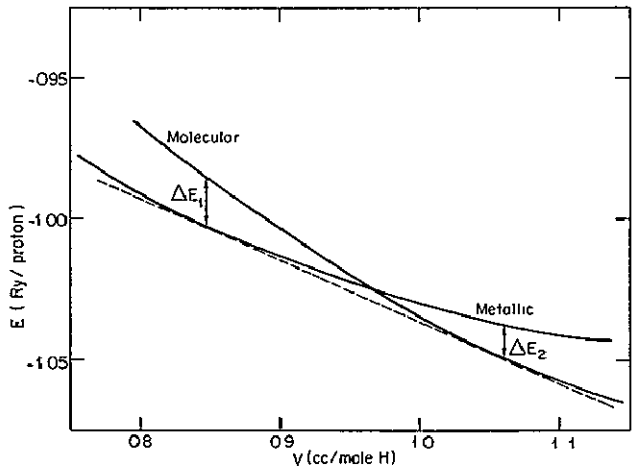


FIG 1.—Internal energy at $T = 0$ K for molecular and metallic phases Dashed line is a common tangent with slope $P = 3$ Mbar See text for discussion of ΔE_1, ΔE_2.

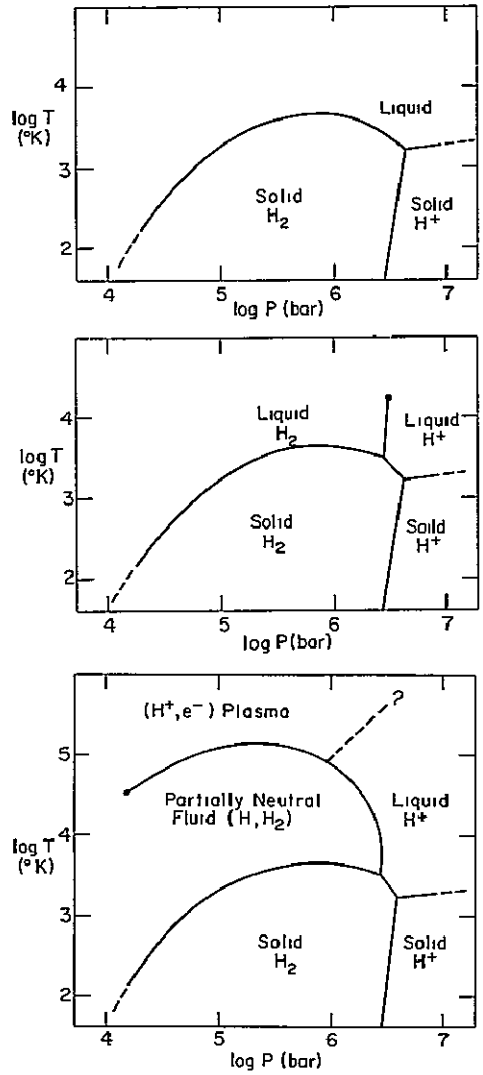


FIG 2.—Several possible phase diagrams of high-pressure hydrogen In (a) (*top*) no critical point exists In (*b*) (*middle*) there is a critical point so that two distinct liquid states coexist In (*c*) (*bottom*) the low-temperature phase diagram of (*b*) is joined in a natural way to the high-temperature phase diagram of Filinov and Norman (1975) The high-temperature dashed line represents the onset of degeneracy or even the possibility of another first-order transition (cf Landau and Zel'dovich 1943) In all these phase diagrams, the solid metallic phase is assumed to exist

hydrogen molecule The estimated critical temperature is comparable to the melting temperature of the molecular phase at $\rho \approx 1\ \mathrm{g\ cm^{-3}}$, but this is purely coincidental. Our model may, however, be misleading and our first hypothesis may not even hold An upper limit to T_c is of order 10^5 K, and any value in the range $10^3 \lesssim T_c \lesssim 10^5$ K cannot presently be discounted In Figure 2, three possible high-pressure phase diagrams of hydrogen are shown to illustrate the large uncertainty. The bottom phase diagram in Figure 2 is highly unconventional, but is a natural extension of a recent suggestion by Filinov and Norman (1975) that hydrogen undergoes a gas-liquid transition, analogous to that of cesium, in which the gas is almost fully ionized nondegenerate atomic hydrogen, and the "liquid" is partially ionized atomic hydrogen This last phase diagram is also in the spirit of the Landau-Zel'dovich (1943) hypothesis.

To conclude, there is a quite high probability that the molecular-metallic transition is first-order in part of the fluid phase The transition is possibly first-order even at 10,000 K, the relevant temperature for the present interior of Jupiter (see Paper II).

b) Helium

Helium is the most difficult element to ionize and the most difficult substance to metallize. Estimates of the insulator-metal transition pressure range from 20 Mbar to 100 Mbar (Simcox and March 1962; Trubitsyn 1967; Brust 1972; Ross 1972; Østgaard 1974, Stevenson 1976*a*), but the most reliable of these estimates are near the upper limit Since this transition is so far removed from the hydrogen transition, we will effectively ignore it, but it may be important in cold stars of low mass

There are two approaches to the thermodynamics of helium At low pressures, an interatomic pair potential compatible with experiment can be used (Trubitsyn 1967). At sufficiently high pressures ($P \gtrsim$ 10 Mbars), a first-principles approach analogous to metallic hydrogen can be used. This approach is accurate provided the band gap (between valence and conduction bands) is less than the valence band width, and does *not* require that the helium actually be metallic. The overlap between the two procedures is substantial and readily leads to a smooth interpolation between the low-pressure and high-pressure limits (Trubitsyn 1967). The considerations in the next three sections are not sensitive to the slight mismatch of the two approaches.

The melting temperature can be estimated from the criterion for freezing of a hard sphere fluid or from Lindemann's rule. At low pressures, the hard sphere criterion predicts $T_M \approx 1700$ K at $P = 1$ Mbar and $T_M \approx 4500$ K at $P = 4$ Mbar (Stevenson 1976*a*). At high pressures, the melting temperature increases less rapidly with

$$T_M \approx 4700\rho^{1/3}\ \mathrm{K} \qquad (4)$$

for ρ in g cm^{-3} (Trubitsyn 1967, Stevenson and Ashcroft 1974) For example, $T_M \approx 10{,}000$ K at $P = 50$ Mbar Like hydrogen, helium also melts at $T = 0$ K for a sufficiently high density (Stevenson and Ashcroft 1974), but this is of no interest for the giant planets

III METALLIC HYDROGEN-HELIUM MIXTURES

We first consider fluid mixtures The existence of miscibility gaps in many liquid metal mixtures is well known experimentally, but is difficult to predict theoretically since it depends on subtle free energy differences between the mixed and separated states Nevertheless, it has recently become possible to predict phase diagrams to roughly 10% accuracy, at least for simple metals where the interactions are well known (Stroud 1973) These calculations are based on a nearly free electron theory of metals, and a hard sphere perturbation theory for the structural properties of the liquid.

Metallic hydrogen-helium mixtures differ from alloys currently accessible in the laboratory, in that there are no "core" electrons to contend with, so the accuracy of a calculation is limited only by our knowledge of the dielectric response of the electron gas and the structural properties of the liquid On the other hand, the "bare" protons and α-particles are rather severe perturbations on the electron gas, so it is desirable to evaluate the electronic response to higher order than the usual low-order (linear response) approximation. A recent calculation (Stevenson 1975) evaluates the Gibbs energy to third-order in the electron-ion interaction, and uses a perturbation theory of fluids This calculation predicts a miscibility gap, the pressure dependence of which is shown in Figure 3. Below the critical line, a mixture containing roughly 40% helium by number will phase-separate into helium-rich and hydrogen-rich phases Below the dashed line, any mixture with a composition between 10% and 70% helium will similarly phase-separate.

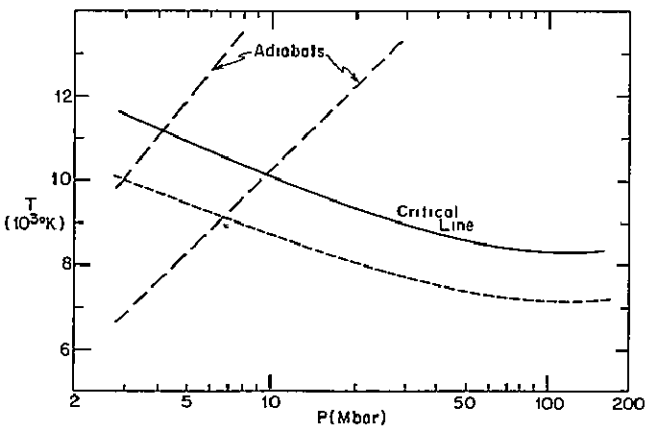


FIG 3.—Critical line for immiscibility in a metallic H-He mixture Also shown (– – –) is the temperature below which a solar composition mixture (10% He by number) would phase separate, and two typical adiabats (–·– ·–) appropriate to Jupiter or Saturn.

Calculations to second-order in the electron-ion interaction (Hansen and Vieillefosse 1976; Firey and Ashcroft 1976) confirm the general features of the phase diagram, but predict somewhat lower critical temperatures The existence of a miscibility gap can be explained merely by consideration of the Madelung energy (the electrostatic energy of the point ions immersed in a *uniform* electron gas), although correct allowance for the nonuniformity of the electron gas appears to increase the gap. The Madelung energy E_M can be adequately approximated by assuming ion-sphere charge averaging (Salpeter 1954), according to which E_M *at constant electron density* is a linear function of ionic concentration However, the comparison of alloy and separated phases must be made *at constant pressure*, and Stevenson (1976*b*) shows that under this constraint, there is a nonlinear dependence of E_M on ionic concentration such that the alloy is unfavorable relative to the separated phases. The crucial point is that at the densities and pressures of interest, the pressure is not just the Fermi contribution (independent of composition), but also has a substantial (negative) contribution from E_M. At much higher pressures (for which the electron gas is relativistic) the miscibility gap may no longer exist, since constant pressure and constant electron density become equivalent (Dyson 1971; Witten 1974). In Figure 3, Madelung energy considerations dominate for $P \gtrsim 10^2$ Mbar, whereas the rise in the critical temperature at lower pressures is explained by higher-order effects (the nonuniformity of the electron gas).

Pollock and Alder (1977) agree with the above conclusions in the high-pressure limit ($P \gtrsim 10^2$ Mbar), but conclude that at the lower pressures relevant to Jupiter, helium may be highly soluble (perhaps soluble in all proportions) However, this conclusion is based on very crude models for the low-density interactions, and it is possible to construct physically realistic models which predict that the helium solubility is *least* at zero pressure and increases monotonically with pressure for $0 \lesssim P \lesssim 10^2$ Mbar. More needs to be known about the electronic structure of helium dissolved in low-density metallic hydrogen before firm conclusions can be reached for the solubility at the lowest pressures We shall adopt the working hypothesis that helium is least soluble in metallic hydrogen at the lowest pressure of interest (i.e , at the molecular-to-metallic hydrogen transition), and that phase separation begins for $T \lesssim 10{,}000$ K at this pressure

Solid hydrogen-helium alloys have been considered by Straus, Ashcroft, and Beck (1977). Their calculations indicate an even larger miscibility gap in the solid state than in the fluid state This suggests that the liquidus for the alloy is lower than at least one of the melting temperatures for the pure phases, at all compositions. This effect of alloying on the melting temperature was suggested by Smoluchowski (1971) on the basis of known trends in metallic alloys It follows that the metallic core of the giant planets is fluid (see Paper II).

IV MOLECULAR HYDROGEN-HELIUM MIXTURES

In contrast to the metallic state, the molecular state is not readily amenable to first-principles calculations, and we are forced to resort to semiempirical pair potentials that are compatible with experimental data, yet are also plausible modifications of first-principles calculations. Experiments have been conducted on molecular H_2-He mixtures for pressures up to 7 kilobars, and a miscibility gap has been observed (Streett 1973) The calculation about to be described for megabar pressures can only be suggestive, and is not as quantitatively reliable as the metallic calculation reviewed in the previous section

The Helmholtz free energy F was calculated by Stevenson (1976*a*) as a function of density, temperature, the fraction x (the number of molecules) of He in the fluid H_2-He mixture. Two different calculations were carried out, one using a simple exponential 6–8 form for all the interaction potentials, with the coefficients for the H_2-H_2, H_2-He, and He-He interactions taken from Ross (1974), Shafer and Gordon (1973), and Trubitsyn (1967), respectively This calculation was carried out for all pressures from 1 kbar up to 5 Mbar. The second calculation used Lennard-Jones 6–12 potentials and was carried out only at low pressures. From F, the Gibbs free energy $G(P, T, x)$ was then obtained For each pressure P, the requirement $\partial^2 G/\partial x^2 = \partial^3 G/\partial x^3 = 0$ gives the critical temperature T_c and the critical helium mole fraction x_c The calculated results for $T_c(P)$ are given in Figure 4 and agree fairly well with Streett's experimental results, especially with regard to slope. The calculated ratio $k_B T_c/G_e(P)$, where G_e is the nonideal gas part of the Gibbs free energy of the critical mixture, varies by only 50% as the pressure changes by two orders of magnitude The slopes of the curves for $G_e(P)$ and $T_c(P)$ are probably fairly reliable, and, in view of the agreement with the experimental data at low pressures,

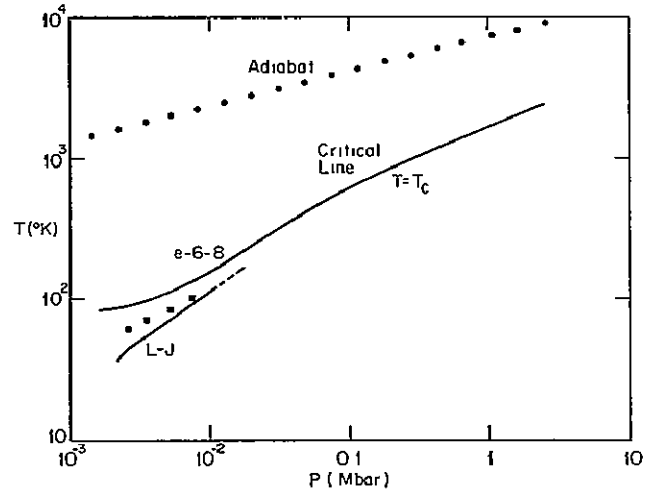


FIG 4.—Critical line for immiscibility in a fluid H_2-He mixture, for exp 6–8 and L-J potentials Also shown are Streett's experimental critical values (■) and a typical Jovian adiabat (●●●)

the critical curve in Figure 4 is better than an order-of-magnitude estimate and perhaps within a factor of 2 of the correct value. The calculated value for the critical helium mole fraction was $x_c \approx 0.55$ at pressures appropriate to Streett's experiment, close to the experimental value of $x_c \approx 0.58$. The calculated value changed little with pressure, decreasing to $x_c \approx 0.50 \pm 0.05$ at $P = 3$ Mbar

To summarize If the intermolecular potentials can all be written in the simple form chosen, then Streett's experimental results have implications for the phase diagram at megabar pressures. It seems likely that at $P \approx 3$ Mbar, $2000\,\mathrm{K} \leqslant T_c \leqslant 6000\,\mathrm{K}$ This is at least a factor of 2 smaller than the critical temperature of the *metallic* mixture at $P \approx 3$ Mbar.

A notable feature of both Streett's experimental results and the above fluid-state calculations is that T_c is very similar to the melting point of either pure phase. The eutectic temperature may be substantially lower, but there is nevertheless uncertainty as to whether fluid-state calculations are relevant. No solid-state calculation has been attempted for the mixture, and all subsequent considerations are confined to the fluid state. This is justified in our discussions in Paper II, since only the evolution prior to immiscibility in the *molecular* phase is considered in detail.

V. THE TOTAL PHASE DIAGRAM

The previous three sections have dealt with three aspects of the hydrogen-helium phase diagram as though they were distinct and unrelated We now unify these into a single, coherent topology for the three-dimensional phase diagram (the dimensions being pressure P, temperature T, and composition x) according to the following model

We consider an arbitrary hydrogen-helium mixture as a constrained ternary system of N protons and helium atoms, in which xN particles are helium atoms, $(1-x)yN$ are unbound protons, and $(1-x)(1-y)N$ are protons bound together as H_2 molecules. The Gibbs energy of the system is approximated as

$$G(P,T) = N\left[\sum_i x_i G_i^{(1)}(T) + \frac{1}{2}\sum_{i,j} x_i P_{ij} G_{ij}^{(2)}(P)\right], \tag{5}$$

where i ranges from 1 to 3, and x_i is the number fraction for each of the three species ($i = 1$ is He, $i = 2$ is H^+, $i = 3$ is bound protons) P_{ij} is the probability that a particle of species i will have a particle of species j as one of its nearest neighbors The $G_i^{(1)}$ incorporate the ideal entropy of mixing and any chemical potential relative to an arbitrarily chosen energy zero. In other words,

$$\begin{aligned} G_1^{(1)} &= k_B T \ln (x/s)\,, \\ G_2^{(1)} &= k_B T \ln [(1-x)y/s] + \tfrac{1}{2}D\,, \\ G_3^{(1)} &= \tfrac{1}{2}k_B T \ln [(1-x)(1-y)/2s]\,, \\ s &\equiv x + (1-x)y + \tfrac{1}{2}(1-x)(1-y)\,, \end{aligned} \tag{6}$$

where D is the dissociation energy of the hydrogen molecule Entropy effects (other than the ideal entropy of mixing) are omitted in these expressions, since thermal contributions are minor perturbations in *cold* systems (these entropy perturbations can be readily reintroduced for evaluating thermal derivatives along phase boundaries) The diagonal elements of $G_{ij}^{(2)}$ are *known* since they correspond to the three pure phases (see § II) The three distinct off-diagonal elements are found by *assuming* numerical values for the three distinct critical temperatures T_c(H-He), T_c(H_2-He), and T_c(H-He). For example, T_c(H-He) is the solution of $\partial^2 G/\partial x^2 = \partial^3 G/\partial x^3 = 0$ for $y \equiv 1$ A random mixture was assumed, so that $P_{ij} = x_j/s$ This simple choice automatically implies the following simple compositions for the critical mixtures: $x_c = 1/2$ for H-He, $x_c = 1/3$ for H_2-He (half H_2, half He), and $y = 1/3$ for H-H_2 (half H_2, half H)—all crude but adequate approximations. The total Gibbs energy for a given x, P, and T is then minimized with respect to y to yield the equilibrium state of the hydrogen At sufficiently low temperatures there are two minima—one corresponding to "metallic" hydrogen, the other corresponding to "molecular" hydrogen Except in special cases, one minimum will be lower than the other and correspond to the equilibrium state. The higher minimum corresponds to the metastable state. If the temperature is too high, or the helium content is too great, then the first-order character of the molecular-metallic transition is "washed out," and there is only one minimum

For each (P, T) the existence of one or more common tangents to the equilibrium Gibbs energy as a function of x determines the coexisting phases and the thermodynamically inaccessible regions. In this way, the phase diagram was mapped out for all P, T, x of interest

We shall describe in detail the results for the choice

$$\begin{aligned} T_c(\text{H-He}) &= 12{,}000\,\mathrm{K}\,, \\ T_c(\text{H}_2\text{-He}) &= 6{,}000\,\mathrm{K}\,, \\ T_c(\text{H-H}_2) &= 18{,}000\,\mathrm{K}\,, \end{aligned} \tag{7}$$

which, according to the discussion of the previous sections, is a possible selection. (For simplicity, the pressure dependence of each T_c is ignored) Figure 5 illustrates the results. Consider, first, diagram (*a*), for which $T = 13{,}000$ K. At each pressure in the range 3–4 6 megabars there coexist a helium-poor metallic phase and a helium-rich molecular phase whenever the total helium content lies within the shaded region. Below the dashed line, the metallic phase is more dense than the molecular phase, whereas the reverse is true above the dashed line. This "density inversion" is a consequence of the competition between the density increase accompanying the addition of helium, and the density decrease accompanying the metallic-molecular transition At sufficiently large helium concentration x, the first-order character of the metallic-molecular transition is lost and there are no excluded regions.

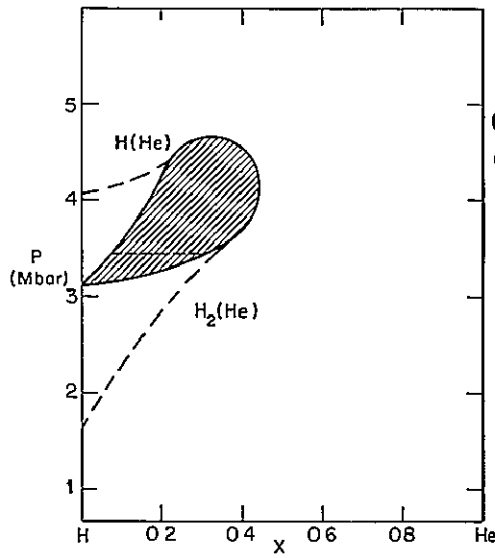


FIG 5a

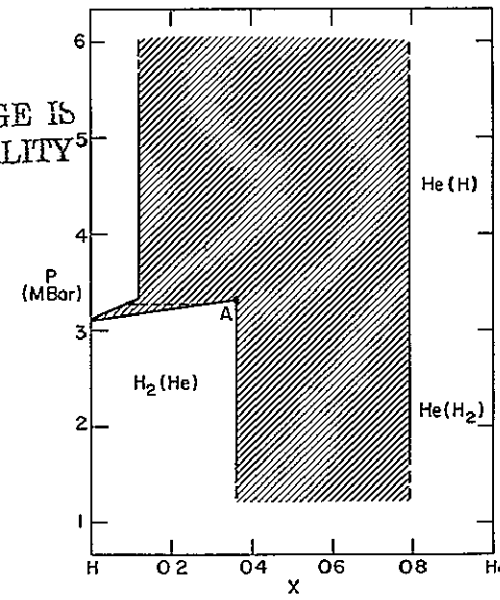


FIG 5c

FIG 5.—Phase diagrams for three different temperatures: (*a*) $T = 13{,}000$ K, (*b*) $T = 7500$ K, (*c*) $T = 4000$ K. In each case, the phase-excluded region is shaded Above the dashed line (---), the phase on the right-hand side of the phase-excluded region has greater mass density than the coexisting phase on the left-hand side Below the lower dot-dashed curve (-·-) the metallic phase ceases to be metastable Above the upper dot-dashed curve the molecular phase ceases to be metastable. Note the presence of a triple point *A* in diagram (*c*).

Consider diagram (*b*) of Figure 5. Since $T = 7500$ K $< T_c$(H-He), there is now a miscibility gap which extends to high pressures. This evolves smoothly from the "loop" of diagram (*a*) Notice that there is no clear distinction between the molecular-metallic transition and the phase separation in the metallic fluid Proceeding smoothly along the lower phase boundary from small x to large x, the fluid progresses smoothly from predominantly molecular to predominantly metallic.

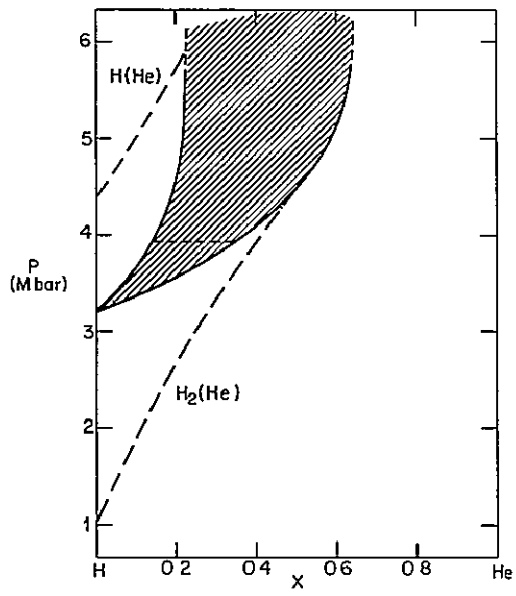


FIG 5b

In diagram (*c*), $T = 4000$ K and there is now a miscibility gap in the molecular fluid. This miscibility gap forms smoothly from diagram (*b*), as T is lowered, in the following way: At some critical temperature, T_c^*, an inflection becomes formed in the lower phase boundary of diagram (*b*). In this model, T_c^* is comparable to T_c(H$_2$-He) For $T < T_c^*$ a minimum in P (as a function of x along the phase boundary) is formed, and the miscibility gap rapidly grows as T is further reduced. Immediately below T_c^* a triple point [marked *A* in diagram (*c*)] is formed Thus there is a line of triple points ending at a critical point $T = T_c^*$ (at $P \approx 3.5$ Mbar) The concentration at the triple point is a sensitive function of temperature, and becomes smaller as the temperature is reduced and the excluded region expands to fill most of (P, x)-space. At low temperatures, the "density inversion" effect eventually vanishes and the immiscibility effects dominate.

For general values of the parameters in equation (7) one can define a "configuration space" in which each point is itself a phase diagram This is shown in Figure 6 for the choice T_c(H-He) $= 2T_c$(H$_2$-He) For given values of T_c(H-He), T_c(H-H$_2$), and T one can find from this "configuration" diagram what the topology of the physical phase diagram is

In the following paper (Paper II) these model phase diagrams will be used in considering specific compositional and thermal histories of an evolving hydrogen-helium planet such as Jupiter

VI. MINOR CONSTITUENTS

It is clear both from atmospheric observations and interior models that the hydrogen-helium planets

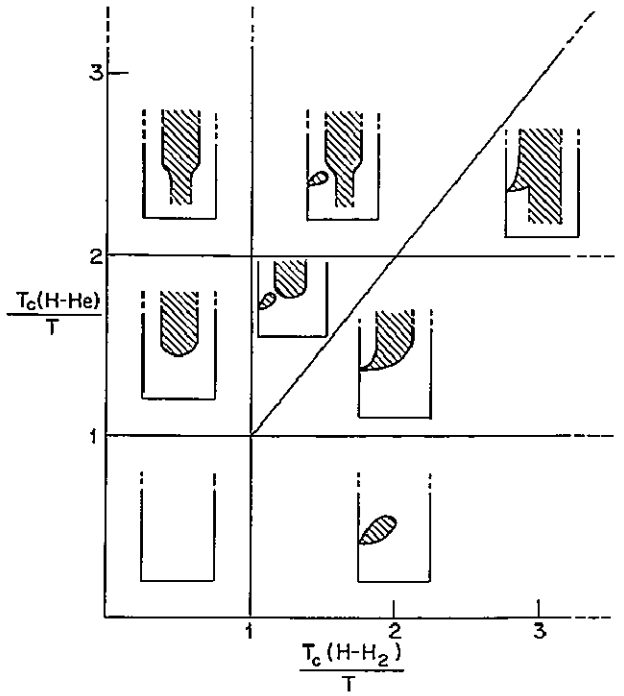


FIG 6—Various possible phase diagrams, assuming $T_c(\text{H-He}) = 2\,T_c(\text{H}_2\text{-He})$ Each small diagram within the figure is a schematic representation of a (P, x)-diagram similar to that in Fig 5.

contain minor constituents at least to the extent of solar abundance. The distribution of these minor constituents is important both for model construction and for relating the observed atmospheric abundance to the total abundance. There is the possibility that an appropriately chosen minor constituent or group of constituents could be very precise "tracers" of internal dynamic processes by virtue of their almost complete partitioning into one of the hydrogen-helium phases. No especially appropriate tracer is indicated by the analysis of this section, which deals primarily with general trends. The special case of deuterium is discussed separately. This section deals only with thermodynamic considerations The actual distribution of constituents within an evolving planet also depends on fluid-dynamic and diffusive processes (Paper II)

a) Deuterium

Both CH_3D (Beer *et al* 1972) and HD (Trauger *et al.* 1973) have been observed in the Jovian atmosphere, and the inferred deuterium abundance has been frequently quoted as indicative of the primordial solar (or even cosmic) abundance. The partitioning of deuterium therefore has an importance out of proportion to its abundance Unlike other minor constituents, the chemical potential of deuterium is readily calculable (as a simple extension of the analysis of ordinary hydrogen)

Consider, first, the partitioning of a small amount of deuterium between pure, coexisting molecular and metallic phases of ordinary hydrogen. Hubbard (1974) concluded that the mass fraction of deuterium in the metallic phase exceeds that in the molecular phase by roughly 15%. His calculation is for the "classical" (i e., high-temperature) limit but neglects the vibrational degrees of freedom for the H_2 and HD molecules, and also neglects dissociation If, instead, one assumes that the vibrational degrees of freedom are fully excited and harmonic, then the chemical constant of HD is increased by $\frac{1}{2}$ ln $(\frac{4}{3})$ relative to H_2, and the mass fraction of deuterium in each phase is exactly the same (This is a general result for the classical limit and not a special property of hydrogen) Excitation of the vibrational modes probably is achieved at 10^4 K, the temperature of interest, since the low-density vibrational temperature for H_2 is 6000 K, and this does not appear to increase at high density (Silver and Stevens 1973). As the temperature is reduced, another effect not considered by Hubbard becomes important: quantum corrections to the translational energy of the protons and deuterons in the metallic state This can be calculated from the Wigner theory as in Stevenson (1975). This positive contribution to the chemical potential is larger for protons than for deuterons and therefore favors partitioning of deuterons into the metallic phase (The competing quantum effect in the molecular phase is negligible) The incomplete excitation of the vibrational modes of H_2 and HD also favors partitioning into the metallic phase. Numerical calculations indicate that the mass ratio of deuterium (metallic) to deuterium (molecular) is essentially unity for $T \gtrsim$ 8000 K, about 1 05 at $T \approx 5000$ K, and 1 25 at $T \approx 2500$ K.

Consider now the partitioning of deuterium between hydrogen-rich and helium-rich metallic phases. In the relevant high-temperature limit, the only free energy contribution tending to produce a partitioning of deuterons different from the partitioning of protons is the quantum translational energy. According to the Wigner theory, the shift in equilibrium is such as to favor less variation of the ionic thermal de Broglie wavenumber. The deuterium-to-hydrogen ratio is thus greater in the helium-rich phase. Numerical calculation, based on the evaluation of F_Q in Stevenson (1975), indicates that this ratio is 10% larger in the helium-rich phase than in the hydrogen-rich phase at $T = 5000$ K, with the difference vanishing at $T \approx 10{,}000$ K.

The deuterium-to-hydrogen ratios in coexisting hydrogen-rich and helium-rich molecular phases should coincide at the temperatures of interest, provided the rotational and vibrational degrees of freedom of the H_2 and HD molecules are not strongly influenced by the fraction of helium in the local environment. In the absence of a detailed model for these modes, no quantitative calculation can be made Substantially unequal partitioning seems unlikely, however

In conclusion, the partitioning of deuterium between the various hydrogen-helium phases appears to preserve the deuterium-to-hydrogen mass ratio, at least for $T \gtrsim 5000$ K. The deuterium content in the

uppermost convective layers of hydrogen-helium planets should therefore be representative of the bulk composition, provided the reservoir of material from which the planet formed had a uniform distribution of deuterium

b) Other Minor Constituents

First, consider the possibility of a phase transition caused by a minor constituent (e g , insolubility of a minor constituent) This could occur independently of the existence of phase boundaries in the hydrogen-helium, but it is improbable for the low concentrations and high temperatures of interest. If the number fraction of a minor constituent is z, then an energy of about $-k_BT \ln z$, which favors the dissolved state, must be compensated by an effect which favors the separated phase For example, water at $T \lesssim 300$ K, pressures of order of a few bars, and abundance $z \approx 10^{-3}$ can preferentially form droplets since $-k_BT \ln z \lesssim 0.2$ eV can be overcome by the binding energy of the liquid water In the deep interior of the planet, however, $-k_BT \ln z \approx 6$ eV, and there is apparently no correspondingly large binding effect. Water is probably insoluble in molecular hydrogen at low enough temperatures or high enough concentrations, but this is probably not relevant to the deep interiors of present giant planets. We shall therefore restrict ourselves to a discussion of partitioning between phases of the hydrogen-helium system.

The degree of partitioning is determined by equating the chemical potentials for the impurity in the two coexisting phases. At high pressures, the chemical potential can be meaningfully separated into four parts. (i) the "nonchemical" electronic contribution (i.e , a part which does not explicitly invoke the symmetry properties or discreet band structure of the electronic spectrum), (ii) residual chemical effects [i e , electronic effects not included in (i)], (iii) configurational (including entropy) effects, resulting from the different size of solute and solvent atoms; and (iv) the ideal free energy of mixing

Consider first the "nonchemical" electronic contribution. In the high-pressure limit, where the electrons can be considered to be a uniform Fermi gas, Stevenson (1976*b*) showed that the miscibility gap in a binary alloy increases as the difference between the nuclear charges of the constituents increases A direct corollary of this result is that ions will partition so as to minimize nuclear charge differences Thus all elements with $Z \geq 3$ will preferentially partition into the helium-rich phase of a hydrogen-helium mixture A more general result, applicable to lower pressures, can be obtained by an extension of the Thomas-Fermi-Dirac (TFD) method The usual TFD procedure for an alloy is to assume volume additivity, whereby the locally evaluated "pressure" at the Wigner-Seitz cell boundary is assumed to be the same for every cell If electron correlation is ignored, or evaluated in a local approximation, then this also implies continuity of the electron density across cell boundaries (Salpeter and Zapolsky 1967). Clearly, this procedure predicts that the chemical potential of a constituent is independent of its environment (at a given pressure) so that no nonuniform partitioning could occur The failure of the TFD method is not so much in the prescription for determining the charge density (which is very accurate at sufficiently high pressure) but in the unphysical procedures for evaluating pressure and assigning boundary conditions We propose that a better, albeit more complicated, procedure is to enforce continuity of the electron density at the cell boundaries, and calculate pressure according to the rigorous (i e., nonlocal) thermodynamic derivative of the total energy with respect to volume Let $\rho \equiv \rho(P)$ be the actual electron density at the Wigner-Seitz cell boundary (approximated by a sphere) at pressure P Let $V_i(\rho)$ be the specific (cell) volume of species i, and $E_i(V)$ be the energy per cell (evaluated as though the substance were purely species i) In accord with the Wigner-Seitz philosophy, the total energy per atom is assumed to be

$$E = \sum_i x_i E_i\{V_i[\rho(P)]\}, \tag{8}$$

where x_i is the number fraction of species i [The energy is *not* a linear function of the x_i, since $\rho(P)$ is also a self-consistently determined function of the alloy composition.] It then follows that in the limit of vanishing concentration for species i, the chemical potential μ_i is

$$\begin{aligned} \mu_i &= \mu_i^0 + \Delta\mu_i , \\ \mu_i^0 &= E_i\{V_i[\rho_i(P)]\} + PV_i[\rho_i(P)] , \\ \Delta\mu_i &= -\frac{1}{2}\left(\frac{\partial P}{\partial V}\right)_i \left[(\rho_0 - \rho_i)\frac{\partial V_i}{\partial \rho}\bigg|_{\rho_i}\right]^2 , \end{aligned} \tag{9}$$

to lowest nonvanishing order in $(\rho_0 - \rho_i)$, where $\rho_i(P)$ is the cell boundary electron density for a pure substance composed of species i, and $\rho_0(P)$ is the cell-boundary electron density for the solvent phase (the relevant hydrogen-helium phase in this case). The TFD procedure (without correlation or with locally evaluated correlation) predicts $\rho_0 \equiv \rho_i$ and $\Delta\mu_i \equiv 0$. The above procedure does not require that the $E_i(V)$ be evaluated according to TFD and, in general, $\rho_0 \neq \rho_i$. The $\Delta\mu_i$ is always positive, and can be regarded as a microscopic "surface energy." The model predicts that a solute preferentially enters the phase in which the cell boundary electron density is most compatible For example, ρ(He) is more similar to $\rho(H_2)$ than ρ(metallic H), and helium therefore prefers the molecular phase, in accord with our discussion in § V.

Unfortunately, the pressure of interest is not high enough for simple generalities based only on nuclear charge For example, Na and Al, elements with similar nuclear charges, behave quite differently. Pseudopotential theory (with polarizable core states) suggests that the essentially monovalent Na has $\rho \approx 0.041 a_0^{-3}$ at $P = 3$ Mbar (a_0 is the first Bohr radius), whereas the trivalent Al has $\rho \approx 0.058 a_0^{-3}$. (For a discussion of

pseudopotential theory, see Ashcroft and Langreth 1967) The corresponding cell boundary densities for hydrogen are $0.06a_0^{-3}$ (metallic) and $0\,035$–$0\,04a_0^{-3}$ (molecular). The metallic value is estimated from Wigner-Seitz calculations (Neece, Rogers, and Hoover 1971) and the molecular value from band structure calculations (Friedli and Ashcroft 1976). If metallic hydrogen is the solvent, then (from eq. [9]), $\Delta\mu_{Na} \approx 2$ eV and $\Delta\mu_{Al} \approx 0$; whereas if molecular hydrogen is the solvent, then $\Delta\mu_{Na} \approx 0$ and $\Delta\mu_{Al} \approx 1\,5$ eV. If other factors were negligible then Al would prefer metallic hydrogen and helium-poor phases, whereas Na would prefer molecular hydrogen and helium-rich phases Further generalization is difficult, and the partitioning of Fe and Mg (for example) is not readily predicted One would expect, however, that atoms or molecules with closed shell configurations at low densities would, in most instances, still have low cell boundary electron densities even at megabar pressures, and prefer molecular or helium-rich phases This might include the abundant "closed shell" species H_2O, CH_4, and NH_3 (but see the discussion on H_2O at the end of this section)

Consider, now, the "chemical" effects that are not implicit in the previous analysis These are difficult to estimate, but appear to be small For example, it might be supposed that a metal would not dissolve in dense molecular hydrogen because the available conduction states in the hydrogen are separated from the valence band by an energy gap. However, the band gap is $\lesssim 1$ eV at the transition pressure (Friedli and Ashcroft 1976), so this effect may be less than that predicted by equation (9). Similarly, the categorization of polar and nonpolar molecules is meaningless at megabar pressures, and the distinctions among covalent, ionic, and metallic bonding become inapplicable.

The configurational contribution to the chemical potential can be estimated for the fluid phase by the hard sphere model (Lebowitz and Rowlinson 1964), with the effective (pressure- and temperature-dependent) hard sphere diameters determined by minimization of the total free energy. Numerical calculations indicate that this contribution is several k_BT at $T \approx 10^4$ K, but that the difference between solute potentials for the various solvent phases is less than $k_BT \approx 1$ eV and therefore usually small compared with electronic differences

The ideal free energy of mixing is $k_BT \ln z$, where z is the number fraction of the solute. Typically, the electronic chemical potential differences between two coexisting phases are a few eV, so that for $k_BT \approx 1$ eV the value of z could change by as much as an order of magnitude as one crosses a phase boundary

We conclude with a brief discussion of the partitioning of H_2O, probably the most abundant minor constituent in Jupiter and Saturn (although possibly underabundant in the Jovian atmosphere, according to Larson *et al.* 1975). According to the preceding analysis, we would expect H_2O to prefer molecular and helium-rich phases However, this assumes that the configuration—and the electronic structure—of H_2O is similar for each phase. Pure water is completely dissociated into H_3O^+ and OH^- at about 200 kilobars (Hamann and Linton 1966) and is metalized at several megabars (Ramsey 1963; Vereschchagin, Yakovlev, and Timofeev 1975*b*), at which pressure nothing is known about the configuration. The dissociation does not significantly modify the previous analysis, since H_3O^+ and OH^- are both isoelectronic with a closed shell atom (neon). However, one should consider the possibility that H_2O enters metallic hydrogen as $2H^+ + O^{n+} + (n+2)e^-$, where $n \geq 0$ Approximate numerical calculations suggest that this is highly improbable, even for $n = 1$, despite the similarity of the first ionization energy of oxygen ($\sim$13 6 eV) and the binding energy per electron of the metallic state. The problem is that the energy reduction gained by "metalizing" the oxygen atom is small, and does not compensate the rather large binding energy of the OH^- ion The chemical potential of H_2O in molecular hydrogen is $\sim$20 eV (relative to the isolated zero-pressure H_2O molecule), whereas the chemical potential for the hypothetical metalized state (with the oxygen in the O^+ form) has a chemical potential $\sim$28 eV at least

VII. TRANSPORT PROPERTIES OF THE METALLIC PHASE

We consider essentially all the "first-order" atomic transport coefficients in the following order: electrical conductivity, thermal conductivity, viscosity, self-diffusion, inter-diffusion, and radiative opacity. There is also a brief discussion of "second-order" (or off-diagonal) transport coefficients such as the Soret coefficient.

a) Electrical Conductivity

This has been evaluated by Stevenson and Ashcroft (1974) using the well-known Ziman theory, and the hard sphere static structure factors In that paper, the temperature scale was only estimated, but subsequent thermodynamic calculations (Stevenson 1975) established the correspondence between hard sphere diameter and temperature for each density. An estimate can also be made for the *dynamic* corrections, using the theory of Baym (1964) and the molecular-dynamics results of Hansen, McDonald, and Pollock (1975) for the one-component plasma The improved temperature scale and the dynamic corrections each modify the results of Stevenson and Ashcroft (1974) by as much as a factor of 2—but in opposite directions The final result is the following approximate formula for the conductivity σ:

$$\sigma \approx \frac{5 \times 10^{20} \rho^{4/3}}{T(1+3x)} \text{ esu}, \tag{10}$$

where ρ is the mass density in g cm^{-3}, and x is the helium number fraction This formula should be correct to within a factor of 2 for $1 \lesssim \rho \lesssim 10^2$ g cm^{-3} and $10^3 \lesssim T \lesssim 10^6$ K, but should only be used for $x \lesssim 0.2$. In the conditions prevailing in the Jovian

core at present, $\sigma \approx 10^{17}$ esu, comparable to that of room-temperature alkali metals. The value of σ given by equation (10) is about a factor of 2 larger than the estimates for *solid* metallic hydrogen by Abrikosov (1964) and Hubbard and Lampe (1969).

b) Thermal Conductivity

In the metallic phase, thermal conductivity is dominated by electronic transport If the electrons are degenerate, and if the Born approximation is valid (see Stevenson and Ashcroft 1974 for a discussion of this point), then the thermal conductivity is related to the electrical conductivity by the Wiedemann-Franz relation The thermometric conductivity κ is then given by

$$\rho C_p \kappa \approx \frac{1.5 \times 10^6 \rho^{4/3}}{(1 + 3x)} \text{ ergs cm}^{-1} \text{ s}^{-1} \text{ K}^{-1}, \quad (11)$$

or, if we assume $C_p \approx 3Nk_B$, where N is the number of ions per gram,

$$\kappa \approx 0.3 \rho^{1/3} \text{ cm}^2 \text{ s}^{-1}. \quad (12)$$

Notice that the temperature T does not appear in equations (11) and (12). The accuracy and validity of these equations is the same as for the electrical conductivity.

c) Viscosity

Unlike the electronic transport properties above, viscosity and atomic diffusion depend explicitly on the dynamic properties of the fluid. There is no generally accepted and successful theory for the dynamics of a dense fluid However, models which work for the conventional alkali metals, such as the Longuet-Higgins and Pople (1956) model, as adapted by Ascarelli and Paskin (1968) and modified by Vadovic and Colver (1971), probably are also satisfactory for metallic hydrogen. The following approximate formula is then deduced:

$$\nu \approx 4 \times 10^{-3} T_4^{-1/2} \text{ cm}^2 \text{ s}^{-1}, \quad (13)$$

for any hydrogen-helium mixture, where T_4 is the temperature in units of 10^4 K The apparent lack of density dependence in this result is only approximate At the temperatures and densities of interest, this result should be correct to at least a factor of 5 (and probably a factor of 2).

This calculation is based on a hard sphere approach. The opposite extreme is the one-component plasma, which can be regarded as the unscreened metallic state. Two calculations for this system (Hansen, McDonald, and Pollock 1975; Vieillefosse and Hansen 1975) agree that

$$\nu \approx 0.1 \omega_p \bar{r}^2 \quad (14)$$

to within a factor of 2, where ω_p is the ion plasma frequency and $\bar{r}$ is the radius of that sphere which contains one ion on the average. This formula yields a value that is typically a factor of 2 smaller than equation (13), at least for $T_4 \approx 1$, and it also predicts a very weak density dependence ($\nu \propto \rho^{-1/6}$).

From equations (12) and (13), we can now estimate the Prandtl number Pr:

$$Pr \equiv \frac{\nu}{\kappa} \approx 10^{-2} T_4^{-1/2} \rho^{-1/3}, \quad (15)$$

provided the helium content satisfies $x \lesssim 0.2$. (Helium-rich fluids may have a substantially lower κ.) Thus, for $T_4 \approx 1$ and $\rho \approx 1$ g cm^{-3}, $Pr \approx 10^{-2}$, which is typical of liquid alkali metals.

d) Self-Diffusion

This transport property may not be of great interest itself, but it provides a means of estimating the more interesting interdiffusion (diffusion of helium in hydrogen) We use the same theory as for the viscosity (Vadovic and Colver 1971), which predicts that the product of self-diffusion D and viscosity ν is given by

$$D\nu \approx 0.17 \sigma^2 \left(\frac{k_B T}{M}\right), \quad (16)$$

where σ is the effective hard sphere diameter, and M the ion mass This result is experimentally verified when σ is chosen by thermodynamic considerations alone. Thus,

$$D \approx 3 \times 10^{-3} \rho^{-2/3} T_4^{3/2} \text{ cm}^2 \text{ s}^{-1}, \quad (17)$$

for both pure hydrogen and pure helium

The one-component plasma studies (Hansen, McDonald, and Pollock 1975, Vieillefosse and Hansen 1975) predict $D \propto \rho^{-2/3} T^{4/3}$ and a magnitude that is typically a factor of 3 smaller than that given by equation (17) This agreement is satisfactory, and suggests that this transport property is not strongly dependent on the details of the ion-ion interaction.

e) Interdiffusion

There is no similarly successful model for interdiffusion, so we shall resort to empirical evidence. Experiments on liquid metal mixtures (Ejima and Yamamura 1973) indicate that the interdiffusion of one atomic species in another differs from the self-diffusion of the most abundant species to the extent that the species differ in "size." Thermodynamic calculations (Stevenson 1975) indicate that the helium pseudoatom (α-particle plus screening cloud of electrons) is 30% larger than the hydrogen pseudoatom The experiments then indicate that a small amount of helium in hydrogen should diffuse about half as rapidly as the self-diffusion of hydrogen Thus

$$D_{\text{H-He}} \approx 1.5 \times 10^{-3} \rho^{-2/3} T_4^{3/2} \text{ cm}^2 \text{ s}^{-1}, \quad (18)$$

and independent of composition to a first approximation.

To see whether diffusion is anomalous near a phase transition, we first express the interdiffusion coefficient D in a more fundamental form (Landau and Lifshitz 1959):

$$D = \frac{\alpha}{\rho}\left(\frac{\partial \mu}{\partial x}\right)_{T,P}, \tag{19}$$

where μ is the helium chemical potential, x is the helium concentration, and α is a "canonical" kinetic coefficient, as explained by Landau and Lifshitz The requirement that entropy increase with time implies that $\alpha > 0$ Consider, now, the specific Gibbs energy in Figure 7*a* (This is a schematic representation of Fig. 2 in Stevenson 1975.) Between A and D, a fluid mixture is energetically unfavorable relative to separated helium-rich and hydrogen-rich phases Between A and B and between C and D the fluid mixtures are *metastable* (i e., $\partial^2 \Delta G/\partial x^2 = \partial \mu/\partial x > 0$). In these regions, phase separation must proceed by nucleation and can be strongly inhibited by the surface energy between the phases. Between B and C, the fluid mixture is unstable to *spinodal decomposition* (the onset of long-wavelength concentration fluctuations), the rate of which is essentially limited only by diffusion rather than by surface energy. In this region, $\partial \mu/\partial x < 0$, and the diffusion coefficient can be regarded as negative in the sense that compositional inhomogeneities tend to grow rather than decay with time. At the points B and C, the diffusion constant is zero In Figure 7*b* the phase diagram for a given pressure is shown and the various regions indicated. Spinodal decomposition has recently been clearly simulated for the first time in computer experiments (Abraham *et al* 1976) and has been the subject of several theoretical investigations (Abraham 1975*a*, *b*)

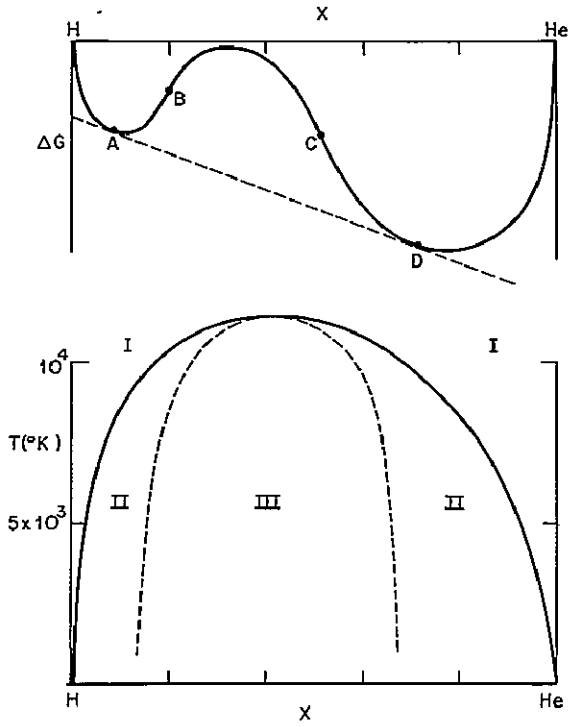


FIG 7 (*a*) (*top*)—Gibbs energy of mixing for a H-He mixture at a given pressure and temperature, as a function of helium concentration x The dashed line is a common tangent to the Gibbs energy curve Regions AB and CD correspond to metastable fluid mixtures, and the diffusion constant is not anomalous, except near B and C. The region between B and C corresponds to unstable mixtures (*b*) (*bottom*) The phase diagram of H-He mixtures for a given pressure. In region I the uniform mixture is thermodynamically favored In region II the uniform mixtures are metastable and diffusion is not anomalous. In region III the uniform mixture is unstable and undergoes spinodal decomposition The dashed line separates regions of normal and "anomalous" diffusion.

The important point for our considerations is that, provided one is not within or near region III in Figure 7*b*, the diffusion coefficient is not anomalous We will return to this point in Paper II, where the dynamics of the phase separation are discussed for a real system.

f) *Radiative Opacity*

At the temperatures of interest ($T \approx 10^4$ K), thermal photons have energies of order 1 eV. At the densities of interest ($\rho \approx 1$ g cm^{-3}), the electron plasmon energy is of order 30 eV. Photons cannot propagate below the plasmon energy and still undergo substantial absorption above the plasmon energy It follows that the radiative opacity exceeds the electron conduction "opacity" by many orders of magnitude in the metallic phase. It can therefore be ignored.

g) *Second-Order Transport Coefficients*

Among the many "second-order" transport coefficients, there are those which characterize the effect of simultaneous concentration, thermal, and pressure gradients in a nonconvecting fluid. First, there is the barodiffusion caused by the pressure gradient. In the applications to be discussed in Paper II, the composition varies over a smaller length scale than the pressure scale height, so the effect of barodiffusion is small (Of course, barodiffusion *does* nevertheless ensure that the zero temperature final state of a self-gravitating body is inhomogeneous.) Second, there is the effect of solute flux on the thermal gradient (the DuFour effect). The Onsager reciprocal relations ensure that this effect is always negligibly small for a dense fluid (Caldwell 1973). Third, there is the effect of the temperature gradient on the solute flux F_x (Landau and Lifshitz 1959),

$$F_x = -\rho D\left(\nabla x + \frac{k_T}{T}\nabla T\right), \tag{20}$$

where x is the fractional concentration of solute (i e , helium) and k_T is the Soret (or thermodiffusion) coefficient. This coefficient is not small in general: it can be as large as of order unity, and can have either sign. In a metal, an apparently successful model for k_T (Bhat and Swalin 1971) evaluates this coefficient as the sum of a "dense gas" contribution (determined by the mass and size of the pseudoatoms) and an electronic contribution, given by Gerl (1967). The

former was evaluated using the hard sphere diameters implied by thermodynamics, and the latter was evaluated using the conductivity calculations of Stevenson and Ashcroft (1974) Both contributions were positive and approximately $0.5x$ each, where x is the (assumed small) helium number fraction. In the situations of interest, we might therefore expect $k_T \approx +0.1$ As in the case of molecular diffusion, this result should be viewed with suspicion if the fluid is near a phase transition A positive value of k_T implies that the helium tends to diffuse toward colder regions. In most of the considerations in Paper II, k_T should be small enough to only slightly modify the solute flux (and certainly not change the direction of flux) We shall therefore ignore it

VIII. TRANSPORT PROPERTIES OF THE MOLECULAR PHASE

We repeat the considerations of the last section, but for the molecular phase.

a) Electrical Conductivity

Except near the molecular-metallic phase transition, molecular hydrogen is an insulator, and the only electrical conduction arises from impurities (Smoluchowski 1972). However, quite general considerations, together with recent band-structure calculations (Friedli and Ashcroft 1976), indicate that the indirect band gap in molecular hydrogen vanishes at or near the molecular-metallic transition Smoluchowski (1975) has pointed out that under these circumstances, the electronic conductivity at the phase transition could be within an order of magnitude of that given by equation (10).

b) Thermal Conductivity

If electrical conduction is almost metallic at the phase transition, then heat can be transported by electrons, with $\kappa \approx 0.1 \text{ cm}^2 \text{ s}^{-1}$ (eq. [12]) If no electronic degrees of freedom are available, then the less efficient molecular motions must be utilized. Neglecting the internal motion of the hydrogen molecule, this implies

$$\kappa \approx c\sigma \left(\frac{k_B T}{M}\right)^{1/2}, \tag{21}$$

where c is a correction factor of order unity, σ is a hard sphere diameter, and M is the mass of the molecule The correction factor can be deduced from Chapman-Enskog theory, or from Monte Carlo results for hard spheres (Alder, Gass, and Wainwright 1970). As usual, the hard sphere diameter is deduced from thermodynamic models (e g , § IV). For a hydrogen-rich fluid, the molecular contribution to κ is then

$$\kappa \approx 10^{-2} T_4^{-1/2} \quad \text{cm}^2 \text{ s}^{-1}, \tag{22}$$

accurate to perhaps a factor of 2, for $\rho \approx 1 \text{ g cm}^{-3}$.

c) Viscosity

Dense molecular fluids, like gases, have a Prandtl number close to unity This property is predicted by kinetic theories and Monte Carlo calculations (Alder, Gass, and Wainwright 1970), which show that both viscosity and thermal conductivity vary linearly as the Enskog correction. We shall not attempt to evaluate the Prandtl number more accurately, so it is adequate to use

$$\nu \approx 10^{-2} T_4^{-1/2} \quad \text{cm}^2 \text{ s}^{-1}. \tag{23}$$

If electronic transport is negligible, then $Pr \approx 1$ If electronic transport is almost metallic, then $Pr \approx 0.1$ or even 0 01.

d) Self-Diffusion

This transport coefficient is comparable to ν, but varies inversely as the Enskog correction and thus has a different density and temperature dependence. Using equation (21), with c given by Monte Carlo results (Alder, Gass, and Wainwright 1970), one finds

$$D \approx 4 \times 10^{-3} \rho^{-5/6} T_4^{3/2} \quad \text{cm}^2 \text{ s}^{-1} \tag{24}$$

for pure hydrogen or pure helium, to within a factor of 2.

e) Interdiffusion

The thermodynamic calculations (§ IV) indicate that the H_2 molecule is 15% larger than the helium atom The diffusion of a small amount of helium in hydrogen should therefore proceed slightly faster than the self-diffusion of hydrogen This effect is smaller than the probable inaccuracies in the calculation, so equation (24) suffices for the interdiffusion. As in the metallic case, this result should be viewed with caution near phase transitions

f) Second-Order Transport Coefficients

The only second-order coefficient that is likely to be important is k_T, the Soret coefficient. The dense-gas theory (Chapman and Cowling 1952) predicts $k_T \approx 0\,5x$, where x is the (assumed small) helium mole fraction The positive value is ensured by the greater mass of the helium atom and the strongly repulsive character of the intermolecular potentials As usual, this result is suspect near phase transitions.

g) Radiative Opacity

Unlike the preceding discussion, which has concentrated on the dense fluid regime ($\rho \approx 0\,1$ to 1 g cm^{-3}), the radiative opacity is of interest for a much wider range of densities and temperatures Interior models of Jupiter, for example, always assume an adiabatic molecular envelope, and do not allow for the possibility that molecular hydrogen may be sufficiently transparent for radiation to transport the internal heat flux subadiabatically Stevenson (1976*a*) has considered this problem, and concludes that molecular hydrogen

alone is sufficiently opaque to ensure convection, except at temperatures and pressures for which the 1500 cm^{-1} to 3000 cm^{-1} window in the hydrogen spectrum is important These calculations are based on the theory and observations of Linsky (1969), Welsh (1969), and Herzberg (1952) In Jupiter, the 1500 cm^{-1} to 3000 cm^{-1} window is most important for 400 K $\leqslant T \leqslant$ 700 K. For $T \leqslant$ 400 K, pure translational and rotation-translational pressure-induced bands provide sufficient opacity to ensure convection, until the optical depth to free space becomes less than unity at $T \approx$ 150 K (Trafton and Stone 1974; Wallace, Prather, and Belton 1974). At $T \geqslant$ 700 K, the vibration-rotation translational band ($\nu \approx$ 4000 cm^{-1}), and higher-order bands ($\nu \approx$ 8000 cm^{-1}, 12,000 cm^{-1}) ensure convection in Jupiter. Since the pressure-induced opacity varies roughly as P^2, where P is the pressure, and since the bands become broadened and overlapping at higher pressures, the radiative heat transport decreases as one goes deeper into the planet At even higher temperatures ($T \geqslant$ 3000 K) free-free absorption, arising from the small number of conduction electrons in the molecular fluid, begins to dominate. Unlike the free-free absorption usually considered (e g., Clayton 1968), the molecular fluid is so dense that the electron-molecule interactions are more important than electron-ion interactions in ensuring momentum conservation

The region 400 K $\leqslant T \leqslant$ 700 K is nevertheless probably convective, but only because of the small amounts of strongly absorbing molecules such as H_2O, CH_4, and NH_3. The opacities of these species are "spiky" at room temperature, with typical strong line separations of about 1 cm^{-1}. However, the pressure broadening exceeds the line spacing for pressures in excess of 5 or 10 bars, so that the opacity becomes quasi-continuous Assuming the validity of the quasi-continuous approximation, Stevenson (1976*a*) estimates that H_2O, CH_4, and NH_3 have sufficient combined opacity to "block" the 1500 cm^{-1} to 3000 cm^{-1} hydrogen window in Jupiter The data used in this calculation were Ferriso, Ludwig, and Thomson (1966) for H_2O; Burch and Williams (1962) and Plyler, Tidwell, and Blaine (1960) for CH_4; and Gille and Lee (1969) and Benedict, Plyler, and Tidwell (1958) for NH_3. Some uncertainty does remain, however, especially in the 2000–2500 cm^{-1} region where none of H_2O, CH_4, or NH_3 is strongly absorbing, so a careful band model is probably desirable

To conclude: A hydrogen-helium mixture is not sufficiently opaque to ensure convection in the deep atmosphere under typical conditions (such as those which prevail in Jupiter). The addition of a solar abundance of minor constituents (H_2O, CH_4, NH_3) probably suffices to reduce the radiative heat transport to less than 10% of the total and ensure an adiabatic thermal structure

IX. CONCLUSION

It is evident from our discussion of the phase diagram that the main uncertainty lies in the value of the critical temperature for the pure molecular-metallic hydrogen transition. Whereas this critical value is only known to about an order of magnitude, the metallic H-He critical temperature is known to perhaps 20%, and the H_2-He critical temperature to perhaps a factor of 2 This uncertainty forces us to consider a wide range of possibilities in Paper II (Stevenson and Salpeter 1977), where specific thermal and compositional evolutions are discussed Improvements in the value of the molecular-metallic hydrogen critical temperature will not be easy from purely theoretical calculations, and some experimental input is highly desirable

The partitioning of minor constituents is clearly difficult to predict quantitatively, with the exception of deuterium It is particularly desirable to understand more about the high-pressure properties of H_2O. Generally speaking, the relevant temperature ($\sim 10^4$ K) is too great for highly nonuniform partitioning of the kind that is observed in the Earth, for example. Constituents such as H_2O, CH_4, and NH_3 probably prefer molecular or helium-rich phases.

With two notable exceptions (electronic conductivity and radiative opacity of the molecular phase), the transport properties are known to within a factor of 3, typically. This is usually quite adequate for the purposes of Paper II The uncertainty in the electronic conductivity of the molecular phase near the molecular-metallic phase transition is of concern, since if electronic degrees of freedom are available for heat transport, then the efficiency of upward transport of helium by convection is generally low (see Paper II). The uncertainty in the radiative opacity is generally only large at those temperatures and pressures for which the opacity is one or more orders of magnitude in excess of that required to transport the heat flux at an adiabatic temperature gradient

Apart from the radiative opacity, where minor constituents are crucial, the effect of such molecules as H_2O, CH_4, and NH_3 on the phase diagram and transport properties is small, provided their abundances are close to solar.

We wish to thank N. W Ashcroft, M E Fisher, W B. Hubbard, and R Smoluchowski for discussions and comments. This work is supported by National Aeronautics and Space Administration grant NGR 33-010-188 and National Science Foundation grant AST 75-21153.

REFERENCES

Abraham, F R 1975*a*, *J Chem Phys*, 63, 157
——— 1975*b*, *J Chem Phys*, 63, 1316
Abraham, F R, Schreiber, D E, Mruzik, M R, and Pound, G M 1976, *Phys Rev Letters*, 36, 261
Abrikosov, A A 1964, *Soviet Astr—AJ*, 31, 112
Alder, B J, Gass, B M, and Wainwright, T E 1970, *J Chem Phys*, 53, 3813
Ascarelli, P, and Paskin, A 1968, *Phys Rev*, 165, 222

Ashcroft, N. W 1968, *Phys. Rev Letters*, **21**, 1748
Ashcroft, N W., and Langreth, D. C 1967, *Phys Rev*, **155**, 682
Aviram, I, Goshem, S, Rosenfeld, Y, and Thieberger, R 1976, *J. Chem Phys*, **65**, 846
Baym, G 1964, *Phys Rev*, **A135**, 1691
Beer, R, Farmer, C B, Norton, R H, Martonchik, J V, and Barns, T G 1972, *Science*, **175**, 1360
Benedict, W S, Plyler, E A, and Tidwell, E D. 1958, *J Chem. Phys*, **29**, 829.
Bhat, B N, and Swalin, R A 1971, in *Atomic Transport in Solids and Liquids*, ed A Lodding and T Lagerwall (Tubingen: Tubingen-Verlag), p 179
Brovman, E G, Kagan, Y, and Kholas, A 1972, *Soviet Phys—JETP*, **35**, 783
Brust, D 1972, *Phys Letters*, **40A**, 255.
Burch, D E, and Williams, D. 1962, *Appl Opt*, **1**, 587.
Caldwell, D R 1973, *J. Phys Chem*, **77**, 2004.
Caron, L G 1974, *Phys Rev*, **9B**, 5025
Chapman, S, and Cowling, T G 1952, *The Mathematical Theory of Non-Uniform Gases* (Cambridge Cambridge University Press), p 254
Clayton, D D 1968, *Principles of Stellar Evolution and Nucleosynthesis* (McGraw-Hill), chap. 3
Dyson, F 1971, *Ann Phys*, **63**, 1.
Ejima, T, and Yamamura, T. 1973, in *The Properties of Liquid Metals*, ed S Takenchi (London: Taylor and Francis), p 537
Etters, R D, Danilowicz, R, and England, W 1975, *Phys Rev*, **12A**, 2199.
Ferriso, C. C, Ludwig, C B, and Thomson, A L. 1966, *J Quant Spectrosc Rad Transf*, **6**, 241
Filinov, V S, and Norman, E G 1975, *Phys Letters*, **55A**, 219.
Firey, B, and Ashcroft, N W 1976, unpublished
Friedli, C, and Ashcroft, N W 1976, unpublished
Gerl, M 1967, *J Phys Chem Solids*, **28**, 725
Gille, J. C, and Lee, T-H 1969, *J. Atm Sci*, **26**, 932
Glyde, H R, Keech, G H, Mazighi, R, and Hansen, J. P 1976, *Phys Letters*, **58A**, 226
Grigoryev, F V, Kormer, S B, Mikhailova, O L, Tolochko, A P, and Urlin, V D 1972, *JETP Letters*, **16**, 201
Hamann, S D, and Linton, M 1966, *Trans. Faraday Soc*, **62**, 2234
Hammerberg, J, and Ashcroft, N W. 1974, *Phys Rev*, **9B**, 409
Hansen, J P., McDonald, I R, and Pollock, E. L. 1975, *Phys Rev*, **A11**, 1025.
Hansen, J. P, and Vieillefosse, P. 1976, *Phys Rev. Letters*, **37**, 391
Herzberg, G 1952, *Ap J*, **115**, 337
Hubbard, W. B 1974, *Ap J*, **190**, 223
Hubbard, W B, and DeWitt, H 1976, unpublished
Hubbard, W B, and Lampe, M. 1969, *Ap J Suppl*, **18**, 297.
Hubbard, W B., and Slattery, W L 1971, *Ap J*, **168**, 131.
——— 1976, in *Jupiter*, ed T Gehrels (University of Arizona Press), p 176
Hubbard, W B, and Smoluchowski, R 1973, *Space Sci. Rev*, **14**, 599.
Kerley, G I 1972, *Phys. Earth Planet. Inter*, **6**, 78
Krumhansl, J, and Wu, S-W. 1968, *Phys Letters*, **28A**, 263.
Landau, L, and Lifshitz, E M 1959, *Fluid Mechanics* (Reading, Massachusetts Addison-Wesley), p 224
Landau, L, and Zel'dovich, G 1943, *Acta Phys Chim (USSR)*, **18**, 194
Larson, H. P, Fink, U, Treffers, R R, and Gautier, T. N 1975, *Ap. J (Letters)*, **197**, L137
Lebowitz, J, and Rowlinson, J. S 1964, *J Chem. Phys*, **41**, 133
Linsky, J L 1969, *Ap. J*, **156**, 989
Longuet-Higgins, H C, and Pople, J A 1956, *J Chem. Phys*, **25**, 884
McMahan, A, Beck, H, and Krumhansl, J 1974, *Phys Rev.*, **9A**, 1852.
Monkhurst, H J, and Oddershede, J 1973, *Phys Rev Letters*, **30**, 797
Neece, G A, Rogers, F J, and Hoover, W. G 1971, *J. Comput Phys*, **7**, 621
Østgaard, E 1974, *Physica*, **74**, 113
Plyler, E A., Tidwell, E D, and Blaine, L R 1960, *J Res NBS*, **64A**, 201
Podolak, M 1977, *Icarus*, **30**, 155
Podolak, M, and Cameron, A G W 1975, *Icarus*, **25**, 627
Pollock, E L, and Alder, B 1977, *UCRL Rept* No 79511.
Pollock, E L, and Hansen, J P 1973, *Phys Rev*, **8A**, 3110
Ramaker, D E, Kumar, L., and Harris, F E 1975, *Phys Rev. Letters*, **34**, 812
Ramsey, W H. 1963, *M N R A S*, **125**, 469
Ree, F H, and Bender, C F 1974, *Phys. Rev Letters*, **32**, 85
Ross, M 1972, *J Chem Phys*, **56**, 4651.
——— 1974, *J. Chem Phys*, **60**, 3634.
Ross, M., and McMahan, A. L 1976, *Phys Rev*, **B13**, 5154.
Salpeter, E E 1954, *Australian J. Phys*, **7**, 353.
Salpeter, E E, and Zapolsky, H. 1967, *Phys. Rev*, **158**, 876
Shafer, R, and Gordon, R G. 1973, *J. Chem Phys*, **58**, 5422
Silver, D M, and Stevens, R M 1973, *J Chem Phys*, **59**, 3378.
Simcox, L. N, and March, N. H 1962, *Proc Phys Soc. London*, **80**, 830
Smoluchowski, R. 1971, *Ap. J*, **166**, 435
———. 1972, *Phys Earth Planet Inter*, **6**, 48
———. 1973, *Ap J. (Letters)*, **185**, L95
——— 1975, *Ap J (Letters)*, **200**, L119
Stevenson, D. J 1975, *Phys Rev*, **12B**, 3999
——— 1976*a*, Ph D. thesis, Cornell University.
——— 1976*b*, *Phys Letters*, **58A**, 282
Stevenson, D. J, and Ashcroft, N W 1974, *Phys Rev*, **9A**, 782
Stevenson, D J, and Salpeter, E E 1976, in *Jupiter*, ed T Gehrels (University of Arizona Press), p 85
———. 1977, *Ap J Suppl*, **35**, 239 (Paper II)
Straus, D M, and Ashcroft, N W 1977, *Phys Rev Letters*, **38**, 415
Straus, D. M, Ashcroft, N W, and Beck, H 1977, in preparation
Streett, W. B 1976, *Ap J.*, **186**, 1107.
Stroud, D 1973, *Phys Rev*, **B7**, 4405
Stroud, D, and Ashcroft, N W 1972, *Phys Rev*, **B5**, 371
Trafton, L M, and Stone, P H 1974, *Ap J*, **188**, 649
Trauger, J T, Roesler, F. L, Carleton, N P, and Traub, W A 1973, *Ap J. (Letters)*, **184**, L137
Trubitsyn, V P 1967, *Soviet Phys—Solid State*, **8**, 2593
——— 1972, *Soviet Astr—AJ*, **16**, 342
Vadovic, C J, and Colver, C P 1971, *Phil Mag*, **24**, 509
Van Horn, H 1967, *Phys Rev*, **157**, 342
Vereschchagin, L F, Yakovlev, E N, and Timofeev, Yu A 1975*a*, *JETP Letters*, **21**, 85.
———. 1975*b*, *JETP Letters*, **21**, 304.
Vieillefosse, P, and Hansen, J. P 1975, *Phys Rev.*, **A12**, 1106
Wainwright, T, and Alder, B 1958, *Il Nov Cimento Suppl*, p. 116
Wallace, L., Prather, M, and Belton, M J. S 1974, *Ap. J*, **193**, 481
Welsh, H L 1969, *J Atm. Sci*, **26**, 835.
Wigner, E., and Huntington, H B 1935, *J Chem Phys*, **3**, 764
Witten, T. A, Jr. 1974, *Ap. J*, **188**, 615.
Zharkov, V. N., Makalkin, A. B., and Trubitsyn, V. P. 1975, *Soviet Astr—AJ*, **18**, 768

E. E. SALPETER. Newman Laboratory of Nuclear Studies, Cornell University, Ithaca, NY 14853

D. J. STEVENSON: Research School of Earth Sciences, ANU, P.O. Box 4, Canberra 2600, Australia

Thermal diffuse x-ray scattering in simple metals*

David M Straus and N. W Ashcroft
Laboratory of Atomic and Solid State Physics and Materials Science Center, Cornell University, Ithaca, New York 14853
(Received 17 November 1975)

Calculations are reported for the ionic structure factor and x-ray scattering cross section of sodium (at $T = 0$ and 90°K) and lithium (both isotopes at $T = 0$°K) within the harmonic approximation An evaluation of the appropriate displacement-displacement correlation function by the special-point method circumvents the need for a multiphonon expansion In the case of sodium, the structure in the one-phonon scattering is straightforwardly accounted for and an approximate expansion is obtained for all multiphonon scattering By treating core and conduction electrons on an equal footing it is shown that information on the conduction-electron system is present in the forward-scattering component In lithium the one-phonon cross section at small angles aids in the determination of the effective electron-ion interaction

I INTRODUCTION

For some years x-ray thermal diffuse scattering (TDS) has been used as a probe of lattice dynamics in simple materials [1-4] Although information on the phonon frequencies and polarizations (and also the extent of anharmonicity) is contained in the TDS,[4,5] it is generally hard to extract.[6] The cross section for the scattering of x rays intimately involves the static structure factor of the ions, $S_{\rm ion}(\vec{k})$.[7] The purpose of this paper is to present calculations of (i) $S_{\rm ion}(\vec{k})$, and (ii) the x-ray scattering cross section for Na and Li in the harmonic approximation and in their ground states. The significant features of the calculation are the use of a special point technique[8,9] in the computation of the equal time displacement-displacement correlation function $\langle \vec{u}_i \vec{u}_j \rangle$ [which enters into $S_{\rm ion}(\vec{k})$] and the separation of the scattering cross section into contributions from core and valence electrons In particular, the special point technique enables us to avoid the customary expansion[4] of the inelastic part of $S_{\rm ion}(\vec{k})$ into terms involving the scattering of a definite number of phonons. We determine the "one-phonon" term explicitly, but we can also calculate *all* higher-order processes without recourse to expansion Further, our treatment of the contribution of the valence electrons to the cross section shows that x-ray scattering should yield information, in light metals, on the effective electron-ion interaction, as we demonstrate for the particular case of Li.

Section II contains a derivation of the x-ray scattering cross section $d\sigma/d\Omega$ in a model of a simple metal which distinguishes between bound and conduction electrons. In Sec. III we outline the calculation of $S_{\rm ion}(\vec{k})$ using the special point technique (discussed in detail in the Appendix), and compare it with the other nonexpansion techniques in the literature Section IV presents numerical results for $S_{\rm ion}(\vec{k})$ and $d\sigma/d\Omega$ for Na (at two temperatures) and for both isotopes of Li. We draw particular attention to the secondary maxima associated with the one-phonon term as observed in certain crystallographic directions. These maxima have special importance in the determination of the electron-ion interaction of Li, and also give information about specific portions of the phonon spectrum directly

II THEORY

The differential cross section for scattering of a photon from a solid of N ions in volume V(at $T = 0$ °K) is proportional to the space-time Fourier transform of the Van Hove correlation function $G_e(\vec{r}, t)$:

$$\frac{d^2\sigma}{d\Omega\, d\omega} = \frac{C}{V} \int d^3r \int_{-\infty}^{\infty} dt\, G_e(\vec{r}, t) \exp(i\vec{k}\cdot\vec{r} - i\omega t), \tag{2.1}$$

where C is a constant,[10-12]

$$G_e(\vec{r}, t) = \int d^3x \langle \hat{n}(\vec{x}, 0)\hat{n}(\vec{x}+\vec{r}, t)\rangle \tag{2.2}$$

and

$$\vec{k} = \vec{k}_i - \vec{k}_f, \quad \omega = \omega_i - \omega_f. \tag{2.3}$$

We are considering the cross section per unit volume for scattering a photon of momentum $\hbar k_i$ and energy $\hbar\omega_i$ into a solid angle $d\Omega$ with energy loss between $\hbar\omega$ and $\hbar(\omega + d\omega)$. The quantities $\hbar k_f$ and $\hbar\omega_f$ are, respectively, the momentum and energy of the scattered photon. In Eq. (2.2), $\hat{n}(\vec{r}, t)$ is the total electron number density operator and the angular brackets $\langle\ \rangle$ refer to a ground-state average Introducing spatial Fourier transforms

$$\frac{d^2\sigma}{d\Omega\, d\omega} = \frac{C}{V} \int_{-\infty}^{\infty} dt\, e^{-i\omega t} \langle \hat{n}(-\vec{k}, 0)\hat{n}(\vec{k}, t)\rangle, \tag{2.4}$$

where $\hat{n}(\vec{k})$ is the Fourier transform of $\hat{n}(\vec{r})$.

We separate $\hat{n}(\vec{r})$ into contributions from core

and valence electrons, and we treat the core electrons as if they were rigidly attached to the ions. Any core excitations or distortions of the ions are therefore neglected; should these occur they must be calculated separately. In practical terms this means that in comparing experiment and theory the Compton scattering from the *core* electrons must first be subtracted from the data. In addition we invoke the adiabatic approximation, so that the conduction electrons (ce) are always in a ground state appropriate to an instantaneous ion configuration (ion). By virtue of the rigid-ion approximation we may write

$$\hat{n}(\vec{k},t)=\sum_i e^{i\vec{k}\cdot\vec{R}_i(t)}f(\vec{k})+\hat{n}_{ce}(\vec{k},t) \tag{2.5}$$

Here $f(\vec{k})$ is the Fourier transform of the average core-electron density about a nucleus at the origin, and $\vec{R}_i(t)$ refers to the instantaneous position of the ion labeled i. From Eqs. (2.4), (2.5), and the adiabatic approximation, we then find

$$\frac{d^2\sigma}{d\Omega\,d\omega}\frac{V}{C}=\int_{-\infty}^{\infty}dt\,e^{-i\omega t}\Big(\Big\langle\sum_{ij}e^{-i\vec{k}\cdot\vec{R}_i(0)}e^{i\vec{k}\cdot\vec{R}_j(t)}\Big\rangle_{ion}|f(\vec{k})|^2+\Big\langle\sum_i e^{-i\vec{k}\cdot\vec{R}_i(0)}f(-\vec{k})\langle\hat{n}_{ce}(\vec{k},t)\rangle_{ce}\Big\rangle_{ion}$$
$$+\Big\langle\sum_i e^{i\vec{k}\cdot\vec{R}_i(t)}f(\vec{k})\langle\hat{n}_{ce}(-\vec{k},0)\rangle_{ce}\Big\rangle_{ion}\Big)+\int_{-\infty}^{\infty}dt\,e^{-i\omega t}\langle\hat{n}_{ce}(-\vec{k},0)\hat{n}_{ce}(\vec{k},t)\rangle_{ce,ion}. \tag{2.6}$$

We suppose that the interaction between conduction electrons and ions can be represented by a weak pseudopotential with Fourier transform $v(\vec{k})$ (as is the case for many simple metals). The density response may then be calculated to linear order in $v(\vec{k})$:

$$\langle\hat{n}_{ce}(\vec{k},t)\rangle_{ce}=\chi_1(\vec{k})v(\vec{k})\sum_i e^{i\vec{k}\cdot\vec{R}_i(t)}, \tag{2.7}$$

with

$$\chi_1(\vec{k})=(k^2/4\pi e^2)[1/\epsilon(\vec{k},0)-1], \tag{2.8}$$

$\epsilon(\vec{k},0)$ being the static dielectric function of the uniform interacting electron gas [13] Equations (2.6)–(2.8) now give

$$\frac{d^2\sigma}{d\Omega\,d\omega}\frac{V}{C}=\int_{-\infty}^{\infty}dt\,e^{-i\omega t}\Big(\Big\langle\sum_{ij}e^{-i\vec{k}\cdot(\vec{R}_i(0)-\vec{R}_j(t))}\Big\rangle_{ion}\times(|f(\vec{k})|^2+2f(\vec{k})\chi_1(\vec{k})v(\vec{k}))+\langle\hat{n}_{ce}(-\vec{k},0)\hat{n}_{ce}(\vec{k},t)\rangle_{ce,ion}\Big) \tag{2.9}$$

In a typical x-ray experiment *all* the radiation emerging at a given angle is initially measured.[14] All possible energy transfers (on the scale of typical electron and phonon energies) $\hbar\omega$ are therefore included, and we pass from the cross section for energy loss $\hbar\omega(d^2\sigma/d\Omega\,d\omega)$ to the total angular cross section $(d\sigma/d\Omega)$:

$$\frac{d\sigma}{d\Omega}=\int_{-\infty}^{\infty}d\omega\,\frac{d^2\sigma}{d\Omega\,d\omega}$$
$$=2\pi\frac{C}{V}\Big(\Big\langle\sum_{ij}e^{-i\vec{k}\cdot(\vec{R}_i-\vec{R}_j)}\Big\rangle_{ion}(|f(\vec{k})|^2+2f(\vec{k})\chi_1(\vec{k})v(\vec{k}))+\langle\hat{n}_{ce}(-\vec{k})\hat{n}_{ce}(\vec{k})\rangle_{ce,ion}\Big). \tag{2.10}$$

Note that the last term is usually considered part of the Compton scattering, and is therefore generally subtracted from the primary data.[15] What will become apparent, in Sec. IV, is that the value of the last term in Eq. (2.10) (the valence electron correlation function) should be readily obtainable from x-ray measurements. The theoretical results we present are therefore best compared to data from which only the *ionic* Compton scattering has been subtracted.

The last term in Eq. (2.10) is difficult to calculate for interacting electrons in the presence of the ions. For purposes of illustration we use the free-electron value.[13,16]

$$N_eS_e(\vec{k})\equiv\langle\hat{n}_{ce}(-\vec{k})\hat{n}_{ce}(\vec{k})\rangle_{ce,free},$$
$$S_e(\vec{k})=\Big(\frac{3k}{4k_F}-\frac{1}{16}\frac{k^3}{k_F{}^3}\Big),\quad 0<k\leq 2k_F \tag{2.11}$$
$$S_e(\vec{k})=1,\qquad k\geq 2k_F.$$

Here N_e is the number of electrons, and k_F the Fermi wave vector. Setting (for a monovalent system) the number of electrons N_e equal to the number of ions N, Eqs. (2.10) and (2.11) give us the final result

$$W\equiv\frac{d\sigma}{d\Omega}\frac{V}{N}\frac{1}{2\pi C}$$
$$=S_{ion}(\vec{k})(|f(\vec{k})|^2+2f(\vec{k})\chi_1(\vec{k})v(\vec{k}))+S_e(\vec{k}), \tag{2.12}$$

where we have set

$$S_{ion}(\vec{k})=\frac{1}{N}\Big\langle\sum_{ij}e^{i\vec{k}\cdot(\vec{R}_i-\vec{R}_j)}\Big\rangle_{ion}\quad\text{for }\vec{k}\neq 0. \tag{2.13}$$

It should be clear that except for the elements of lowest atomic number (e.g., Li), $S_e(\vec{k})$ makes a small contribution to W for all but the smallest wave vectors $k<2k_F$.

III. CALCULATION OF IONIC STRUCTURE FACTOR

We now proceed to a calculation of $S_{\text{ion}}(\vec{k})$ in a model in which the solid is treated as a harmonic crystal. Letting $\vec{R}_i = \vec{X}_i + \vec{u}_i$, where $\vec{X}_i$ is the equilibrium position of the ith ion and $\vec{u}_i$ its displacement

$$S_{\text{ion}}(\vec{k}) = \frac{1}{N}\sum_{ij} e^{-i\vec{k}\cdot\vec{X}_{ij}}\langle e^{-i\vec{k}\cdot(\vec{u}_i-\vec{u}_j)}\rangle_{\text{ion}}. \tag{3.1}$$

Here, $\vec{X}_{ij} = \vec{X}_i - \vec{X}_j$, and the average is to be taken over the states appropriate to a harmonic crystal. With the definitions

$$\langle (u_{i\alpha}-u_{j\alpha})(u_{i\beta}-u_{j\beta})\rangle_{\text{ion}} \equiv \lambda_{\alpha\beta}(\vec{X}_i-\vec{X}_j) \tag{3.2}$$

and

$$\vec{u} = (u_1, u_2, u_3),$$

we have the result[4,17]

$$\langle e^{-i\vec{k}\cdot(\vec{u}_i-\vec{u}_j)}\rangle_{\text{ion}} = e^{-k_\alpha k_\beta \lambda_{\alpha\beta}(\vec{X}_i-\vec{X}_j)/2}, \tag{3.3}$$

where

$$\lambda_{\alpha\beta}(\vec{X}_i-\vec{X}_j) = \frac{\hbar}{MN}\sum_{\vec{q}j}^{\text{BZ}} (1-\cos\vec{q}\cdot\vec{X}_{ij}) e_\alpha(\vec{q}j) \times e_\beta(\vec{q}j)\frac{1}{\omega(\vec{q}j)}\coth[\tfrac{1}{2}\beta\hbar\omega(\vec{q}j)],$$

$$\beta = \frac{1}{k_B T}, \tag{3.4}$$

and M is the mass of an ion. In Eq. (3.4), $\omega(\vec{q}j)$ and $\vec{e}(\vec{q}j)$ are the frequency and polarization vector of the normal mode of wave vector $\vec{q}$ and polarization index j ($j = 1, 2, 3$). The $\vec{q}$ sum extends over the entire first Brillouin zone (BZ). Using the translational symmetry of the lattice, Eqs. (3.1)–(3.4) yield

$$S_{\text{ion}}(\vec{k}) = \sum_i e^{-i\vec{k}\cdot\vec{X}_i} e^{-k_\alpha k_\beta \lambda_{\alpha\beta}(\vec{X}_i)/2}. \tag{3.5}$$

Next we separate $\lambda_{\alpha\beta}(\vec{X}_i)$ as follows:

$$\lambda_{\alpha\beta}(\vec{X}_i) = \Lambda_{\alpha\beta}(0) - \Lambda_{\alpha\beta}(\vec{X}_i), \tag{3.6}$$

with

$$\Lambda_{\alpha\beta}(\vec{X}_i) = \frac{\hbar}{MN}\sum_{\vec{q}j} e_\alpha(\vec{q}j) e_\beta(\vec{q}j)\frac{1}{\omega(\vec{q}j)} \times \coth[\tfrac{1}{2}\beta\hbar\omega(\vec{q}j)]\cos(\vec{q}\cdot\vec{X}_i). \tag{3.7}$$

Note that

$$\Lambda_{\alpha\beta}(\vec{X}_{ij}) = 2\langle u_{i\alpha}u_{j\beta}\rangle_{\text{ion}} \tag{3.8}$$

We see, therefore, that $\Lambda_{\alpha\beta}(\vec{X})$ is the displacement-displacement correlation function for two ions separated (on average) by $\vec{X}$. Clearly $\Lambda_{\alpha\beta}(0)$ is the displacement-displacement autocorrelation function. For a cubic system,

$$\langle u_{i\alpha}u_{i\beta}\rangle_{\text{ion}} = \delta_{\alpha\beta}\tfrac{1}{3}\langle \vec{u}_i\cdot\vec{u}_i\rangle, \tag{3.9}$$

so that

$$\Lambda_{\alpha\beta}(0) = \delta_{\alpha\beta}\frac{1}{3}\frac{\hbar}{MN}\sum_{\vec{q}j}\frac{1}{\omega(\vec{q}j)}\coth[\tfrac{1}{2}\beta\hbar\omega(\vec{q}j)] \equiv \delta_{\alpha\beta}\Lambda^0. \tag{3.10}$$

This defines Λ^0, which is closely related to the Debye-Waller[4] factor e^{-2W}.

$$2W = \tfrac{1}{2}k_\alpha k_\beta \Lambda_{\alpha\beta}(0) = \tfrac{1}{2}k^2\Lambda^0 \tag{3.11}$$

Substituting Eqs. (3.6)–(3.11) into Eq. (3.5), we find

$$S_{\text{ion}}(\vec{k}) = \sum_i e^{-i\vec{k}\cdot\vec{X}_i} e^{-k_\alpha k_\beta[\Lambda_{\alpha\beta}(0)-\Lambda_{\alpha\beta}(\vec{X}_i)]/2} = \sum_i e^{-i\vec{k}\cdot\vec{X}_i} e^{-k^2\Lambda^0/2} e^{k_\alpha k_\beta \Lambda_{\alpha\beta}(\vec{X}_i)/2}. \tag{3.12}$$

To proceed from this point the usual approach is to expand the last exponential in a power series in $\Lambda_{\alpha\beta}(\vec{X}_i)$. The leading (i.e., constant) term gives the elastic (Bragg) scattering peaks, the second gives the one-phonon scattering, the third the two-phonon scattering, and so forth. Beyond the one-phonon contribution each term is increasingly laborious to evaluate. We can avoid this expansion however, by writing $S_{\text{ion}}(\vec{k})$ as follows:

$$S_{\text{ion}}(\vec{k}) = \sum_i e^{-i\vec{k}\cdot\vec{X}_i} e^{-k^2\Lambda^0/2} + \sum_i e^{-i\vec{k}\cdot\vec{X}_i} e^{-k^2\Lambda^0/2}[\tfrac{1}{2}k_\alpha k_\beta \Lambda_{\alpha\beta}(\vec{X}_i)] + \sum_i e^{-i\vec{k}\cdot\vec{X}_i} e^{-k^2\Lambda^0/2}[e^{k_\alpha k_\beta \Lambda_{\alpha\beta}(\vec{X}_i)/2} - 1 - \tfrac{1}{2}k_\alpha k_\beta \Lambda_{\alpha\beta}(\vec{X}_i)] \equiv S_0(\vec{k}) + S_1(\vec{k}) + S_M(\vec{k}) \tag{3.13}$$

Here $S_0(\vec{k})$ gives the elastic scattering, i.e.,

$$S_0(\vec{k}) = Ne^{-k^2\Lambda^0/2}\sum_{\vec{K}}\delta_{\vec{k},\vec{K}}, \tag{3.14}$$

the $\vec{K}$ being the vectors of the reciprocal lattice. The one-phonon scattering term $S_1(\vec{k})$ is easily seen to be

$$S_1(\vec{k}) = e^{-k^2\Lambda^0/2}\frac{\hbar}{2M}k_\alpha k_\beta \times \sum_j e_\alpha[\vec{q}(\vec{k})j]\, e_\beta[\vec{q}(\vec{k})j]\frac{1}{\omega[\vec{q}(\vec{k})j]} \times \coth\{\tfrac{1}{2}\hbar\beta\omega[\vec{q}(\vec{k})j]\}, \quad (3\ 15)$$

where $\vec{q}(\vec{k})$ is the vector $\vec{k}$ reduced by an appropriate $\vec{K}$ to the first Brillouin zone [i.e., $\vec{q}(\vec{k}) = \vec{k} - \vec{K}$]. Finally, the remainder $S_M(\vec{k})$ will be calculated by *direct computation of* $\Lambda_{\alpha\beta}(\vec{X}_i)$, so that all higher-order phonon terms are automatically taken into account. The reason for adopting this procedure is to assure convergence in the sum over i in $S_M(\vec{k})$. This will be clarified in what follows.

Our method of calculation of $\Lambda_{\alpha\beta}(\vec{R}_i)$ and Λ^0 makes use of *special points* in the first Brillouin zone[8,9] to evaluate the integral of Eq. (3.7). By calculating the integrand at these relatively few special points, one obtains a good approximation to the entire integral. This procedure differs markedly from ordinary numerical integration in that (as shown in the Appendix) one is effectively using an expansion of the integrand in symmetrized plane waves. In connection with this method we draw attention to the behavior of $\Lambda_{\alpha\beta}(\vec{X})$ for large $\vec{X}$. At large $\vec{X}$ the dominant contribution to the integral in Eq. (3.7) comes from small $\vec{q}$, and it can be shown[18] that at $T = 0°K$,

$$\lim_{X\to\infty} \Lambda_{\alpha\beta}(\vec{X}) \sim 1/X^2. \quad (3.16)$$

Thus to ensure convergence in $S_M(\vec{k})$ it is necessary to make the separation indicated in Eq. (3.13).

The method may be compared with the nonexpansion calculations of $S_{ion}(\vec{k})$ by (i) Lomer,[19] who calculates the ionic structure factor directly using the results of a computer experiment; (ii) Semenovskaya and Umanskii,[20] who calculate $\Lambda_{\alpha\beta}(\vec{X})$ in closed form for a model sinusoidal phonon dispersion law; and (iii) Reid and Smith,[21] who calculate the multiphonon scattering $S_M(\vec{k})$ for crystals whose sizes range between 100 and 1000 unit cells. Their evaluation of $\Lambda_{\alpha\beta}(\vec{X})$ is achieved by summing over only those $\vec{q}$ corresponding to the normal modes of such a finite crystal. By separately calculating the $\vec{q}\to 0$ portion of the integral in Eq. (3.7), they find that a crystal of 500 unit cells gives essentially the same $S_M(\vec{k})$ as an infinite crystal, for $\vec{q}(\vec{k})$ belonging to the set of normal modes of the finite crystal.

The method of Reid and Smith appears to be the most accurate and practical, but has the disadvantages that one can calculate $S_M(\vec{k})$ at relatively few points, and that the matrices $\Lambda_{\alpha\beta}(\vec{X})$ for a real crystal are inaccessible. We are able to circumvent these limitations by *directly* calculating the correlation matrices $\Lambda_{\alpha\beta}(\vec{X})$. (These are of considerable interest, of course, in a wide range of problems.)

We illustrate the method by its application to Na and Li. In both cases the phonon spectrum was calculated from a force-constant model designed to fit the experimental data. The corresponding $S_{ion}(\vec{k})$ has been calculated for Na at two temperatures (0 and 90°K) and for both isotopes of Li (at $T = 0°K$).

In the case of Na the force constants were those that fit the data at $T = 90°K$.[22] A simple estimate (supported by some theoretical results[23]) indicates that the change in phonon frequencies between 0 and 90°K is everywhere less than the experimental error. Hence the only effect of temperature we allow is through the hyperbolic cotangent function in Eq. (3.7).[24] To simplify the calculation we use the $T = 0°K$ value of $\Lambda_{\alpha\beta}(\vec{X}_i)$ for $\vec{X}_i \neq 0$ in the 90°K calculation, but use the $T = 90°K$ value of $\Lambda_{\alpha\beta}(0)$.[25] *The 90°K results are therefore meant to be indicative of the effects of temperature, but they are only approximate.* We use the value of r_s determined from the 5°K lattice constant measurement,[26] i.e., $r_s = 3\ 931$ a.u. (r_s is defined by $\frac{4}{3}\pi(r_s a_0)^3 = V/N_e$, where a_0 is the Bohr radius.)

The force constants for ^{7}Li were similarly taken to be those which fit the experimental phonon dispersion[27] measured at $T = 98°K$. The value of r_s was also deduced from the lattice constant,[26] in this case at 78°K ($r_s = 3.248$ a.u.). To calculate $S_{ion}(\vec{k})$ we have set $T = 0°K$. In order to obtain $S_1(\vec{k})$, Λ^0, and $\Lambda_{\alpha\beta}(\vec{X}_i)$ for ^{6}Li, we have assumed that both substances are truly harmonic. This gives

$$\omega \propto M^{-1/2}, \quad \Lambda_{\alpha\beta}(\vec{X}) \propto M^{-1/2} \text{ for all } \vec{X}, \quad \text{and} \quad S_1(\vec{k}) \propto M^{-1/2}. \quad (3.17)$$

IV. RESULTS

In this section we present numerical results for both $S_{ion}(\vec{k})$ and the x-ray scattering cross sections for Na and Li. The structure factor calculations were carried out as described above. As regards the cross sections, we give two sets of results. One corresponds to the theory outlined in Sec. II:

$$W = \frac{d\sigma}{d\Omega}\frac{V}{N}\frac{1}{2\pi C} = S(\vec{k})[|f(\vec{k})|^2 + 2f(\vec{k})\chi_1(\vec{k})v(\vec{k})] + S_e(\vec{k}), \quad (4\ 1)$$

while the other corresponds to the more commonly used expression

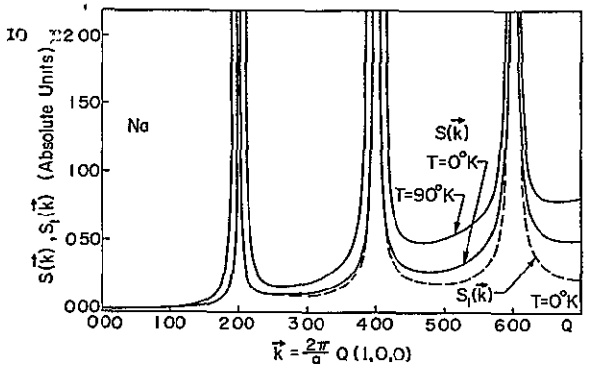


FIG 1 Structure factor $S(\vec{k})$ and the one-phonon contribution $S_1(\vec{k})$ for Na at $T=0$ °K and at $T=90$ °K along [100]

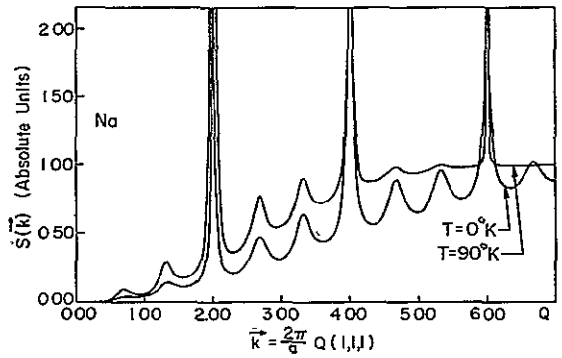


FIG. 3. Structure factor $S(\vec{k})$ for Na at $T=0$ °K and $T=90$ °K along [111].

$$W_a=\left(\frac{d\sigma}{d\Omega}\frac{V}{N}\frac{1}{2\pi C}\right)_a=S(\vec{k})|f_a(\vec{k})|^2. \tag{4.2}$$

Here $f_a(\vec{k})$ is the Fourier transform of the average electron density of an assumed neutral atom, and we write (and shall continue to do so) $S(\vec{k})$ in place of $S_{ion}(\vec{k})$. Both the ionic (Na^+, Li^+) and the atomic (Na, Li) form factors were taken from Ref 28 The Geldart and Vosko[29] modified form of the Hubbard dielectric function $\epsilon(\vec{k})$ was used, as well as an empty-core pseudopotential to represent the effective electron-ion interaction.[30] Figures 1–4 show $S(\vec{k})$ for Na, and Figs. 5–7 show $S(\vec{k})$ for both isotopes of Li.[31] We present both cross sections W and W_a for Na (at $T=0$ °K) in Figs. 8–11, and in Figs. 12–14 we show W for Li (at $T=0$ °K) with two choices of the core radius appearing in the empty core pseudopotential.

The most noticeable feature of the structure factor plots is the sizable structure between the Bragg peaks along all directions *except* the [100] and [110] directions (for a bcc lattice). These maxima are a direct consequence of the behavior of the one-phonon term.[32] Their occurrence is completely general, and has been noted for quite some time.[33] For the sake of simplicity, however, we can most easily explain them in terms of a (polarization-independent) Debye model. Here (at $T=0$°K),

$$S_1(\vec{k})=e^{-k^2A^0/2}\frac{\hbar}{2M}k^2\frac{1}{\omega(\vec{q}(\vec{k}))}, \tag{4.3}$$

where $\omega(\vec{q})=cq$ is independent of polarization and c is the approximate speed of sound. We have plotted in Fig. 15 lines along which the function $1/c|\vec{q}(\vec{k})|$ has constant value for a (001) plane of the reciprocal lattice of a bcc crystal. In any direction (except [100] and [110]), and as a consequence of periodicity alone, the one-phonon term displays secondary maxima as one passes over the ridges of the function shown. Replacing Eq (4 3) with Eq. (3.15) introduces three frequencies (one for each polarization j) at every point, each weighted by the factor $[\vec{k}\cdot\vec{e}(j\vec{q}(\vec{k}))]^2$. For example, Fig. 15 would indicate two secondary maxima between the points[34] (0, 0, 0) and (3, 1, 0), whereas Fig 4 shows only one. The value of the one-phonon term at the point along[310] marked P on Fig. 10 is determined

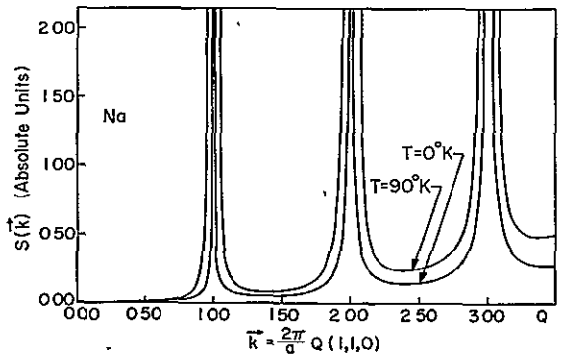


FIG. 2. Structure factor $S(\vec{k})$ for Na at $T=0$ °K and $T=90$ °K along [110]

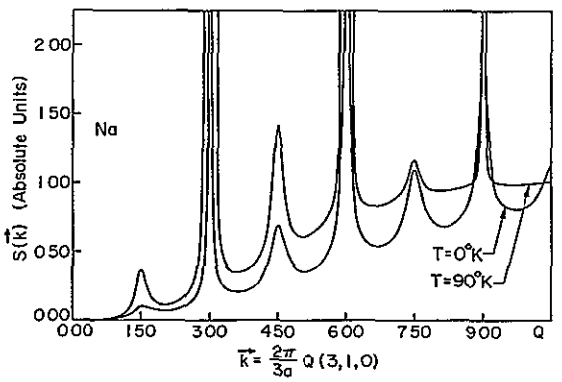


FIG. 4. Structure factor $S(\vec{k})$ for Na at $T=0$ °K and $T=90$ °K along [310].

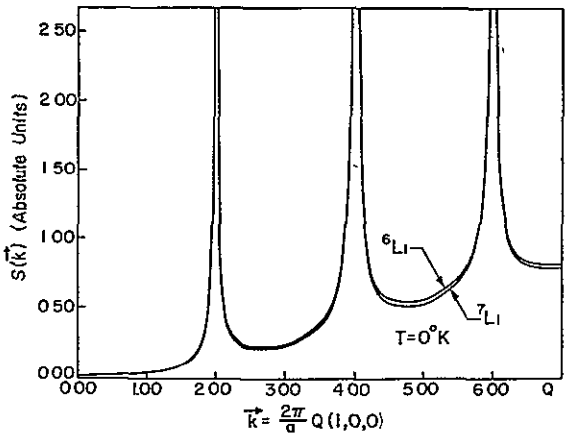


FIG 5. Structure factor $S(\vec{k})$ for ^{6}Li and ^{7}Li at $T=0$ °K along [100]

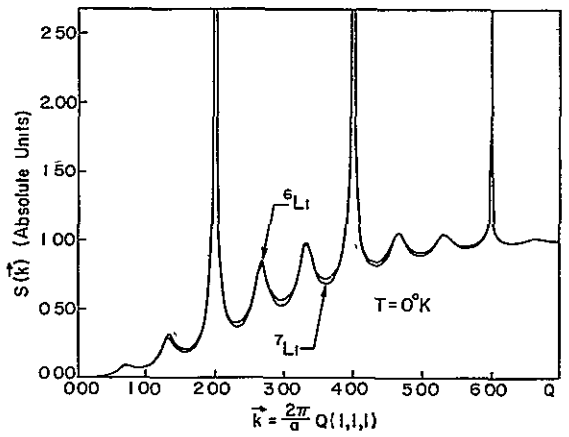


FIG 7. Structure factor $S(\vec{k})$ for ^{6}Li and ^{7}Li at $T=0$ °K along [111]

by the phonons at the point $\vec{q}=(\frac{1}{2},\frac{1}{2},0)$ in the first Brillouin zone. At $(\frac{1}{2},\frac{1}{2},0)$, Na has an anomalously low transverse frequency.[22] Furthermore, since $\vec{q}$ is nearly perpendicular to the [310] direction, the factor $\sum_j[\vec{e}(j\vec{q})\cdot\vec{k}]^2$ will select out the *transverse* frequencies. The resulting single, large, maximum swamps any other effects Thus we see that *any particularly low phonon frequency will cause a sequence of one-phonon maxima along the appropriate direction.* This property of the one-phonon scattering has been widely used to study soft modes,[35] but the discussion is often set in real space. In terms of identifying the maxima with a particular vibrational mode we see that it is advantageous to treat the problem in reciprocal space.

The comparison of W and W_a for Na in Figs. 8–10 shows that at large $\vec{k}$ the only significant difference is a shift arising from the term $S_e(\vec{k})$ in W, which is a constant for $k>2k_F$. However, at small $\vec{k}$ Fig. 11 shows that the presence of $S_e(\vec{k})$ in W contributes to a difference in shape between W and W_a. The small $\vec{k}$ portion of the x-ray cross section (with only *ionic* Compton scattering subtracted out) thus gives us information about the conduction electrons.[36] Note also that for Na the presence of the pseudopotential $v(\vec{k})$ in W seems to make little difference in the final cross section. This is *not* so for elements of very low atomic number. For example, in Figs. 12 and 13 we plot W for ^{7}Li at low values of $\vec{k}$, for two choices of the core radius appearing in the empty-core pseudopotential.[37] The maximum percentage difference is slight in both cases, but in Fig. 13 the actual shape of the one-phonon maximum is noticeably altered. In fact, the differences between pseudopotentials will always be most noticeable in low-$\vec{k}$ one-phonon maxima. In order for $v(\vec{k})$ to have any influence in Eq. (4.1), we need to have $k<2k_F$ (otherwise χ_1 is exceedingly small) and $S(\vec{k})$ to be not too small. Figure 14 emphasizes this point: Here we plot $W-S_e(\vec{k})$, so we subtract *all*

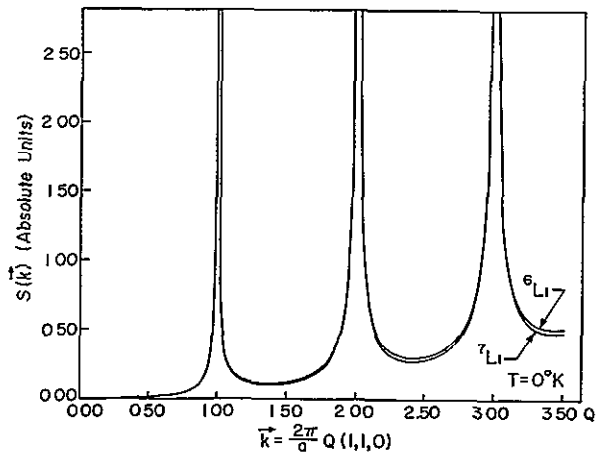


FIG. 6. Structure factor $S(\vec{k})$ for ^{6}Li and ^{7}Li at $T=0$ °K along [110].

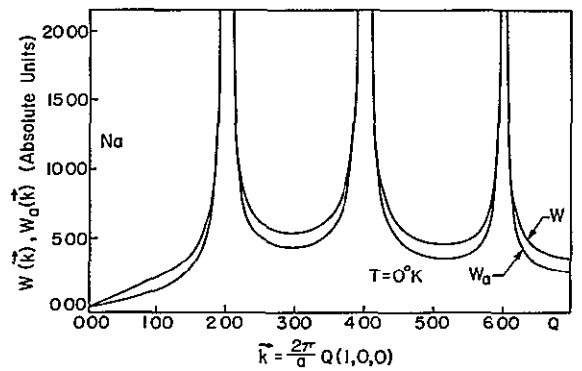


FIG. 8. Cross sections $W(\vec{k})$ and $W_a(\vec{k})$ for Na at $T=0$ °K along [100].

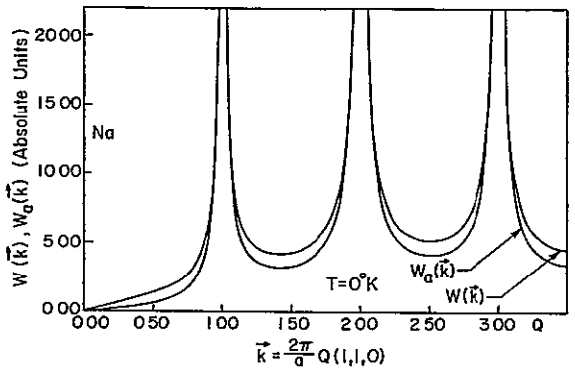


FIG 9 Cross sections $W(\vec{k})$ and $W_a(\vec{k})$ for Na at $T=0$ °K along [110].

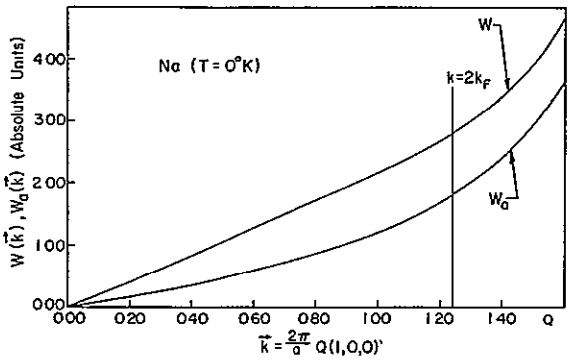


FIG. 11. Cross sections $W(k)$ and $W_a(k)$ for Na at $T=0$ °K along [100] Note the expanded vertical and horizontal scales, and the position of $k=2k_F$.

the Compton scattering. What remains shows a marked dependence on the pseudopotential.

We should discuss the relative composition of the TDS [i.e., of $S(\vec{k})$]. Figure 1 shows the contribution of the one-phonon term, and we see that at large $\vec{k}$ the many-phonon terms become quite important. From Eqs. (3.11) and (3.13), we have[38]

$$S_M(\vec{k}) = e^{-k_\alpha k_\beta \Lambda_{\alpha\beta}(0)/2}\left[e^{k_\alpha k_\beta \Lambda_{\alpha\beta}(0)/2} - 1 - \tfrac{1}{2}k_\alpha k_\beta \Lambda_{\alpha\beta}(0)\right] + \sum_{i\neq 0} e^{-i\vec{k}\cdot\vec{X}_i} e^{-k_\alpha k_\beta \Lambda_{\alpha\beta}(0)/2} \times \left[e^{k_\alpha k_\beta \Lambda_{\alpha\beta}(\vec{X}_i)/2} - 1 - \tfrac{1}{2}k_\alpha k_\beta \Lambda_{\alpha\beta}(\vec{X}_i)\right] \quad (4.4)$$

From the Appendix we also note that for Na, $\mathrm{Tr}\Lambda_{\alpha\beta}(\vec{X}_i) \ll \mathrm{Tr}\Lambda_{\alpha\beta}(0)$ (for $\vec{X}_i \neq 0$) Typically at least 90% of $S_M(\vec{k})$ in Na comes from the first term in Eq (4.4), i e.,

$$S_M(\vec{k}) \approx 1 - e^{-k_\alpha k_\beta \Lambda_{\alpha\beta}(0)/2}\left[1 + \tfrac{1}{2}k_\alpha k_\beta \Lambda_{\alpha\beta}(0)\right] \quad (4.5)$$

In Eq (4.5) we have confirmed a well-known approximation (Eldridge and Lomer[32]).

In spite of the fact that the $\vec{X}_i$ sum in $S_M(\vec{k})$ converges roughly as $\sum_i(\vec{X}_i)^{-4}$, we have found it adequate to take only nine shells (136 vectors) in the sum. [Taking only seven shells changes $S(\vec{k})$ for Na by considerably less than 1%, for example.] This can be understood by noting that

$$\mathrm{Tr}\Lambda_{\alpha\beta}(\vec{X}_9) \ll \mathrm{Tr}\Lambda_{\alpha\beta}(\vec{X}_1) \ll \mathrm{Tr}\Lambda_{\alpha\beta}(0), \quad (4.6)$$

where $\vec{X}_1$ and $\vec{X}_9$ are typical vectors in the first and ninth shells. The point is that the asymptotic limit of $\Lambda_{\alpha\beta}(\vec{X}_i)$ ($\propto 1/X_i^2$) is only reached at large $\vec{X}_i$ where the structure factor is almost independent of the contribution of the remaining shells. In addition, the $\vec{X}_i$ sum actually converges more quickly than $\sum_i 1/X_i^4$ since the term $e^{-i\vec{k}\cdot\vec{X}_i}$ in Eq. (4.4) introduces (except for $\vec{k}=\vec{K}$) considerable self-cancellation

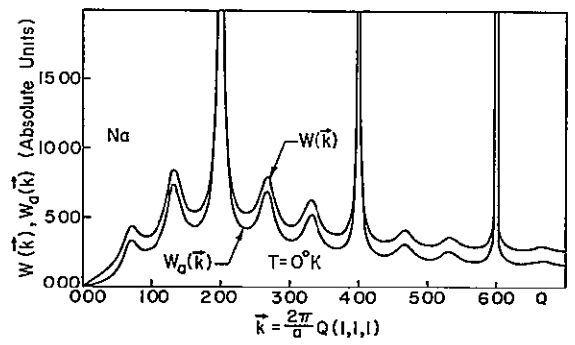


FIG 10 Cross sections $W(\vec{k})$ and $W_a(\vec{k})$ for Na at $T=0$ °K along [111].

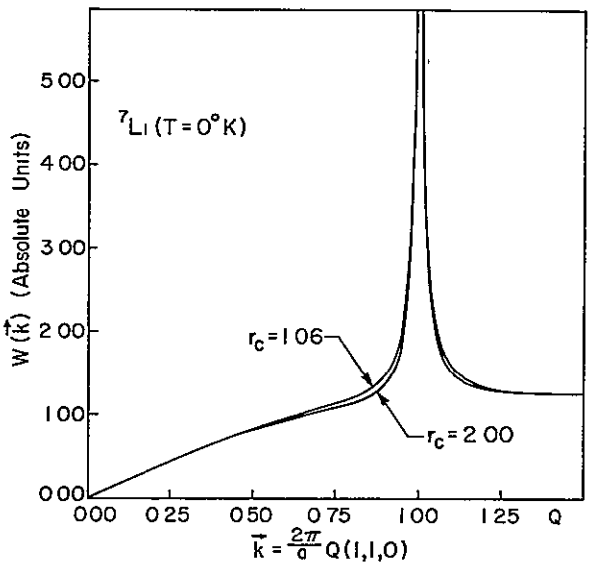


FIG. 12 Cross section $W(\vec{k})$ for ^{7}Li at $T=0$ °K along [110] for two different values of the core radius, $r_c=1.06$ and $r_c=2.00$. Note the expanded horizontal scale.

V. DISCUSSION

The extension of our method of calculation of the ionic structure factor to systems without cubic symmetry and to systems with a basis is completely straightforward. (Special-points-have-been found[9] for systems of hexagonal symmetry, and they can be generated for systems of any symmetry.) The occurrence of one-phonon maxima is equally general. The ability to calculate the $\Lambda_{\alpha\beta}(\vec{X}_s)$ by a procedure which avoids a difficult three-dimensional numerical integration should prove valuable in a variety of contexts, including, for example, the self-consistent harmonic theory of phonons[17] and the computation of static lattice Green's functions.[39]

Much of the theory of x-ray scattering from simple metals presented in Sec. II can be extended to liquid metals. Egelstaff, March, and McGill[40] have derived a formula for the x-ray cross section in liquid metals that is identical to Eq. (2 6), except that they do not make the adiabatic approximation in the terms involving the correlation of conduction electrons with the ions. Making that approximation, and introducing the pseudopotential $v(\vec{k})$, we conclude that Eq. (2 12) is as valid for liquid metals as it is for crystals.

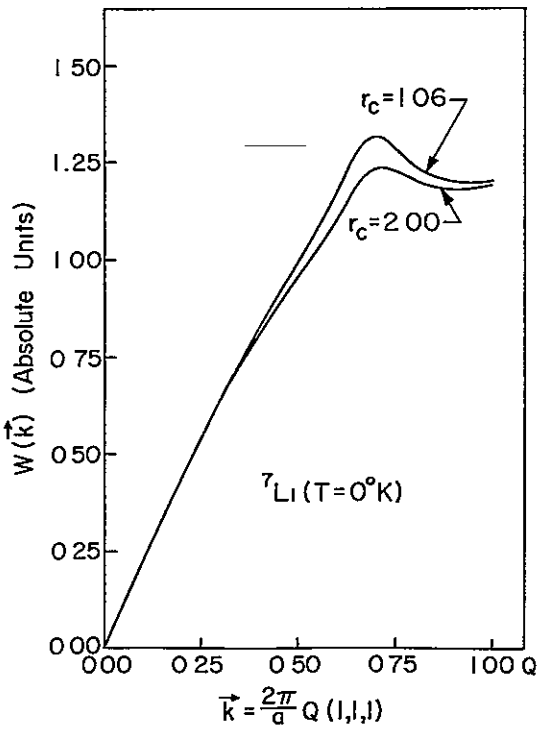


FIG. 13 Cross section $W(\vec{k})$ for ^{7}Li at $T=0$ °K along [111] for two different values of the core radius, $r_c=1.06$ and $r_c=2.00$. Note the expanded vertical and horizontal scales

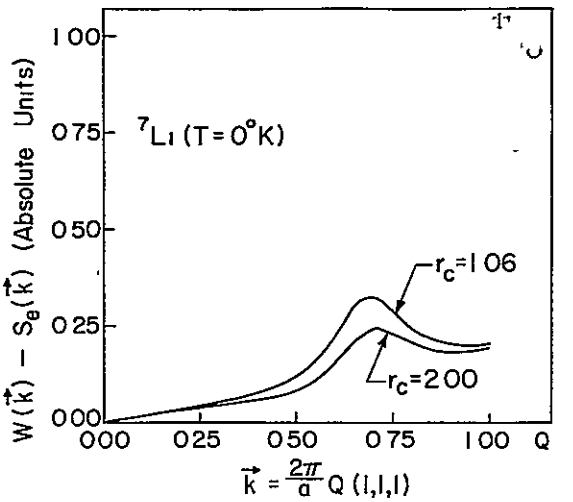


FIG. 14. Cross section with *all* Compton scattering subtracted, $W(\vec{k})-S_e(\vec{k})$, for ^{7}Li at $T=0$ °K along [111], for two different values of the core radius, $r_c=1$ 06 and $r_c=2$ 00 Note the expanded vertical and horizontal scales.

Finally, our calculation has neglected possible anharmonic effects. Those anharmonic terms which are retained in the self-consistent phonon theory[17] are in a sense taken into account here The formalism we have presented is not altered by using the self-consistent theory, but the frequencies are changed from their harmonic values

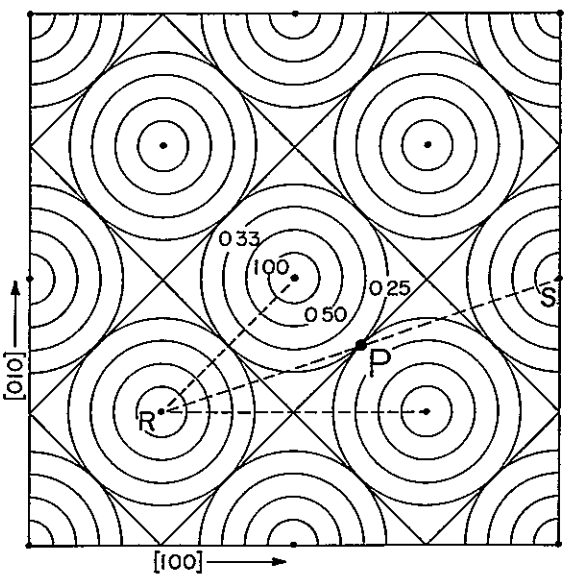


FIG. 15. Lines of equal value of the function $1/c|\vec{q}(\vec{k})|$ in a (001) plane of the lattice reciprocal to the bcc lattice. R is the point $(2\pi/a)(0,0,0)$, P the point $(2\pi/a)(\frac{3}{2},\frac{1}{2},0)$, and S is the point $(2\pi/a)(3,1,0)$, where a is the lattice constant. The numbers 1.00, 0 50, 0.33, and 0.25 indicate the relative value of the function.

In the case of sodium, this change is small.[23] Other anharmonic effects are not taken into account For example, the interference between one- and two-phonon scattering can cause a noticeable change[7,41] in $S_{ion}(k)$ As shown by Glyde,[7] however, it amounts to only a small shift in the one-phonon scattering for Na at high temperatures Since both the anharmonic frequency shifts and the inverse phonon lifetimes become quite small at low temperatures,[23] the size of this contribution should decrease correspondingly Interference effects, as well as other effects due to anharmonicity, may of course be of somewhat greater importance in the case of lithium

APPENDIX

We briefly review the special point method,[8,9] which was designed for the integration of quantities varying *slowly* over the first Brillouin zone Here, by a slight modification, we use it to evaluate the integral of oscillating functions [see Eq. (3.7)].

The general integral to be evaluated is

$$\frac{1}{N}\sum_{\vec{q}}^{BZ} f(\vec{q}) = \frac{\Omega_c}{(2\pi)^3}\int_{BZ} d^3q\, f(\vec{q}), \tag{A1}$$

where $f(\vec{q})$ is assumed to be invariant under the operations of the crystal point group, and Ω_c is the primitive cell volume. [If $f(\vec{q})$ is not symmetric, it can, of course, be easily symmetrized.] One expands $f(\vec{q})$ in symmetrized plane waves $A_m(\vec{q})$:

$$f(\vec{q}) = f_0 + \sum_{m=1}^{\infty} f_m A_m(\vec{q}), \tag{A2}$$

with

$$A_m(\vec{q}) = \sum_{\vec{X},m} e^{i\vec{q}\cdot\vec{X}} \tag{A3}$$

and

$$f_m = \frac{1}{N_m}\frac{\Omega_c}{(2\pi)^3}\int_{BZ} d^3q\, f(\vec{q}) A_m(\vec{q}). \tag{A4}$$

$\vec{X}; m$ refers to all lattice vectors $\vec{X}$ with the same length X_m that are related by point group operations. N_m is the number of vectors in this mth shell, and the sum in Eq. (A2) is ordered so that those shells with lowest X_m come first

A set $\{\vec{q}_i\}$ of special points is defined as a set of n points in the BZ with associated weights α_i which satisfy

$$\sum_{i=1}^{n} \alpha_i A_m(\vec{q}_i) = 0 \quad \text{for } m = 1, \ldots, N, \tag{A5}$$

$$\sum_{i=1}^{n} \alpha_i = 1. \tag{A6}$$

Using Eqs. (A5) and (A6) in Eq. (A2),

$$f_0 = \sum_{i=1}^{n} \alpha_i f(\vec{q}_i) - \sum_{i=1}^{n} \alpha_i A_{N+1}(\vec{q}_i) f_{N+1} + \cdots \tag{A7}$$

Since f_0 is the desired integral, Eq. (A7) gives an approximation to the integral consisting of an evaluation of $f(\vec{q})$ at a (small) set of points The first neglected term can be shown to be $\pm f_{N+1}$. *Not all* coefficients f_m for $m > N$ have been neglected, as Eq. (A5) is always satisfied for an infinite number of shells. The index of the first shell for which Eq. (A5) is *not* satisfied is $N+1$. With increasing number of points n in the set, both the number and the magnitude of the neglected terms become smaller.

At $T = 0$°K, $\mathrm{Tr}\Lambda_{\alpha\beta}(0) \propto \sum_{\vec{q},j} 1/\omega(j\vec{q})$ is a smooth function, and we may apply the special point method. Although the expansion coefficients f_m decrease slowly with increasing m for large m, they are much smaller than $\mathrm{Tr}\Lambda_{\alpha\beta}(0)$ itself. Thus we expect increasing the number of special points n to have a small effect on $\mathrm{Tr}\Lambda_{\alpha\beta}(0)$. From Table I we see the convergence is more rapid for $T = 0$°K than for $T = 90$°K.

The calculation of $\Lambda_{\alpha\beta}(\vec{X}_i)$, $\vec{X}_i \neq 0$ is more troublesome, and we illustrate by examining the trace of this matrix. Symmetrizing the integrand of $\Lambda_{\alpha\beta}(\vec{X}_i)$:

$$\mathrm{Tr}\Lambda_{\alpha\beta}(\vec{X}_i) \propto \sum_{\vec{q},j}^{BZ} \frac{1}{\omega(j\vec{q})} A_i. \tag{A8}$$

Applying the special-point method to this integral means neglecting some of the coefficients $\bar{f}_m$ whose form is (we are at $T = 0$°K)

$$\bar{f}_m \propto \sum_{\vec{q},j}^{BZ} \frac{1}{\omega(j\vec{q})} A_i A_m. \tag{A9}$$

Now $A_i A_m$ is itself a sum of symmetrized plane

TABLE I $M^0 = \frac{1}{2}(2k_F)^2 \Lambda^0$ (in units of 10^{-2}) $M(\vec{R}) = \frac{1}{2}(2k_F)^2\ \mathrm{Tr}\,\Lambda_{\alpha\beta}(\vec{R})$ (in units of 10^{-2}). N is the number of special points (Na, $T = 0$ °K)

	$N=8$	40	240
$M^0(T=0\,°K)$	3.4367	3 4762	3 4832
$M^0(T=90\,°K)$	7 9897	8 5890	8 8258
$M(\vec{R}=(1,1,1))$	1 126	1 134	1 133
$M(\vec{R}=(2,0,0))$	0 538	0 541	0 540
$M(\vec{R}=(2,2,0))$	0 283	0 261	0 259
$M(\vec{R}=(3,1,1))$	0 240	0 223	0 221
$M(\vec{R}=(2,2,2))$	0.473	0 479	0 477
$M(\vec{R}=(4,0,0))$	0.174	0.167	0 164
$M(\vec{R}=(3,3,1))$	0.169	0.152	0 148
$M(\vec{R}=(4,2,0))$	0 140	0 099	0 095
$M(\vec{R}=(4,2,2))$	0.116	0.137	0 113

waves

$$A_i(\vec{q})A_m(\vec{q}) = \sum_j a_j(i,m)A_j(\vec{q}), \tag{A10}$$

where the first j for which $a_j(i,m) \neq 0$ is that for which $\vec{X}_j = |\vec{X}_m - \vec{X}_i|$. From Eqs. (A8)–(A10) it is clear that the $\bar{f}_m$ for large m will be much less than $\mathrm{Tr}\Lambda_{\alpha\beta}(\vec{X}_i)$ only if the $\bar{f}_m$ themselves decrease rapidly with increasing m. This, however, is not the case, for just as in Eq. (3.16),

$$\lim_{X_i \to \infty} \sum_{\vec{q}}^{BZ} \frac{1}{\omega(j\vec{q})} A_i(\vec{q}) \sim \frac{1}{X_i^2}. \tag{A11}$$

The origin of this behavior is the $1/q$ behavior of $1/\omega(j\vec{q})$ as $\vec{q} \to 0$ (see Ref. 18 and Schober *et al.*, Ref. 39).

To circumvent this difficulty one must find a matrix $M_{\alpha\beta}(\vec{q})$ whose behavior at the origin is the same as that of $\sum_j [1/\omega(j\vec{q})] e_\alpha(j\vec{q}) e_\beta(j\vec{q})$, and which leads to an integral $\int_{BZ} d^3q\, M_{\alpha\beta}(\vec{q}) \cos(\vec{q}\cdot\vec{X})$, which can be evaluated analytically. Then we write

$$\Lambda_{\alpha\beta}(X_i) = \frac{\hbar}{MN} \sum_{\vec{q}}^{BZ} \left[\left(\sum_j \frac{1}{\omega(j\vec{q})} e_\alpha(j\vec{q}) e_\beta(j\vec{q}) \right) - M_{\alpha\beta}(\vec{q}) \right] \cos(\vec{q}\cdot\vec{X}_i) + \frac{\hbar}{MN} \sum_{\vec{q}}^{BZ} M_{\alpha\beta}(\vec{q}) \cos(\vec{q}\cdot\vec{X}_i), \tag{A12}$$

and compute the first integral by the special point method. Since the integrand has no troublesome $1/q$ behavior, its expansion coefficients $\bar{f}_m$ should then decrease rapidly, and the number of special points then needed for an accurate determination of $\Lambda_{\alpha\beta}(\vec{X}_i)$ should be (and is in fact) correspondingly small.

To simplify the calculation, we have actually only treated the *trace* of $\Lambda_{\alpha\beta}(\vec{X}_i)$ in the above fashion, subtracting off a function $M(\vec{q})$ whose behavior as $\vec{q} \to 0$ is approximately that of $\frac{1}{3}\sum_j 1/\omega(j\vec{q})$. {As $\vec{q} \to 0$, $\sum_j 1/\omega(j\vec{q}) \to 1/d(\hat{q})q$, where $d(\hat{q})$ is a function of direction. We have approximated $d(\hat{q})$ with $[\sum_j \int d\Omega_q 1/c_j(\hat{q})]^{-1}$, where the $c_j(\hat{q})$ are the three speeds of sound.} Tables I and II show the elements of $\Lambda_{\alpha\beta}(\vec{X}_i)$, for $\vec{X}_i$ in the first nine shells ($T = 0°K$), $\mathrm{Tr}\Lambda_{\alpha\beta}(\vec{X}_i)$, and $\Lambda_{\alpha\beta}(0)$ for $T = 0°K$ and $T = 90°K$. Three different (bcc) special point sets were used, with $n = 8$, 40, and 240. Although one can only expect $\mathrm{Tr}\Lambda_{\alpha\beta}(\vec{X}_i)$ to converge well, the individual matrix elements also show good convergence.

TABLE II $M_{\alpha\beta}(\vec{R}) = \frac{1}{2}(2k_F)^2 \Lambda_{\alpha\beta}(\vec{R})$ (in units of 10^{-3}) N is the number of special points (Na, $T = 0°K$)

	N	M_{xx}	M_{xy}	M_{yy}	M_{xz}	M_{yz}	M_{zz}
$\vec{R}=(1,1,1)$	8	3.754	2 610	3.754	2 610	2 610	3 754
	40	3 780	2.664	3.780	2 664	2.664	3 780
	240	3 778	2.666	3 778	2 666	2 666	3.778
$\vec{R}=(2,0,0)$	8	0 822	0	2 278	0	0	2 278
	40	0 716	0	2 345	0	0	2 345
	240	0 708	0	2 345	0	0	2 345
$\vec{R}=(2,2,0)$	8	1 278	0 698	1 278	0	0	0 270
	40	1 215	0 740	1 215	0	0	0 184
	240	1.207	0 744	1.207	0	0	0 181
$\vec{R}=(3,1,1)$	8	0.715	0 225	0 842	0 225	0 316	0.842
	40	0 557	0 230	0 836	0 230	0 444	0.836
	240	0 541	0.233	0.832	0 233	0 448	0.832
$\vec{R}=(2,2,2)$	8	1.578	1 039	1.578	1 039	1 039	1 578
	40	1 598	1 133	1 598	1.133	1.133	1.598
	240	1 589	1 139	1 589	1 139	1 139	1 589
$\vec{R}=(4,0,0)$	8	0 581	0	0 581	0	0	0 581
	40	0 212	0	0 727	0	0	0 727
	240	0 186	0	0 730	0	0	0 730
$\vec{R}=(3,3,1)$	8	0 680	0 528	0.680	0 073	0 073	0 331
	40	0.668	0.458	0.668	0 120	0 120	0.188
	240	0.653	0.464	0.653	0 125	0 125	0.179
$\vec{R}=(4,2,0)$	8	0 465	0 275	0.465	0	0	0 465
	40	0 356	0.167	0 400	0	0	0 234
	240	0 331	0.171	0 391	0	0	0 227
$\vec{R}=(4,2,2)$	8	0.388	0 304	0 388	0.304	0	0 388
	40	0.386	0 208	0 491	0 208	0 303	0 491
	240	0.363	0 214	0.482	0.214	0.314	0 482

*Work supported by the NSF Grant No DMR74-23494, NASA Contract No. NGR-33-010-188, and the Materials Science Center Grant No DMR72-03029

[1]R. Colella and B. W. Batterman, Phys. Rev B 1, 3913 (1970).
[2]C B. Walker, Phys. Rev 103, 547 (1956)
[3]S. L Schuster and J W. Weymouth, Phys. Rev. B 3, 4143 (1971).
[4]A. A. Maradudin, E. W. Montroll, G. H Weiss, and I. P. Ipatova, in *Solid State Physics*, 2nd ed., edited by H. Ehrenreich, F Seitz, and D Turnbull (Academic, New York, 1971), Suppl. 3
[5]G. Albanese and C. Ghezzi, Phys Rev. B 8, 1315 (1973).
[6]Y. Kashiwase and J. Harada, J. Phys. Soc. Jpn. 35, 1711 (1973)
[7]H R. Glyde, Can. J. Phys 52, 2281 (1974).
[8]A. Baldereschi, Phys. Rev. B 7, 5212 (1973).
[9]D. J. Chadi and Marvin L. Cohen, Phys Rev. B 8, 5747 (1973)
[10]A. Sjolander, in *Phonons and Phonon Interactions*, edited by T A. Bak (Benjamin, New York, 1964)
[11]P Nozières and D. Pines, Phys. Rev. 113, 1254 (1959).
[12]L. Van Hove, Phys. Rev. 95, 249 (1954)
[13]D Pines and P. Nozières, *The Theory of Quantum Liquids* (Benjamin, New York, 1966), Vol. I.
[14]For a discussion of the one-phonon cross section when this is no longer the case, see P. M Platzman and N. Tzoar, Phys. Rev. B 7, 2450 (1973) This calculation is not restricted to weak pseudopotentials, but agrees with the one-phonon contribution to Eq. (2 6) in the appropriate limit.
[15]C. B. Walker, Phys. Rev 103, 558 (1956)
[16]For a liquid metal, the free-electron structure factor is a rough approximation to the true one. See J. Chihara, in *Proceedings of Second International Conference on the Properties of Liquid Metals, Tokyo*, edited by S. Takeuchi (Taylor and Francis, London, 1973), p. 137.
[17]P. Choquard, *The Anharmonic Crystal* (Benjamin, New York, 1971)
[18]G. A. Wolfe and B Goodman, Phys. Rev. 178, 1171 (1969).
[19]T R. Lomer, Proc Phys. Soc Lond 89, 135 (1966)
[20]S. V Semenovskaya and Ya S. Umanskii, Fiz. Tverd. Tela 6, 2963 (1964) [Sov. Phys.-Solid State 6, 2362 (1965)].
[21]J. S. Reid and T. Smith, J. Phys C 3, 1513 (1970).
[22]A D B Woods, B. N. Brockhouse, R. H March, A T. Stewart and R. Bowers, Phys Rev. 128, 1112 (1962)
[23]H. R. Glyde and R. Taylor, Phys Rev. B 5, 1206 (1972)
[24]We neglect thermal expansion.
[25]By applying the special point method carefully (see the Appendix), it is, of course, straightforward to calculate the $\Lambda_{\alpha\beta}(\vec{X}_i)$ at any temperature
[26]R W G Wyckoff, *Crystal Structures*, 2nd ed. (Interscience, New York, 1963), Vol. I, p. 16
[27]H G Smith, G. Dolling, R. M. Nicklow, P R Vijayaraghavan, and M. K Wilkinson, in *Neutron Inelastic Scattering* (IAEA, Vienna, 1968), Vol I, p. 149.
[28]*International Tables for X Ray Crystallography* (Kynoch, Birmingham, England, 1962), Vol III, p. 202.
[29]D. J. W. Geldart and S. H. Vosko, Can J Phys 44, 2137 (1966)
[30]N W. Ashcroft, Phys. Lett 23, 48 (1966); J. Phys. C 1, 232 (1968).
[31]Numerical results are available from the authors upon request
[32]For other examples of this behavior see Ref 33 and Y Kashiwase, J. Phys. Soc. Jpn. 34, 1303 (1973); J. E. Eldridge and T. R. Lomer, Proc Phys Soc. Lond 91, 459 (1967). The "extra spots" in the x-ray photographs of Ref. 33 correspond to this one-phonon structure.
[33]M Born, Rept. Prog. Phys 9, 294 (1942-43)
[34]All k vectors are quoted in units of $2\pi/a$, where a is the lattice constant
[35]R. Comès, in *Lecture Notes in Physics One-Dimensional Conductors*, edited by J Ehlers, K. Hepp, and H A Weidenmuller (Springer-Verlag, Berlin, 1975), Vol. 34
[36]Platzman and Tzoar, in Ref 14 precisely stress this point.
[37]N. W Ashcroft and David C Langreth, Phys Rev 155, 682 (1967).
[38]Reference 5 similarly divides the multiphonon scattering.
[39]In this case we do not need $\Lambda_{\alpha\beta}(\vec{X}_i)$ but a very similar integral. See H. R. Schober, M Mostoller, and P. H. Dederichs, Phys. Status Solidi B 64, 173 (1974).
[40]P A Egelstaff, N. H March and N C. McGill, Can. J. Phys. 52, 1651 (1974).
[41]W. J. L Buyers and T. Smith, Phys Rev 150, 758 (1966); R. A Cowley and W J. L Buyers, J. Phys. C 2, 2262 (1969), R. A. Cowley, E. C. Svensson, and W. J L. Buyers, Phys. Rev Lett 23, 325 (1969)

Aluminum under high pressure. I. Equation of state*

Carlos Friedli and N. W Ashcroft
Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853
(Received 14 July 1975)

A curve of applied pressure P versus lattice constant a is calculated for single-crystal aluminum. It results from an application of the method of structural expansions for deriving the energies of simple metals, a method known to give reasonable results for the elastic constants even at second order in the effective electron-ion interaction The latter (in the present calculation) is taken from Fermi-surface analysis and it is verified (with this essentially experimental information) that the extant face-centered cubic structure remains the preferred crystalline phase up to the highest pressures considered. Arguments are given to suggest that the P versus a curve should have reasonable *a priori* accuracy, and can admit of possible improvement if experimental data in the intermediate-presssure region can be provided to refine the (in principle) energy-dependent pseudopotential. At three megabars the lattice constant is reduced by only 22%; the ion cores at this pressure are still very well separated.

I INTRODUCTION

Among the simple metals, aluminum is in many ways one of the simplest, being cubic close packed under normal conditions and possessing ion cores occupied by electrons in levels of s and p symmetry. It is mainly a consequence of the latter that its nearly-free-electron band structure can be interpolated so accurately by a spatially local pseudopotential, a feature which distinguishes it somewhat from the alkali metals. Although the Fermi surfaces of the alkali metals are a good deal simpler than that of aluminum, the apparent complexity of its multiply-connected Fermi surface can be used to advantage in a study of the transport properties at high pressure. This will be the content of a later work; for the present we are concerned with the equation of state of Al, a necessary preliminary in discussing the dependence of transport properties on pressure.[1] Effects of temperature (for normal conditions) are quite small, and our aim here is therefore to express the equation of state in terms of pressure versus lattice constant. Such a relation can only be considered potentially useful if no crystalline phase changes are likely to occur.[2] We show by a series of arguments that the common face-centered cubic phase of Al appears to remain the stable phase for pressures exceeding 3Mbar. In terms of the lattice constant (or equivalently the r_s electron spacing parameter) these colossal pressures represent a rather modest change of around 20%. The electron density is increased, but not greatly. It is not unreasonable to suppose, therefore, that the method based on structural expansions about the uniform interacting electron gas will continue to function as it does for the system taken at more reasonable pressures The method is summarized in Sec. II, and in the course of discussing the standard second-order theory[3] we comment on the importance of higher-order corrections to the present calculations.

Section III describes the application of the formalism to the problem of deciding which of several possible simple structures (including fcc) will possess the lowest Gibbs energy. For the fcc phase, a curve of pressure versus lattice constant a is presented (Sec. IV); up to and above 3 Mbar, the changes in a are quite monotonic. Up to about 800 kbar, our calcuations, based on the method of structural expansions in a weak pseudopotential, can be compared directly with the results of Ross and Johnson[4] who obtain the equation of state of aluminum from an *a priori* calculation of the band structure by the augmented-plane-wave (APW) method.

We estimate that not until pressures of over 100 Mbar are reached will the ion cores of Al be substantially contiguous. This is a very different situation from the one prevailing in ionic crystals where the pressure scale is founded largely on assumed short-range interactions.[5] Although the atomic number of Al is relatively low, it may compete reasonably well in x-ray scattering power with NaCl and may, therefore, be an alternative candidate for calibration and use as a pressure scale.

II. ENERGY OF SIMPLE METALS

On account of the compactness of its ion core (and the absence of filled d-shell levels) the pseudopotential in Al, although energy dependent to a small degree[6] is remarkably local and provides an excellent interpolation to *a priori* band structures. Invoking an adiabatic approximation.

we shall take it that an ion of the dynamic lattice of Al carries with it a bare pseudopotential, $v(\vec{k})$ known (at the Fermi energy) from Fermi surface analysis.[7] It is a function which as is well-known oscillates in sign as k increases, a fact which reflects the finite size of the Al ion core. Since we shall shortly need to consider the possibility of corrections arising from dynamic lattice effects, it is convenient to set down a Hamiltonian for the *electron* system that is written[8] for instantaneous positions $\vec{r}(\vec{R})$ of the ions near equilibrium sites $\vec{R}$, i.e.,

$$H = H_{ee} + H_M + H_{ei}, \tag{1}$$

where for the present H_{ee} can be taken as the standard Hamiltonian for the interacting electron gas (uniform compensating positive background) and the ionic Hamiltonian H_M leads to the Madelung energy NE_M of point ions. In rydbergs it can be written (for ZN electrons)

$$ZNE_M = \sum_{k\neq 0} \frac{8\pi}{k^2}[S(\vec{k}) - 1], \tag{2}$$

where for the ions in a volume V the structure factor for the ionic system is

$$S(\vec{k}) = (1/N)\langle \hat{\rho}_{\vec{k}} \hat{\rho}_{-\vec{k}} \rangle - N\delta_{\vec{k},0}\ , \tag{3}$$

with

$$\rho_{\vec{k}} = \sum_{\vec{R}} e^{i\vec{k}\cdot\vec{r}(\vec{R})},$$

and the average in (3) being taken over the states of the crystal. The final term in (1), H_{ei}, is the electron ion interaction in which it is convenient to include the largely compensating zeroth Fourier component of all the long-range interactions; that is, a term E_0 which although independent of structure is always difficult to calculate from first principles. It can, however, be eliminated by exploiting a fragment of experimental information such as, for example, the equilibrium density.[9]

Accordingly we write

$$H_{ei} = E_0 + \sum_{\vec{k}\neq 0} \hat{\rho}_{\vec{k}} v(\vec{k}) \hat{\rho}^s_{-\vec{k}}, \tag{4}$$

where for the electrons the density operator is written

$$\hat{\rho}^s_{-\vec{k}} = \sum_i e^{-i\vec{k}\cdot\vec{r}_i} \tag{5}$$

We turn first to the static lattice case for which the contribution of E_M to the thermodynamic function is known, at least for most simple structures. The problem of calculating the energy of a simple metal then reduces to an expansion (relative to the structureless electron gas system) in orders (beginning at the second) of H_{ei}. Since the ionic correlation function [for example, $S(\vec{k})$] are then δ functions on the reciprocal lattice they reduce the resulting summations in the perturbation series to lattice sums. Thus, in addition to the ground-state energy[10] from H_{ee} (and E_M) we have, as the first term of the structural expansion, a second-order band-structure contribution $E^{(2)}_{BS}$ of the form

$$E^{(2)}_{BS} = \tfrac{1}{2} \sum_{[\vec{K}]}{}' |v(\vec{K})|^2 \chi^{(1)}(\vec{K})/\epsilon(\vec{K}); \tag{6}$$

$[\vec{K}]$ reciprocal-lattice set,

where $\epsilon(\vec{k})$ is the dielectric function of the interacting electron gas and $\chi^{(1)}(\vec{k})$ its (static) first-order polarizability. At this level of approximation the internal energy is then

$$E = (E_{ee} + E_M + E_0) + E^{(2)}, \tag{7}$$

and it is interesting, before proceeding further, to examine their relative contributions to the pressure at a given volume V, or what is equivalent, a mean electron spacing r_s [$V/NZ = (r_s a_0)^3 \frac{4}{3}\pi$]. Table I shows[11] that as pressure increases the contribution from $E^{(2)}$ becomes progressively a smaller fraction of the total. Since we know[9] the ground-state energy and compressibility of Al to be quite well given near $P = 0$ by (7) and its derivatives, we may conclude that even at high pressures the higher-order band-structure contributions to E are not likely to be an important factor in limiting the accuracy of a calculation of P vs a. The most significant of these corrections is the third-order band-structure energy. If the electron gas is treated, for example, within the random-phase approximation, this term can be written[3,12]

$$\frac{1}{3} \sum_{\vec{K}\vec{K}'} \frac{v(\vec{K})}{\epsilon(\vec{K})} \frac{v(\vec{K}')}{\epsilon(\vec{K}')} \frac{v(\vec{K}-\vec{K}')}{\epsilon(\vec{K}-\vec{K}')} \chi^{(2)}(\vec{K}, \vec{K}', \vec{K}-\vec{K}'), \tag{8}$$

where $\chi^{(2)}$ is the second-order polarizability of the

TABLE I. The quantities E_{ee}, E_M, and E_0 are present at any order of the calculation and are convenient to group together in the comparison of the relative pressure contributions. The first column gives an estimate of the pressure (in Mbars) from $E_{ee} + E_M - E_0$ and the second column for $E^{(2)}_{BS}$. Energies are given in rydbergs.

r_s	$P\ (E_{ee}+E_M-E_0)$	$P\ (E^{(2)}_{BS})$
2.07	0.48 (−1.29)	−0.48 (−0.097)
1.9	1.39 (−1.24)	−1.07 (−0.138)
1.8	2.38 (−1.189)	−1.62 (−0.176)
1.7	3.95 (−1.110)	−2.37 (−0.227)
1.6	6.47 (−0.993)	−3.35 (−0.292)

electron system. As remarked earlier, $v(\vec{K})$ for Al (and indeed any non-point-ion system) alternates in sign as its argument increases and, as a consequence, there is substantial self cancellation in (8). Furthermore, relative to $\mathcal{E}_F$, the $|v(\vec{K})|$ are considerably less than ~0.1 (for example, $|v_{111}/\mathcal{E}_F| = 0.0209$, and $|v_{200}/\mathcal{E}_F| = 0.0657$). It follows that the higher-order band-structure energies are quite small in comparison with $E_{BS}^{(2)}$. This has already been noticed by others,[13] although we must recognize that the derivatives of the higher-order terms (in the elastic constants for example) need not always be unimportant.

As far as a calculation of the pressure is concerned it seems a reasonable approximation to neglect the higher-order band-structure energies. The approximation would appear less justifiable in the calculation of the ground-state energy for various crystal structures. But in fact it remains numerically valid. The concern is that differences in Gibbs energy for different crystal structures are quite small, about 4–6 mRy between hcp and fcc per electron if calculated with a second-order expression. And these can be less than typical third-order energies. However, we need not the *absolute* third-order energies, but their *differences* for different structures; these are in turn smaller by about an order of magnitude. We shall see in a moment that inclusion of dynamic effects are likely to reduce the third-order differences still further, so that a calculation of the energy at second order is sufficient for the present purposes.

Relaxing the static lattice assumption requires (a) the inclusion of phonon energy term, if indeed the excitations are to be described by phonons, and (b) the reintroduction in (6) and (7) of the corresponding ionic correlation functions, for example, $S(\vec{k})$ [Eq. (3)]. If $\vec{u}(\vec{R})$ is the displacement of an ion from site $\vec{R}$, then

$$S(\vec{k}) = \frac{1}{N}\sum_{\vec{R}\vec{R}'} e^{i\vec{k}\cdot(\vec{R}-\vec{R}')}\langle e^{i\vec{k}\cdot\vec{u}(\vec{R})} e^{-i\vec{k}\cdot\vec{u}(\vec{R}')}\rangle, \tag{9}$$

and if the $\vec{u}(\vec{R})$ may be developed as a linear synthesis of phonon operators, it follows that[14]

$$S(\vec{k}) = \frac{1}{N}\sum_{\vec{R}\vec{R}'} e^{i\vec{k}\cdot(\vec{R}-\vec{R}')} \exp\tfrac{1}{2}\{-[\vec{k}\cdot\vec{u}(\vec{R})]^2 - [\vec{k}\cdot\vec{u}(\vec{R}')]^2 + [\vec{k}\cdot\vec{u}(\vec{R})][\vec{k}\cdot\vec{u}(\vec{R}')]\}, \tag{10}$$

and this replaces the sequence of δ functions which led to the lattice sum in the second-order term (6). The correlation function corresponding to (9) and appearing in the third-order expression is easily seen to be of the form

$$\sum_{\vec{R}\vec{R}'\vec{R}''} e^{i\vec{k}\cdot\vec{R}} e^{i\vec{q}\cdot\vec{R}'} e^{-i(\vec{q}+\vec{k})\cdot\vec{R}''} \exp\{-\tfrac{1}{2}\langle [\vec{k}\cdot\vec{u}(\vec{R})]^2 + [\vec{q}\cdot\vec{u}(\vec{R}')]^2 + [(\vec{q}+\vec{k})\cdot\vec{u}(\vec{R}'')]^2 + 2[\vec{k}\cdot\vec{u}(\vec{R})\vec{q}\cdot\vec{u}(\vec{R}')] - 2[\vec{q}\cdot\vec{u}(\vec{R}')(\vec{k}+\vec{q})\cdot\vec{u}(\vec{R}'')] - 2[\vec{k}\cdot\vec{u}(\vec{R})(\vec{k}+\vec{q})\cdot\vec{u}(\vec{R}'')]\rangle\}, \tag{11}$$

which is straightforward to generalize to higher orders.

For metals with substantial Debye temperatures (in which category we may place Al) one method of handling (10) and (11) is to proceed by a multiphonon expansion. The zero-phonon term leads immediately back to (6) and (8). The one-phonon term leads, when combined with the kinetic energy of the phonon system,[15] to the internal energy of the phonons. The remaining multiphonon terms, as is known from the analysis of thermal diffuse x-ray scattering are quite small. Thus we may, with a sufficient accuracy, treat the phonons independently of the electron system and calculate the Gibbs energy of the latter assuming a rigid lattice. The internal energy can then be written

$$E = (E_{es} - E_0 + E_M) + E_{BS}^{(2)} + E^{ph}, \tag{12}$$

where E^{ph} is the internal energy of the phonon system.

III STRUCTURAL CONSIDERATIONS

From the known Fermi surface of Al (and the assumption of a static lattice) the values of $v(K)$, $\vec{K} \equiv (1,1,1,),(2,0,0)$ can be extracted and these can be interpolated and extended by an empty-core pseudopotential $[v(k) = (-8\pi Z/k^2)\cos kr_c]$. The range of validity (in k) of such a simple form is quite sufficient to assure convergence of the sums in (6), and hence of the band-structure energy. Since $v(k)$ is a property of the ion we may repeat the procedure at any chosen volume or density. Assuming for the moment that this is fixed we must examine the structure-dependent terms in (12) as the ions are rearranged in a variety of possible

crystal structures.

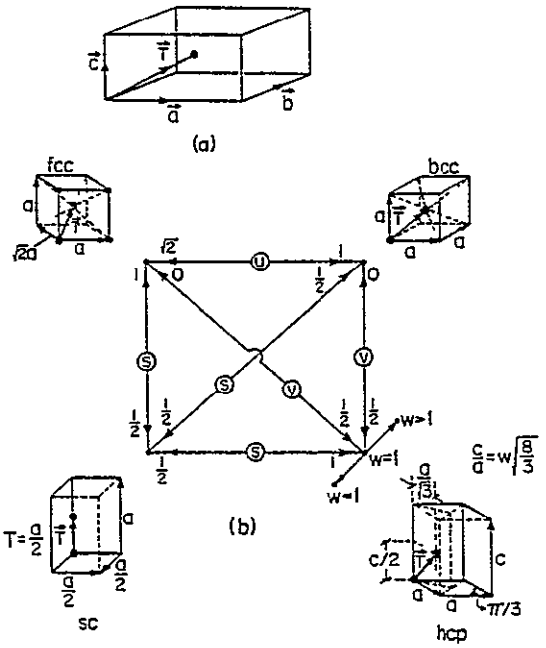
To begin with we consider the electronic terms (and Madelung energy) and allow ourselves at this point the freedom of a structure with a two atom basis. The task is to ascertain which of the structures (at least, which of the simple structures) is preferred for Al: to this end we will select carefully a system of primitive and basis vectors which will allow us continually to deform between different structures by means of a smooth variation of parameters.[16] Refer now to Fig. 1(a). We take $\vec{a}$, $\vec{b}$, and $\vec{c}$ as primitive vectors which are written in the form

$$\vec{a}=a(s,0,0),\quad \vec{b}=a(v',\xi,0),\quad \vec{c}=a(0,0,\eta). \tag{13}$$

Direct lattice vectors are then written

$$\vec{R}=n\vec{a}+p\vec{b}+q\vec{c}.$$

We take the basis vectors



	s	u	v	w
sc	$\frac{1}{2}$	arbitrary	arbitrary	arbitrary
bcc	1	1	0	1
fcc	1	$\sqrt{2}$	0	1
ideal hcp	1	arbitrary	$\frac{1}{2}$	1
hcp	1	arbitrary	1/2	$\neq 1$

(c)

FIG. 1. (a) General structure defined. (b) Some particular cases and representations of continuous one-parameter transformations of them into each other. (c) Values of the parameters for these particular cases. The parameters are defined by Eqs. (13)–(15).

$$\vec{b}_1=0,\quad \vec{b}_2=\vec{T}=\tfrac{1}{2}a(2s-1,(2s-1)\zeta,\eta). \tag{14}$$

In (13) and (14) the parameters v', ζ, η, and ξ are chosen in the following way:

$$\begin{aligned} v'&=(2s-1)v,\\ \xi&=u-v(2u-\sqrt{3})+(1-s)[1-2u+2v(2u-\sqrt{3})],\\ \eta&=w+2vw(\sqrt{\tfrac{8}{3}}-1)+2(1-s)[1-w-2vw(\sqrt{\tfrac{8}{3}}-1)],\\ \zeta&=u-2v(u-1/\sqrt{3}), \end{aligned} \tag{15}$$

with s, u, v, w taken as independent parameters. Transformation (15) is only one of many ways of continually deforming the standard simple crystal structures. We have selected it because it permits us to examine single-cubic (sc), face-centered cubic (fcc), body-centered cubic (bcc), and hexagonal closed packed (hcp) with variable (c/a) ratio. As an example, note that when $s=\frac{1}{2}$ we have (whatever finite values u, v, w may assume) a simple-cubic structure. On the other hand, if $s=1$, $v=0$, and $w=1$, the structure is fcc for $u=\sqrt{2}$, and bcc with $u=1$. Further, if $s=1$, $v=\frac{1}{2}$, and $w=1$, we have hcp with ideal ratio. These are summarized on Figs. 1(b) and 1(c). Although it cannot be deduced simply from the results we shall give, it is interesting to note that the transformation we have chosen moves the atoms in a very natural way, keeping them well apart, and proceeding as directly as possible from one structure to another. In a sense we are moving the atoms along valleys in the energy-structure space.

The lattice reciprocal to (13) is spanned by primitive vectors

$$\begin{aligned} \vec{A}&=(2\pi/a)(1/s,-v'/s\xi,0),\\ \vec{B}&=(2\pi/a)(0,1/\xi,0),\\ \vec{C}&=(2\pi/a)(0,0,1/\eta), \end{aligned} \tag{16}$$

and the reciprocal-lattice vectors are

$$\vec{K}=h\vec{A}-l\vec{B}-m\vec{C},$$

which we use to define in Al ($Z=3$)

$$\vec{x}=(2k_F)^{-1}\vec{K},$$

$$x=\left(\frac{\pi}{6Z}s\xi\eta\right)^{1/3}\left[\frac{h^2}{s^2}+\left(\frac{l-hv'/s}{\xi}\right)^2+\frac{m^2}{\eta^2}\right]^{1/2} \tag{17}$$

With the choice of basis given in (13) the structure factor, per ion, is

$$\tfrac{1}{2}(1+e^{-i\beta}),$$

where

$$\beta=\vec{K}\cdot\vec{T}=\pi[h(2s-1)/s+(l-hv'/s)(2s-1)\zeta/\xi+m]. \tag{18}$$

Accordingly, the band-structure energy (in Ry/elec-

tron) becomes

$$E_{BS}^{(2)} = \sum_{\vec{x}\neq 0} \left(\frac{1}{\epsilon(\vec{x})} - 1 \right) \frac{k_F^2 x^2}{8\pi Z^2} |v(x)|^2 n(1+\cos\beta). \quad (19)$$

In (19), $\epsilon(\vec{x})$ (the dielectric function of the interacting electron gas) can be written

$$\epsilon(x) = 1 + (\lambda^2/x^2) f(x) g(x); \quad \lambda^2 = 1/(\pi a_0 k_F),$$

with

$$f(x) = \tfrac{1}{2} + \frac{1-x^2}{4x} \ln\left|\frac{1+x}{1-x}\right|,$$

and $g(x, r_s)$ a correction for exchange and correlation. We have not found the latter to make any important correction in the matter of deciding between relative structures at second order.

Using Ewald's method we can determine the Madelung energy in the standard form (again in Ry/electron)

$$E_M = C_M Z^{2/3}/r_s. \quad (20)$$

To find C_M, we normalize the direct lattice vectors by the Wigner-Seitz radius

$$r_{WS} = (3s\xi\eta/8\pi)^{1/3} a,$$

i.e., we define

$$\vec{\rho} = \vec{R}/r_{WS},$$

where

$$\rho = (8\pi/3s\xi\eta)^{1/3}[(ns+pv')^2 + p^2\xi^2 + q^2\eta^2]^{1/2}.$$

Similarly, put

$$\vec{t} = \vec{T}/r_{WS},$$

where

$$|\vec{\rho}+\vec{t}| = (8\pi/3s\xi\eta)^{1/3}\{[ns+pv' + \tfrac{1}{2}(2s-1)]^2 + [p\xi + \tfrac{1}{2}(2s-1)\eta)]^2 + (q+\tfrac{1}{2})^2\eta^2\}^{1/2}$$

Finally, put

$$\vec{G} = r_{WS}\vec{K},$$

with

$$G = 2(9\pi Z/4)^{1/3}\bar{x},$$

then

$$C_M = -\frac{3}{P^2} + \frac{3}{2P^2} \sum_{K\neq 0} (1+\cos\beta) \frac{e^{-K^2/P^2}}{K^2/P^2} - \frac{P}{\sqrt{\pi}} + \frac{\operatorname{erfc}(\frac{1}{2}Pt)}{t} - \sum_{\vec{\rho}\neq 0} \left(\frac{\operatorname{erfc}(\frac{1}{2}P\rho)}{\rho} - \frac{\operatorname{erfc}(\frac{1}{2}P|\vec{\rho}+\vec{t}|)}{|\vec{\rho}+\vec{t}|} \right), \quad (21)$$

where $P(>0)$ is Ewald's dimensionless parameter and erfc denotes the complementary error function. Then at second order, we evaluate (12) by using (6) for $E_{BS}^{(2)}$ [with $v(K)$ there replaced by $\frac{1}{2}(1+e^{-\beta})v(K)$] and (20) and (21) for the Madelung energy. For a given structural choice (corresponding to a particular selection of s, u, v, w) we determine E_0 by the zero-pressure condition $(\partial E/\partial r_s)_{r_{s0}} = 0$. Expressed as an energy per electron, E_0 always has the form

$$\alpha/(\tfrac{4}{3}\pi r_s^3),$$

where α is a property of the ion alone and is assumed not to alter under reasonable variations of density. Since the total energy near zero pressure contains small contributions from the omitted higher-order band-structure terms, the imposition of the zero-pressure condition forces their inclusion in a crude way through the choice of α. To the extent that these terms are not seriously density dependent the subsequent use of this α will therefore continue to incorporate such terms If one takes the Noziéres-Pines form for the correlation energy,[17] it is easy to see that

$$\alpha = \frac{4\pi}{9}\left\{ r_{s0}^2 \left[0.916 + Z^{2/3} C_M + 0.031 r_{s0} + r_{s0}^2 \left(\frac{\partial E_{BS}^{(2)}}{\partial r_s} \right)_{r_{s0}} \right] - 4.42 r_{s0} \right\}, \quad (22)$$

where for the fcc structure observed for Al in its ground state[18] $r_{s0} = 2.0647$. What is required in (22) is $E_{BS}^{(2)}(r_s)$, and this can be calculated by a combination of a direct numerical summation (out to a chosen reciprocal-lattice shell) augmented by integration for the remainder This remainder, designated by $S(x_1, r_s)$ (where x_1 is the radius of the shell) is independent of structure and depends[19] very weakly on r_s. Its contribution is in any event, quite small. At $r_s = r_{s0}$ and for $x_1 = 2.5$ we find $S = 0.005$ Ry/electron, which amounts to 5% of $E_{BS}^{(2)}$ and 0.4% of E

IV ENERGIES AND PHASES RESULTS AND DISCUSSIONS

In Fig. 2 we show a selection of the results we obtain for the Helmoltz free energy E as the crystal structure is continuously deformed from fcc to hcp ($c/a = \sqrt{\frac{8}{3}}$). In this example fcc is lower in energy at all densities considered. This result remains true for other structures, the two that are always closest in energy (at least of the simple structures we consider) being fcc and hcp. It is a straightforward matter to compute the PV term and, hence, in the ground state the Gibbs energy for different phases We find fcc Al (with an assumed static lattice) to have the lowest Gibbs energy and to be the preferred structure, even up to theoretical pressures in excess of 3 Mbar

Contributions to the thermodynamic functions from the ionic degrees of freedom can be estimated from the Debye model; in particular, the zero-point energy is of order $\frac{3}{8}k_B\Theta_D$ per *electron* (about 0.001 Ry) and for temperatures less than the Debye temperature will remain of this order. *Changes* in this energy accompanying changes in crystal structure will be much less than 0.001 Ry. The contribution of the phonons to the pressure is readily shown to be $(\frac{2}{3}\gamma)nk_B\Theta_D$, where $n=N/V$ is the ionic density and γ is the Grüneisen constant. Even for changes of 50% in the equilibrium value of n, the phonons change the pressure calculation above by at most a few kilobars. Figures 3 and 4 give the Gibbs energy as a function of pressure for fcc and hcp, and (for comparative purposes) as a function of r_s for sc, fcc, bcc, and hcp. In Fig. 5 we plot the pressure on a single crystal of Al (under pure hydrostatic strain) as a function of its lattice constant a (rather than r_s) at a nominal temperature of 300 °K The equation of state given there may also be appropriate to polycrystalline samples under less than pure hydrostatic conditions. It is worth remarking that at 3 Mbar, where $a=3.14$ Å, and the nearest-neighbor separation is $(1/\sqrt{2})a$ $=2.22$Å, the distance between ion cores (taking them to have a radius of 0 59 Å) is still 1.04 Å. For the pressure range in Fig. 5 the energy (and the corresponding pressure) is dominated by the terms aris-

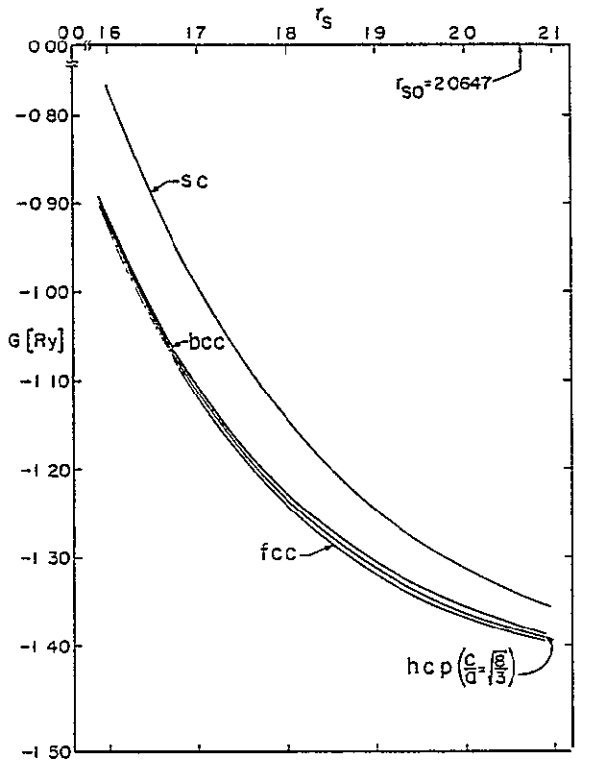


FIG. 3. Gibbs free energy as a function of r_s for several common structures Compare with Fig. 4 where G is plotted against the natural variable pressure.

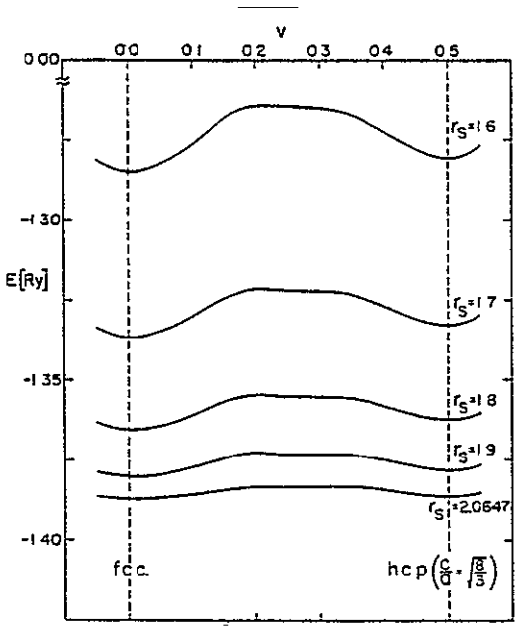


FIG. 2. Helmholtz free energy as a function of r_s and v; the other parameters fixed at their fcc values; varying v here takes the structure from fcc to hcp.

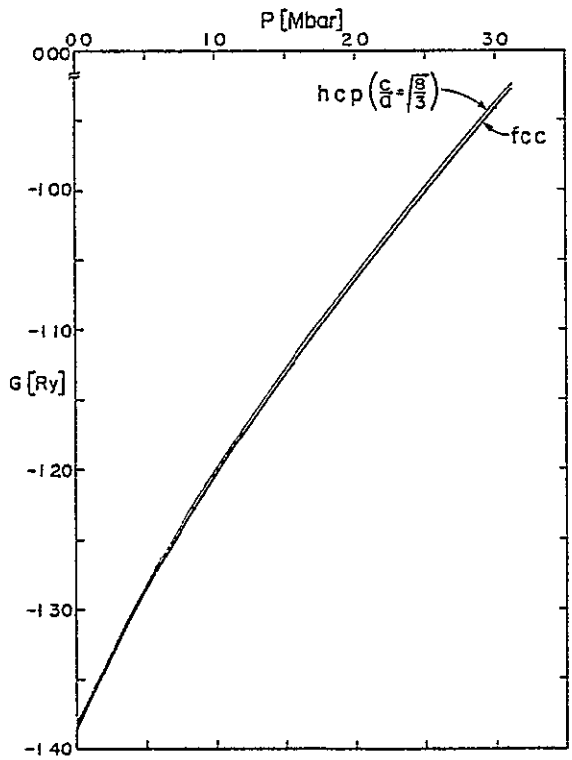


FIG. 4. Gibbs free energy as a function of pressure for the fcc and hcp structures, these have the lowest Gibbs free energy for any fixed pressure P.

ing from electron gas, Madelung energy, and to a much lesser extent, band structure. Energies arising from the direct overlap of ion cores (so called core-core exchange, or Born-Mayer terms) are evidently not important, although it is conceivable that at very much higher pressures (we estimate they will be in excess of 100 Mbar) they could be. This kind of term is difficult to calculate with confidence from first principles, and is normally parameterized in an exponential form (or even as a power law) in expressions giving its contribution to the internal energy. In pressure scales based on these forms, the concern (aside from the implicit pair force approximation) is that the low-pressure determined-parameters may not remain valid in a region of substantial ion-core wave-function overlap. At 3 Mbar we have only a 22% reduction in lattice constant, and core-core overlap is still a small effect, its neglect leads to errors which will be far less important than those arising from the neglect of, for example, the higher-order band-structure contributions to the energy.

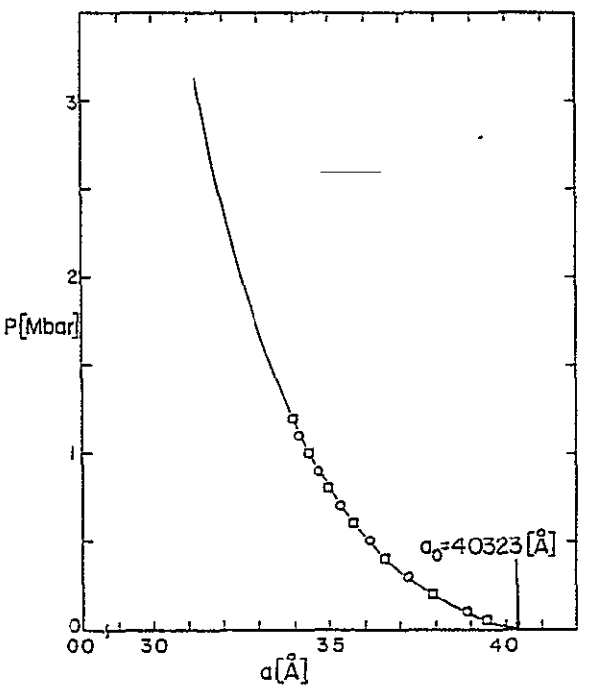


FIG. 5. Pressure as a function of lattice constant for the fcc structure, and experimental points obtained from reduced shock-wave data for two dilute alloys (○ 2024 Al, □ 921-T Al; see Ref. 22) assuming their zero-pressure lattice constant is equal to that of pure Al.

As far as the use of Al in high-pressure devices is concerned it suffers from the disadvantage that its atomic number is quite low It should, however, be visible to x rays in a diamond cell, and the curve presented in Fig. 5 is therefore amenable to experimental test, provided, of course, that sufficiently hydrostatic conditions can be arranged.

If a test of this kind were found to establish as numerically sound the basic curve up to, say, 0.5 Mbar (corresponding to $a = 3.61$ Å), then according to the arguments we have given about it would then appear reasonable to accept the balance of the curve leading to the ultra-pressure region.[20] An independent determination of the pressure can also be used to refine, for example, the form of the pseudopotential used in the high-density regime. It is worthwhile mentioning that the equation of state obtained here agrees within experimental error with the results in the range from 0 to 0 2 Mbar obtained by Roy and Steward.[21] It also agrees very well with shock-wave results for 2024 aluminum and 921-T aluminum up to 1.2 Mbar.[22] Assuming these dilute alloys behave as pure aluminum (with the same lattice constant at zero pressure), we get from the reduced shock data the points plotted in Fig. 5. Small changes in the actual lattice constant are to be expected, and in addition we must expect minor effects from the different pseudopotentials and valences of the impurities. But in homogeneous *dilute* alloys these can only displace the experimental points slightly from those plotted in Fig 5. Finally, our curve is almost parallel to the corresponding one extracted from Ross and Johnson's paper,[4] but is shifted to the left by $\Delta(V/V_0)$ about −0.06. Although some of this difference may be due to numerical inaccuracy (e.g., the APW calculations take only a few points in the fundamental symmetry element of the Brillouin zone) and some due to questions surrounding the correct choice of local exchange potential, probably the bulk of the discrepancy can be traced to the different methods of handling of the zero-pressure condition. In the method of structural expansions,[8] the contribution to the total energy of the zeroth Fourier component of all the interactions is eliminated with the zero-pressure condition at the corresponding experimentally *known* r_s: the *a priori* calculations (such as those in Ref. 4) seek to obtain every term in the ground-state energy from first principles.

The reasons for choosing Al (the paradigm of small-core, close-packed-cubic nearly-free-electron metals) do not exclude other metals displaying similar features, and it may well be that the principles leading to the choice of a metal rather than an ionic crystal for the measurement

of pressure, can be applied to metals such as In, or Pb, providing, of course, that closer attention is paid to problems arising from spin-orbit coupling, nonlocal effects, and the nature of neighboring levels above the Fermi energy.

ACKNOWLEDGMENT

One of us (C.F.) wishes to thank the Universidad Católica, Santiago, Chile for partial support during the tenure of this work.

---

*Work supported by NASA under Contract No. NGR-33-010-189.

[1]These are most naturally calculated as functions of volume (for a given temperature) which is eliminated in favor of the pressure only if the equation of state is known.

[2]This does not preclude changes in electron structure such as might occur when a Fermi surface changes its topology.

[3]For a review of the techniques, see V. Heine and D. Weaire, Solid State Phys. 24, 249 (1970).

[4]M Ross and K. W. Johnson, Phys. Rev. B 2, 4709 (1970).

[5]D. L. Decker, J. Appl. Phys. 36, 157 (1965), W. A. Bassett, T. Takahashi, T. W. Stook, Rev. Sci. Instrum. 38, 37 (1967).

[6]See, for example, K. Sturm and N. W. Ashcroft, Phys. Rev. B 10, 1343 (1974).

[7]N. W. Ashcroft, Philos. Mag. 8, 2055 (1963). The Fourier components so determined are *folded* in the sense that they are extracted from experiment with the aid of a low-order secular equation (see Ref. 6). In addition they incorporate small Debye-Waller type corrections (see Sec. IV).

[8]C. J. Pethick, Phys. Rev. B 2, 1789 (1970); J. Hammerberg and N. W. Ashcroft, *ibid.* 9, 409 (1974).

[9]N. W. Ashcroft and D. C. Langreth, Phys. Rev. 155, 682 (1967). Any other contributions to the energy from bound charge (fluctuations in the ion cores, for example) will be subsumed in this E_0.

[10]We will take up the (small) effects of temperature in Sec. IV.

[11]For the purposes of this comparison it is sufficient to choose any of the common structures. Here we have given the figures appropriate to fcc.

[12]P. Lloyd and C. Sholl, J. Phys. C 1, 1620 (1968); E. G. Brovman, and Yu Kagan, Zh. Eksp. Teor. Fiz. 57, 1329 (1969) [Sov. Phys.—JETP 30, 721 (1970)].

[13]For example, E. G. Brovman and G. Solt, Solid State Commun. 8, 903 (1970).

[14]See, for example, N. W. Ashcroft and N D. Mermin, *Solid State Physics* (Holt, Rinehart, and Winston, New York, 1975).

[15]D. Stroud and N. W. Ashcroft, Phys. Rev. B 5, 371 (1972).

[16]We are therefore performing a partial scan of Bravais lattice space using a technique not unlike that of E. G. Brovman, Yu Kagan, and A. Kholas, Zh. Eksp. Toer. Fiz. 61, 2429 (1971) [Sov. Phys.—JETP 34, 1300 (1972)].

[17]P Nozières and D. Pines, *Quantum Liquids* (Benjamin, New York, 1966).

[18]This is determined, for $T=0$ from data on thermal expansion recorded in *Properties of Materials at Low Temperature* (Pergamon, New York, 1961); NBS part II; 2.132; and combined with the lattice constant measured by A. S. Cooper, Acta Crystallogr 15, 578 (1962).

[19]For large enough x, $S(x_1, r_s)$ can be calculated by integration (rather than summation) where the asymptotic result can be written $S(x_1, r_s) \sim -(2Z/9\pi^2 x_1^3) \times(1-3 \sin y/y + \cdots)$, where $y=4(9\pi/4)^{1/3}(r_c/r_s)x_1$.

[20]In this respect it is worth noting that the energy dependence of the pseudopotential is a subtlety that, although expected to give small corrections (see Ref. 5) for small over-all energy changes, might well require proper inclusion for large changes in density.

[21]N. N. Roy and E. G. Steward, Nature (Lond.) 224, 905 (1969). The agreement of their results with the predictions of the Murnaghan equation of state is also within experimental error.

[22]R. G. McQueen, S. P. Marsh, J. N. Fritz, and W. J. Carter, *High-Velocity Impact Phenomena*, edited by R. Kinslow (Academic, New York, 1970), pp. 530, 531.

Short-range interaction between hydrogen molecules*

A. K McMahan, H. Beck,† and J. A. Krumhansl
Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14850
(Received 30 January 1974)

Recent calculations of the ground-state energy of a system of four hydrogen atoms are reviewed from the point of view of discerning the short-range interaction potential between two hydrogen molecules. Consistency amongst the results of these calculations suggests that the potential for intermolecular separations in the region 1–2 5 Å can now be specified to about 10% with considerable confidence Analytic fits to spherical averages of these results are presented For calculations of properties of high-density solid molecular hydrogen, the bare pair potential may thus be regarded as well determined. The role of multicenter terms can then be examined, as for example, recent reported work seems to indicate that pairwise additivity is not altogether valid in practice

I INTRODUCTION

The purpose of this paper is to review recent calculations of the short-range, repulsive part of the interaction potential between two hydrogen molecules. Uncertainty in this portion of the potential has led to widely differing determinations of the equation of state for molecular hydrogen at very high pressures, and contributed to variations by more than an order of magnitude amongst predictions of the molecular to atomic phase-transition pressure.[1–10] We demonstrate in this review that recent calculations[11–25] of the short-range part of the potential are in sufficient agreement with each other as to suggest that this part of the potential may now be fairly well established Unfortunately, there are still significant discrepancies with the limited experimental information available.[26–31] Most of the calculations that we discuss have appeared in the chemical physics literature, and many have been motivated by other concerns such as the four-center exchange mechanism between two impinging hydrogen molecules. Since this review is intended for a more general audience, we have included a brief description of the so-called *ab initio* techniques that have been used. It is not the purpose of this paper to give a complete review of the H_4 calculations, and we refer the reader to the paper by Rubinstein and Shavitt[18] for a more thorough list and discussion of the earlier efforts.

The organization of the paper is as follows. In Sec.II we describe the *ab initio* techniques, and in Sec. III the numerical results for the H_2-H_2 interaction energy that have been obtained with these methods. Possible analytic forms for the short-range part of the potential are discussed in Sec. IV In Sec. V we comment on the applicability of these various results to calculations of the ground-state energy of molecular solid hydrogen Finally, our summary is presented in Sec. VI.

II MATHEMATICAL TECHNIQUES

We describe in this section the *ab initio* techniques by which the ground-state energy of a system of four hydrogen atoms has been determined.[11–25] It is customary to begin by making the Born-Oppenheimer approximation and neglecting any zero-point motion of the four nuclei. The nuclear position vectors $\vec{R}_A$, and thus the geometry of the system, are accordingly parameters in the problem. The desired energy is then the ground-state eigenvalue of the Hamiltonian

$$H=\sum_{A<B}\frac{1}{R_{AB}}+\sum_{i}\left(-\frac{1}{2}\nabla_i^2-\sum_{A}\frac{1}{r_{iA}}\right)+\sum_{i<j}\frac{1}{r_{ij}}\,, \tag{1}$$

where the indices A and i run over the four nuclei and four electrons respectively, $R_{AB}=|\vec{R}_A-\vec{R}_B|$, $r_{iA}=|\vec{r}_i-\vec{R}_A|$, and atomic units[32] have been used. The methods by which this energy has been approximately determined have in general been variational,[33] and thus have given upper bounds. These methods may be categorized according to the generality of the trial wave function used.

Heitler-London (HL)

The simplest calculation would appear to be a generalization of the well-known Heitler-London approach for the hydrogen molecule. In the case of four hydrogen atoms, one has

$$\psi_{\rm HL}=\tfrac{1}{2}[(a\bar{b}c\bar{d})-(\bar{a}bc\bar{d})-(a\bar{b}\bar{c}d)+(\bar{a}b\bar{c}d)], \tag{2}$$

$$(\bar{a}\bar{b}\bar{c}\bar{d}) \equiv \frac{1}{\sqrt{4!}} \sum_P (-1)^P P[\chi(|\vec{r}_1 - \vec{R}_a|)\chi(|\vec{r}_2 - \vec{R}_b|)\chi(|\vec{r}_3 - \vec{R}_c|)\chi(|\vec{r}_4 - \vec{R}_d|)\alpha(1)\beta(2)\alpha(3)\beta(4)], \quad (3)$$

$$\chi(r) \equiv (\zeta^3/\pi)^{1/2} e^{-\zeta r}. \quad (4)$$

As usual, the two spin functions are indicated by α and β. In Eq (2), the bars placed over certain letters indicate the arrangement of the spin functions as shown in Eq. (3). The permutation operator P runs over all 24 permutations of four objects, and permutes both spatial and spin variables. Since it is presumed that the ground state will be an eigenfunction of the total spin with eigenvalue zero, it is necessary to combine four Slater determinants as is done in Eq. (2). This is a covalent (as contrasted to ionic) wave function, in that each of the four atomic orbitals (one centered on each nucleus) is singly occupied. If one substitues Eq. (3) into Eq. (2), the spin functions may be grouped in the form of a singlet state for the electrons on nuclei a and b, multiplied by a singlet state for those on c and d One considers this wave function to describe a state in which covalent bonds exist between atoms a and b, and between atoms c and d. It is possible to construct two more covalent wave functions, corresponding to bonds between other pairs of atoms, although only two of the three wave functions are linearly independent The given geometric arrangement of the nuclei dictates which of the three (if any) is the best choice.

The Heitler-London wave function has no variational parameters (unless the effective nuclear charge ζ is varied), and so one must only evaluate

$$E_{HL} = \frac{\langle \psi_{HL}|H|\psi_{HL}\rangle}{\langle \psi_{HL}|\psi_{HL}\rangle}. \quad (5)$$

The interaction energy between molecules may then be found by subtracting the energy of two isolated molecules—also calculated in the Heitler-London approximation. This is not a trivial exercise, for two reasons. The first is that for a general geometry, Eq (5) involves some 64 distinct electron-nucleus attraction and electron-electron repulsion integrals [34] Cancellation amongst these various terms results in the interaction energy being one or more orders of magnitude smaller than the size of some individual terms Second, simple analytic expressions for the 39 three- and four-center integrals do not exist, and only in the last ten years have these integrals been accurately evaluated by rather elaborate computer programs [35] In the early work, de Boer[36] neglected three- and four-center integrals altogether, while Evett and Margenau[37] and Mason and Hirschfelder[38] attempted to approximate them. Because of the extensive cancellation mentioned, such approximation-schemes are not reliable While giving reasonable dependence of the interaction energy on intermolecular separation, the calculations of de Boer and of Mason and Hirschfelder, for example, overestimate the orientation dependence by more than a factor of 2 We return to this point later.

Full configuration interaction

The two linearly independent covalent wave functions are referred to as configurations. Given our set of four atomic orbitals, one centered about each nucleus, it is also possible to construct 12 singly ionized configurations of the form

$$\psi_{s,ion} = (1/\sqrt{2})[(\bar{a}\bar{b}\bar{c}c) - (\bar{a}\bar{b}c\bar{c})] \quad (6)$$

and six doubly ionized configurations of the form

$$\psi_{d,ion} = (\bar{a}\bar{a}c\bar{c}). \quad (7)$$

Each is a linearly independent wave function, satisfying the Pauli principle, and a spin-zero eigenfunction of the total spin. They correspond to the 20 possible ways of placing four indistinguishable electrons on four protons (using only 1s states) consistent with zero total spin

A variational calculation of the ground-state energy in which the trial wave function is composed of a sum of these configurations, each multiplied by a variational parameter, is referred to as a "configuration-interaction" (CI) calculation. In a full configuration-interaction calculation, all configurations consistent with the geometric symmetry of the ground state are employed. To be more precise, the configurations referred to here are actually linear combinations of the original configurations which transform according to the appropriate irreducible representation of the point group of the four-atom system Thus, for the linear geometry (see Fig. 1), only 12 (out of 20) configurations are needed.

A full CI calculation may be improved by enlarging the basis So far, we have considered what is known as a 1s-Slater-type basis, meaning that we used four atomic orbitals obtained by centering a 1s-Slater-type orbital [Eq. (4)] about each of the four nuclei. This is known as a "minimal" basis set in that only the 1s orbital is occupied in the ground state of an isolated hydrogen atom. Williams,[11] Magnasco and Musso,[13] and Wilson and Goddard[16] have used this basis set in their full CI calculations on the H_4 system

Rubinstein and Shavitt,[18] and Silver and Stevens[23] have used a $1s, 1s'$-Slater-type basis set. The use of two $1s$ orbitals (having different exponents) in this "double-zeta" basis appears to be simply a convenient device whereby minor improvements can be made over the minimal basis wave function, most importantly in the region between the atoms. Bender and Schaefer[22] have gone a step further by adding p orbitals, using a $1s, 1s', 2p_x, 2p_y, 2p_z$-Gaussian-type basis in their calculations. This $1s$ orbital is a "contracted" sum of three Gaussians, while the $1s'$ orbital is a single Gaussian. Amazingly enough, full CI calculations with Gaussian orbitals have proved quite successful. Among the advantages of their use is the easy evaluation of multicenter integrals, while a disadvantage is that generally a large enough basis must be used so as to at least crudely be capable of representing a Slater function. A discussion of the philosophy behind these various choices of basis sets is given in the book by Schaefer.[39] One fact should be borne in mind: the number of configurations involved increases dramatically with the size of basis chosen. A full CI calculation for the linear geometry, for example, involves 12, 176, and 2172 configurations, respectively, for the $1s$; $1s, 1s'$; and $1s, 1s', 2p_x, 2p_y, 2p_z$ basis sets.

"Self-consistent field"

The "self-consistent-field" (SCF) calculation, as referred to in the papers of interest to us in this review, is a particular version of the Hartree-Fock approach. One seeks to minimize the energy using a wave function of the form

$$\psi_{\mathrm{SCF}} = \frac{1}{\sqrt{4!}} \sum_P (-1)^P P[\Phi_1(1)\Phi_1(2)\Phi_2(3) \times \Phi_2(4)\alpha(1)\beta(2)\alpha(3)\beta(4)] . \tag{8}$$

However, in contrast to the most general Hartree-Fock approach, the molecular orbitals Φ_1 and Φ_2 are restricted in this method to be linear combinations of whatever basis functions are being used. In the case of the minimal basis set, then

$$\Phi_1(\vec{r}) = C_{1a}\chi(|\vec{r}-\vec{R}_a|) + C_{1b}\chi(|\vec{r}-\vec{R}_b|) + C_{1c}\chi(|\vec{r}-\vec{R}_c|) + C_{1d}\chi(|\vec{r}-\vec{R}_d|) , \tag{9}$$

and the coefficients C would be the quantities to be determined. Actually, for such a small basis, geometric symmetry alone will often be sufficient to determine these coefficients. Bender and Schaefer[22] and Tapia and Bessis[19-21] have used $1s, 1s', 2p_x 2p_y, 2p_z$ and $1s, 1s', 1s'', 2p_x, 2p_y, 2p_z$-Gaussian bases in their SCF calculations.

Both the SCF and the Heitler-London (HL) wave functions are contained as special cases within the corresponding full CI wave function. They offer shorter computing time at the cost of less-accurate results. In general, the SCF wave function exhibits too little spatial correlation amongst the four electrons; the HL wave function, too much. The SCF wave function is best suited to geometries in which all four atoms are closely spaced; the HL wave function, when the atoms are far apart. In any case, for a given basis, the full CI calculation always yields lower upper-bounds on the ground-state energy than either the HL or SCF methods.

Other methods

The same full CI wave function may be arrived at from either the valence-bond point of view, in which ionic configurations are added to the covalent configurations, or from the molecular-orbital point of view, in which excited configurations are added to the SCF configuration. There are a number of limited CI calculations (i.e., not full) based on one or the other of these viewpoints. These methods include the "group function" approach of Magnasco, Musso, and McWeeny,[14] and the "GI" method of Wilson and Goddard.[16,17] The "SCF + CI" method, which we shall take to mean the SCF configuration plus all singly and doubly excited configurations, has proved to be very successful for at least the linear geometry.[22] Bender and Schaefer,[22] Tapia and Bessis,[21] Kochanski *et al.*,[24] and Ree and Bender[25] have used this approach.

III SURVEY OF NUMERICAL RESULTS

This section reviews numerical results obtained for the ground-state energy of the H_4 system by the *ab initio* techniques described previously. We first make use of these results to give some indication of when the concept of interacting H_2 molecules is valid and where it breaks down. Then we specialize to the problem of the angular (viz., Fig. 1) and intermolecular separation dependence of the H_2-H_2 interaction energy. At this stage quantitative comparison of the various computational methods is made.

Interacting H_2 molecules

One may identify a particular pair of hydrogen atoms as constituting an H_2 molecule if, when considered as a function of the distance between these two atoms, the energy of the full H_4 system is near a local minimum. A system of four infinitely separated hydrogen atoms has an energy of -2.00 hartrees.[32] The energy may be lowered to -2.35 hartrees by grouping the atoms into two infinitely

separated pairs, with the distance between atoms composing a given pair being 1 40 bohrs. The H_2 molecule binding energy, 0.17 hartree, accounts for this energy reduction.[40] The energy of the H_4 system increases when the two pairs of hydrogen atoms are pushed close together; i.e , there is a repulsive short-range interaction between the two H_2 molecules.

One would expect the concept of interacting H_2 molecules to remain valid down to separations for which the interaction energy approaches the binding energy in magnitude. This appears to be borne out by the calculations. In Fig. 2, we show the Silver and Stevens[23] results for the rectangular geometry. The abscissa specifies one side of the rectangle(R_2); the curves are labelled according to the other (R_1). It is evident that the lowest energies are obtained when one side is near 1 40 bohr (the equilibrium H_2 bond length), and the other side is large. Decreasing this larger side (the intermolecular distance) results in exponential-like increase as seen in the curve labelled $R_1 = 1.4$. The effect of intermolecular distance on the local potential well associated with the H_2 bond length can be seen in the dotted portion of the curves, where R_1 is to be taken now as the intermolecular distance; and R_2, the bond length The calculations of Conroy and Malli,[33] in particular their Fig. 6, suggest that the obvious trend here does indeed result in an eventual loss of the barrier for $R_2 > 1.4$ bohrs as R_1 is further decreased below 1.8 bohrs Somewhat before this point, the vibrational zero-point energy of the two molecules associated with the coordinate R_2 (about 0.02 hartree as estimated from the curvature at $R_2 = 1.4$ bohrs) will result in loss of the H_2 bonds.

Does the optimal bond length change as the two molecules are pushed closer together? Analytic fits to the potential wells shown in Fig. 2 yield minima within a percent of 1.40 bohrs for the range of intermolecular separations from 2 8 to 1.8 bohrs On the other hand, Conroy and Malli,[33] Wilson and Goddard,[17] and Tapia *et al.*[19,21] have reported results for the same rectangular geometry suggesting the optimal bond length shrinks as the intermolecular distance is decreased From the first two of these papers, the shrinkage may be estimated to be about 4% for intermolecular separations near 2.2 bohrs. For two H_2 molecules approaching each other in a linear manner, the results of Wilson and Goddard,[16] as seen in their Fig 18, suggest a similar shrinking of the optimal bond length. Extrapolation of their data suggests about a 4% effect for intermolecular separations near 3.1 bohrs. Recent work of Ree[41] implies the optimal bond length decreases for all geometries shown in Fig. 1. He obtains somewhat larger effects. The important point to bear in mind, as can be seen in Fig. 2, is that these uncertainties in the bond length lead to errors in

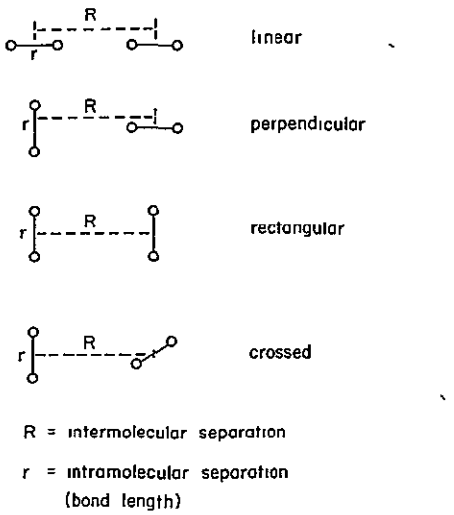


FIG. 1 Geometries of the H_4 system. The linear, perpendicular, and rectangular arrangements lie in the plane of the paper as shown. In the crossed geometry, the intramolecular axis of the right-hand molecule is perpendicular to the plane of the paper

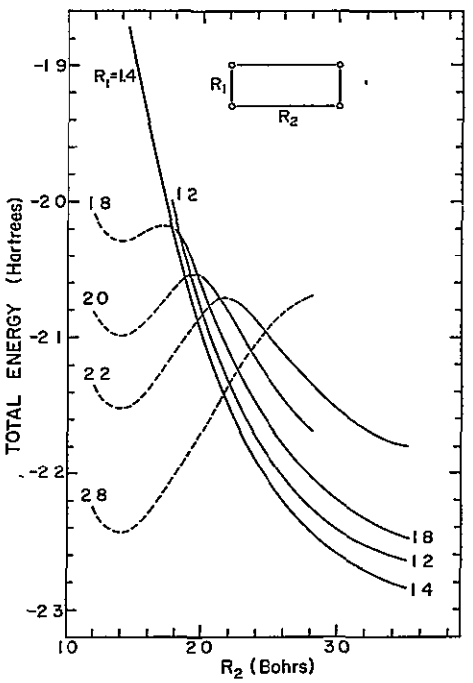


FIG 2 Total energy of the H_4 system for the rectangular geometry The abscissa specifies one side of the rectangle R_2 and the curves are labelled according to the length of the other side R_1. Both lengths are in bohrs The H_2 bond length and the H_2-H_2 intermolecular separation may be identified with R_1 and R_2, respectively, for the solid curves, and the reverse, for the dashed curves These results are from Silver and Stevens (Ref 23).

the interaction energy generally less than a few percent. Accordingly, calculations of the H_2-H_2 interaction energy based on a fixed bond length of 1.40 bohrs should be valid to within this same accuracy.

As a rough summary one might say that the idea of the H_2 bond, and an associated length more or less equal to 1.4 bohrs, are relevant down to separations where the distance between the nearest atoms on two approaching molecules is about equal to, or perhaps half again as large, as this bond length On further contraction, both the local potential wells signifying the bonds, and the associated length are lost The Bender and Schaefer[22] results for the linear H_4 system, for example, show that in this regime it is energetically favorable to equally space the four atoms rather than trying to maintain the 1.4-bohr bond length (see Fig. 3). For lower (linear) densities this equally spaced geometry, while a bound state with respect to four separated atoms, is clearly unstable with respect to the formation of H_2 molecules.

Interaction energy

A partial judgement of the relative merit of the computational techniques can be made by checking their results for the ground-state energy of a single H_2 molecule (see Table I). Since these calculations are variational, the results are quite as expected: Lower energies are obtained by using larger basis sets, and by including all possible configurations (full CI) which may be constructed from the given basis set. This table is only indirectly related to our problem, however, since we are interested in *relative* changes in the energy of the H_4 system as the constituent H_2 molecules are moved about. The interaction energy of two H_2 molecules is calculated as the energy of the H_4 system less the energy of two infinitely separated molecules evaluated in the *same* approximation Thus, for example, the large-basis SCF calculations of the interaction energy are superior to the minimal-basis full CI results, in spite of the fact that the latter technique gives the lower H_2 molecule ground-state energy.

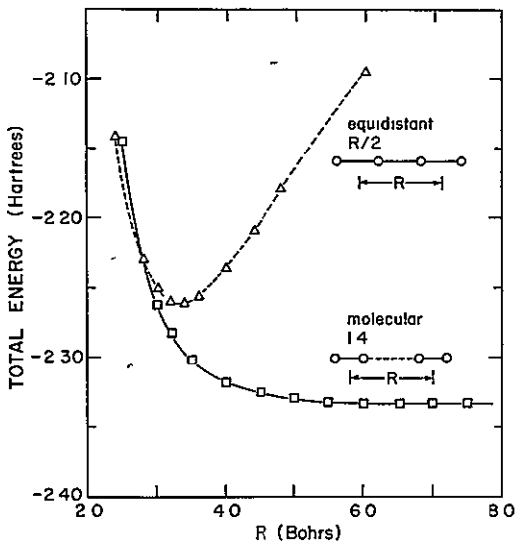


FIG. 3 Total energy of the H_4 system for the linear geometries. The solid curve corresponds to the "molecular" arrangement in which the atoms are grouped into two pairs, as shown, with a "bond length" of 1 4 bohrs The dashed curve corresponds to the "atomic" arrangement in which the atoms are equally spaced, the interatomic separation being $R/2$ The curves intersect for $R=2$ 8 bohrs These results are from Bender and Schaefer (Ref. 22)

The results of minimal-basis full CI calculations by Magnasco and Musso,[13] Williams,[11] and Wilson and Goddard[16] are shown in Fig. 4 for the linear and rectangular geometries. The density dependence is roughly exponential, $e^{-\alpha R}$, with α ranging between 1 80 and 1.85 bohrs^{-1} for the linear and 1.67–1.90 bohrs^{-1} for the rectangular geometry as the intermolecular distance R is increased from 3 to 5 bohrs The Williams[11] results place the energy of the crossed and of the perpendicular geometries, respectively, about 15% below and 50% above those of the rectangular geometry. In contradiction to the statement made by Hoover *et al.*,[10] it is clear that the interaction energy of the linear geometry as calculated with the minimal basis set is only about a factor of 2 larger than that of the rectangular geometry We have also included in Fig 4 the results of the Heitler-London calculation using correct multicenter integrals.[44] The fairly close agreement with the full CI results clearly points out the danger of using approximate multicenter integrals as in the early Heitler-London calculations by de Boer,[36] and Mason and Hirschfelder.[38] An angular dependence more than twice as large as seen here was reported in those papers.

The results of CI calculations using larger bases (specified in Table I) are shown in Figs 5 and 6.[45] The results of Bender and Schaefer[22] and of Silver and Stevens[23] shown here are from full CI calculations. Those of Tapia and Bessis[21] and of Kockanski *et al.*[24] are from the SCF + CI technique, which gives values for the interaction energy within a few percent of full CI values for the linear case.[22] For intermolecular separations R around 3 bohrs, the curves in Fig 5 have about the same dependence on this parameter as in the minimal-basis calculations, i e., $e^{-\alpha R}$ with $\alpha=1$ 81 and 1 62 bohrs^{-1} for the linear and rectangular cases, respectively. The actual values of the interaction

TABLE I Ground-state energy of the H_2 molecule as calculated in various approximations

Basis	Type	Exponent	r_0 [a]	Reference	HL [b]	SCF [b]	SCF+CI [b]	Full CI [b]
1s	Slater	1 193	1 4166	Magnasco and Musso[13] Williams[11]	−1.138 20 [c]	−1 127 98		−1.147 89
		1 05	1 4	Wilson and Goddard[16]		−1 128 19		−1 147 91 [d]
1s, 1s′	Slater	0 9, 1 2	1 4	Silver and Stevens[23]				−1 151 75
1s, 1s′, 1s″, $2p_x$, $2p_y$, $2p_z$ [e]	Gaussian		1 4166	Tapia and Bessis[21]		−1 125 50	−1.135 10	
1s, 1s′, $2p_x$, $2p_y$, $2p_z$ [f]	Gaussian		1 4	Bender and Schaefer[22] Ree and Bender[25]		−1 131 20	−1 165 19	−1 165 70
1s, 1s′, 1s″, $2p_x$, $2p_y$, $2p_z$, $2p'_x$, $2p'_y$, $2p'_z$ [g]	Gaussian		1 4	Kochanski *et al.* [24]			−1.168 02	−1.168 60
James-Coolidge function	100 terms		1 4	Kolos and Roothaan[42] Kolos and Wolniewicz[40]		−1 133 64		−1 174 47
Experiment				Stoicheff[43]				−1 174 45

[a] Bond length in bohrs, 1 bohr = 0 529 17 Å.
[b] Energy in hartrees, 1 hartree = 27 211 eV
[c] Evaluated for this choice of parameters by the authors
[d] r_0 = 1 4166
[e] Each function is a single Gaussian
[f] The 1s function is a contracted sum of three Gaussians
[g] The 1s function is a contracted sum of four Gaussians

energy, however, are smaller by 36% and 12% respectively. This reduces the rectangular-to-linear variation from about a factor of 2.2 to 1.6 In Fig. 6 it is seen that the interaction energy of the perpendicular geometry has also been reduced relatively more strongly than that of the rectangular case, so that only about a 15% variation in energy is involved in changing the orientation from the crossed to the rectangular, and then to the perpendicular geometry.

For values of the intermolecular separation greater than about 4 bohrs, Fig. 5 shows that the interaction energy begins to fall off considerably faster than an exponential. This behavior, which was only barely suggested by the minimal-basis calculations, reflects the importance of the attractive van der Waals or dispersion forces in this region. In fact, Tapia and Bessis,[21] Bender

FIG. 4 Minimal-basis calculations of the H_2–H_2 interaction energy for linear and rectangular geometries. The full CI results of Magnasco and Musso (Ref 13), Williams (Ref. 11), and Wilson and Goddard (Ref. 16) are shown Results of the Heitler-London calculations (Ref 44) are included for comparison The two curves differ by a factor of 2 2, 1 9, and 1 8 for intermolecular separations of 3, 4, and 5 bohrs, respectively The uppermost two points of Wilson and Goddard were obtained with a bond length of 1 4 bohrs. The lowest two points of Magnasco and Musso are from their limited CI calculations (Ref 12). Some of Williams' questionable (Ref 12) large separation results have been omitted

and Schaefer,[22] and Kochanski *et al.*[24] have all observed some form of attractive van der Waals minimum (depth ~ 10^{-4} hartree) in the interaction energy for intermolecular separations around 6.5–7.0 bohrs. Kochanski *et al.* note that calculations in this region are extremely sensitive to the choice of basis, and that a $2p$ orbital with a small exponent is essential. In contrast to the orientation dependence seen for smaller separations, Kochanski *et al.* find the perpendicular geometry to be most stable for intermolecular separations greater than about 4.5 bohrs. There does not appear to be any one type of force responsible for this fact, as they note that the valence, quadrupole, and the dispersion forces all contribute to this stability.

Margenau and Kestner[46] have argued that an SCF calculation of the interaction energy cannot include dispersion effects. This seems intuitively clear in that electron-electron correlations (aside from those originating from the antisymmetrization) are not incorporated in the SCF wave function, and such correlations would appear to be essential to an induced dipole-induced dipole interaction. In Fig. 7 we show the results of SCF calculations by Bender and Schaefer[22] and by Tapia and Bessis[21] which are consistent with these expectations. For intermolecular separations less than 3 bohrs, these results are in fairly close agreement with the CI calculations. For larger separations they fall off too slowly, roughly exponentially, and do not display an attractive van der Waals minimum. For very large separations, greater than 12.5 bohrs, the SCF calculations of Bender, Schaefer, and Kollman[47] are in quantitative agreement with the predicted classical quadrupole-quadrupole interaction.

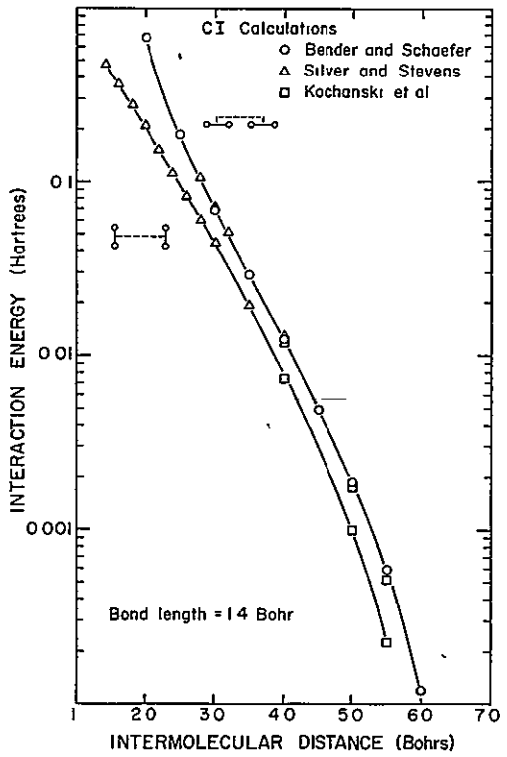


FIG 5 Extended-basis calculations of the H_2–H_2 interaction energy for linear and rectangular geometries. The results of Bender and Schaefer (Ref. 22) and of Silver and Stevens (Ref. 23) are full CI, while those of Kochanski *et al.* (Ref. 24) were obtained by the SCF + CI technique. The bases used are specified in Table I. The two curves differ by a factor of 1.6, 1.6, and 1.9 for intermolecular separations of 3, 4, and 5 bohrs. For these same separations, the linear results are lower by 36%, 27%, and 31%, respectively, in comparison to the corresponding minimal-basis results (Fig. 4); while the rectangular results are lower by 12%, 12%, and 34%, respectively, in comparison to the rectangular results in Fig. 4.

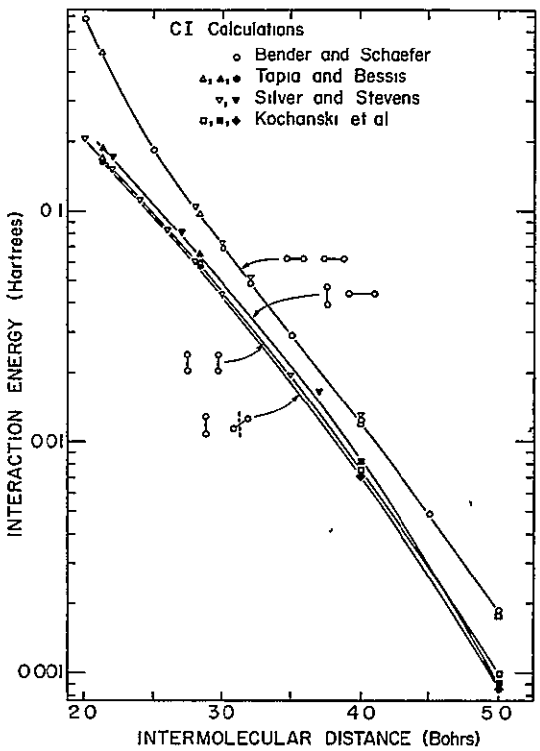


FIG 6 Extended-basis calculations of the H_2–H_2 interaction energy for various geometries. The results of Bender and Schaefer (Ref. 22) and of Silver and Stevens (Ref. 23) are full CI, while those of Tapia and Bessis (Ref. 21) and of Kochanski *et al.* (Ref. 24) were obtained by the SCF + CI technique. The bases used are specified in Table I. The results for the linear and rectangular geometries (open symbols) are identical to those in Fig. 5. For intermolecular separations from 3 to 4 bohrs, the results for the perpendicular and crossed geometries (closed symbols) are, respectively, about 10% above and 5% below those for the rectangular geometry.

To summarize this section, we note that CI calculations using an extended basis that includes a diffuse $2p$ orbital appear to be necessary to accurately determine the H_2-H_2 interaction energy for all separations. There is sufficient numerical agreement for intermolecular separations between 2 and 5 bohrs to suggest that the curves in Figs. 5 and 6 are correct to within better than 10% Furthermore, these results are expected to include *all* contributions to the interaction energy.

IV ANALYTIC EXPRESSIONS

The interaction energy of two hydrogen molecules is generally subdivided into contributions from (i) the short-range valence (overlap, or exchange) forces, (ii) the long-range dispersion forces, and (iii) the electrostatic quadrupole-quadrupole forces. Analytic expressions for the latter two contributions are fairly well established.[34,48-50] We confine our attention to the short-range part of the interaction energy.

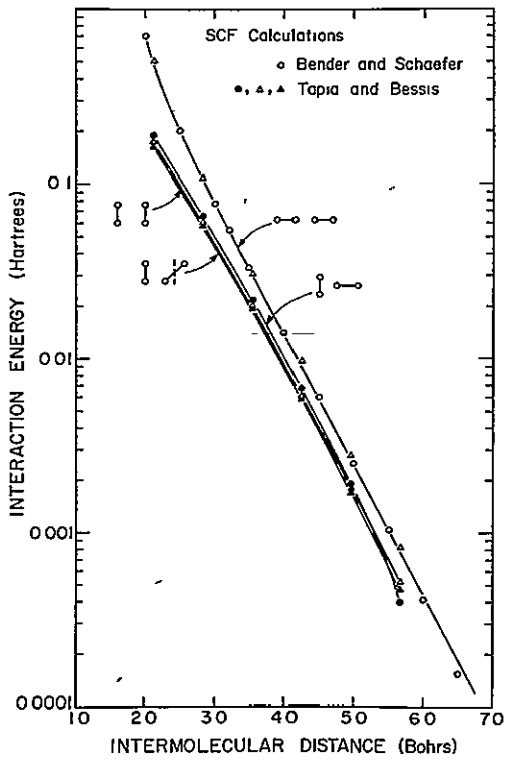


FIG 7 Extended-basis SCF calculations of the H_2-H_2 interaction energy for various geometries The SCF results of Bender and Schaefer (Ref 22) and Tapia and Bessis (Ref 21) are shown The choice of basis is specified in Table I For intermolecular separations less than about 3 0 bohrs these results are generally within a few percent agreement with the CI results shown in Figs 5 and 6 For intermolecular separations around 5 bohrs, these SCF results are higher than the CI results by about (35–90%), depending on geometry

Both de Boer[36] and Abrikosov[2] chose forms for the valence contribution which may be interpreted as representing pairwise interactions between the *atoms* making up the two H_2 molecules:

$$\Phi^{(\mathrm{val})} = \epsilon(R_{ac}) + \epsilon(R_{ad}) + \epsilon(R_{bc}) + \epsilon(R_{bd}) . \qquad (10)$$

Atoms a and b constitute one molecule; c and d, the other While de Boer chose an exponential for the function $\epsilon(R)$, Abrikosov used an appropriate average of the singlet and triplet interactions between two hydrogen atoms. In light of the results discussed in Sec. III, however, there are serious objections to the general form given by Eq. (10). If the intramolecular separation is taken to be near 1.4 bohrs, any choice for the function $\epsilon(R)$ giving the right dependence on intermolecular separation for some particular geometry results in an orientation dependence of about a factor of 5. Yet all *ab initio* calculations have shown an overall orientation dependence of a factor of 2 or less. It is to be emphasized in particular, that the de Boer potential can *not* adequately represent any of the results discussed in Sec. III, including the minimal basis work. Neece *et al.*[8] were able to fit the Magnasco and Musso[13] results with a de Boer potential only because the Magnasco and Musso work did not include any of the high-energy geometries such as the perpendicular or linear arrangements. These facts are illustrated in Fig. 8, where the de Boer potential with the choice of parameters used by Neece *et al.*, is plotted for the standard geometries, and compared to minimal-basis full CI calculations for the rectangular and linear cases.

Equation (10) can be made to yield an overall dependence on orientation of about a factor of 2 if the intramolecular separation is artificially chosen to be a third or so smaller than 1.4 bohrs. However, in this case the perpendicular geometry still falls midway between the linear and rectangular results, and the dependence of the interaction energy on intermolecular separation can not be made satisfactory for all geometries.

The close agreement of the Heitler-London calculations with the minimal-basis full CI results shown in Fig 4, might suggest that de Boer's original goal of selecting out a few dominant terms from the Heitler-London expression might still be achieved. Unfortunately, there are simply too many equally large, and partially cancelling terms for this to be feasible. The angular dependence immediately suffers from such selection processes For example, in their book Margenau and Kestner[34] make a slight approximation in the Heitler-London expression based on neglecting the fourth power of the ratio of the inter- to intramolecular overlap

integrals. Evaluation of this expression using correct multicenter integrals yields results about 20% higher for the rectangular geometry and about 100% higher for the linear geometry, in comparison with the full Heitler-London result

The angular dependence of the interaction potential appears to be of rather high order, as is evidenced in Fig. 6. Low-order terms of the form

$$(\cos^2\theta_1+\cos^2\theta_2)f(R),$$

where θ_1 is the angle between the axis of the first molecule and the line joining the centers of mass of the two molecules, would place the perpendicular results halfway between those for the rectangular and linear cases. This is clearly not the case.

The problem of fitting the angular behavior may be avoided in first approximation by performing some form of average over the angular variables, as is done by Hoover *et al.*[10] and by Ree and Bender [25] Hoover *et al* arrive at the potential

$$\Phi=\frac{8\ 2e^2}{a_0}\ e^{-1.74x}-\frac{e^2}{a_0}(13x^{-6}+116x^{-8})e^{-400x^{-6}}, \tag{11}$$

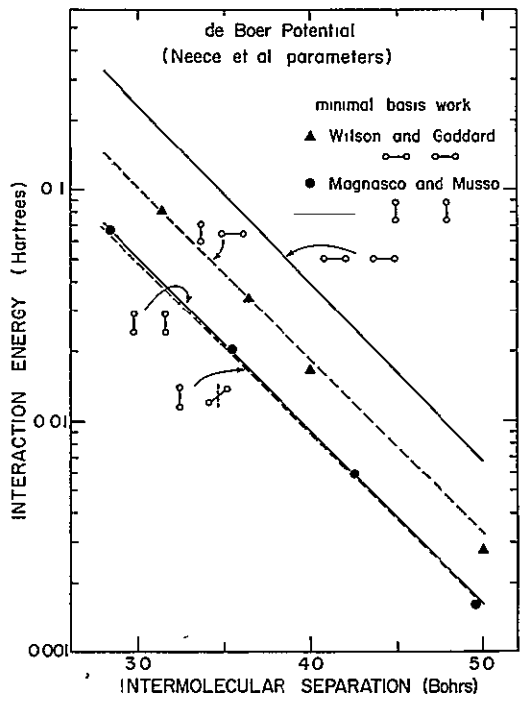


FIG. 8 de Boer potential for various geometries The de Boer potential is plotted for the choice of parameters used by Neece *et al* (Ref 8), i e , $\epsilon(R)=3\ 2e^{-1.7639R}$ [atomic units, see Eq (10)] While the curve for the rectangular case is in close agreement with the calculations of Magnasco and Musso (Ref 13), the curve for the linear case is too high by about a factor of 2 in comparison to the calculations of Wilson and Goddard (Ref 16)

where $x=R/a_0$ and $a_0=1$ bohr $=0.52917$ Å The first term is the valence energy, which they obtain by a spherical average over SCF calculations for the four standard geometries. The second term is the usual expression for the dispersion energy,[49,50] multiplied by a short-range cutoff factor as suggested by Trubitsyn.[5] From a similar spherical average of their SCF + CI calculations for the four geometries, Ree and Bender obtain

$$\Phi=(3.536e^2/a_0)\ e^{-1.242x-0.06784x^2}, \tag{12}$$

for $2.5<R<4\ 5$ bohrs As noted earlier, this expression should already include dispersion effects. A spherical average of the results illustrated in Fig. 6 may be fit by

$$\Phi=(2\ 184e^2/a_0)e^{-0.8431x-0\ 13811x^2}, \tag{13}$$

which agrees to within 10% of the Ree and Bender expression throughout the range 3–4 5 bohrs. The Evett and Margenau[37] averaging procedure yields results only a few percent different from that of Hoover *et al* ,[10] which we have used in arriving at Eq (13).

The various potentials [Eqs (11)–(13)] are shown in Fig. 9. On purely formal grounds, the extended-basis CI results [solid curves, Eqs.(12) and (13)] must be considered the most reliable determinations of the spherically averaged interaction between two hydrogen molecules. They represent agreement to within about 10% of most of the recent *ab initio* CI calculations, and incorporate the dispersion effects in a fundamental manner In contrast, the expression of Hoover *et al.* [dashed line, Eq. (11)] relies on the presumption that the standard long-range expression for the dispersion energy may also be applied for short intermolecular separations It is in fact this contribution which is responsible for the significantly weaker repulsion of Eq (11) as compared to Eqs (12) and (13). We also show in Fig. 9 the potential used by Neece *et al.*,[8] which consists of a de Boer form for the valence contribution plus the Margenau[49] result for the dispersion energy. Since their calculation of the energy of the molecular solid was based on the "α-nitrogen" structure, we have plotted their potential (dotted curve) for the near-neighbor molecular orientations of this structure This geometry is close in energy to the perpendicular case, and so the de Boer potential has significantly overestimated the repulsive energy.

In spite of the consistency evidenced amongst the recent extended-basis CI calculations for the H_2-H_2 interaction potential, there is not good agreement between theory and experiment The shaded region in Fig. 9 represents the determination by Hoover *et al.*[10] of bounds on an effective

pair potential which would be consistent with the shock experiments of Dick[28] and van Thiel *et al.*[27,29] More recently, van Thiel *et al.*[31] have reported shock experiments on deuterium which are in excellent agreement with an analysis based on Eq. (11) (dashed curve in Fig 9). Experimental determinations of the pair potential are evidently a factor of 2 or so smaller than the *ab initio* theoretical calculations. The recent work of Ree and Bender[25] suggests that this discrepancy is due to the breakdown of pairwise additivity for short-range interactions amongst hydrogen molecules in the bulk.

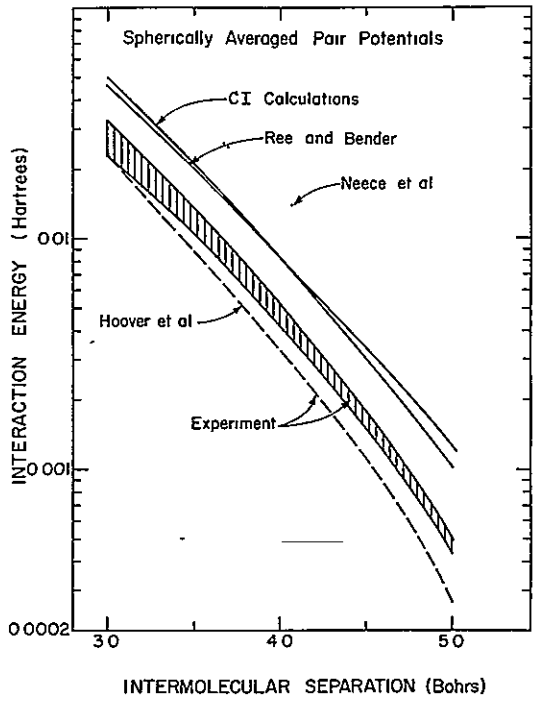


FIG. 9 Spherically averaged H_2-H_2 interaction potential The solid curves labelled "CI calculations" and "Ree and Bender" are from *ab initio* calculations, and are plots of Eqs. (13) and (12), respectively The shaded region labelled "Experiment" corresponds to the determination by Hoover *et al* (Ref. 10) of bounds on an effective pair potential consistent with shock experiment (Ref. 29). More recent shock experiments (Ref. 31) are consistent with analyses based on the dashed curve, which is a plot of Eq (11), the potential determined by Hoover *et al* The dotted curve is a plot of the potential used by Neece *et al* (Ref 8) for the molecular orientations characteristic of near neighbors in the α-nitrogen structure Calculations of the $T=0$ molecular-to-atomic phase transition pressure by Neece *et al.* (using the dotted curve), Hoover *et al* (using the upper bound to the shaded region), and by van Thiel *et al.* (Ref. 31) (using the dashed curve) yield 0 84, 1 7, and 4 2 Mbar, respectively In each case the atomic calculations of Neece *et al* were used

V APPLICABILITY TO THE SOLID

The assumption of pairwise additivity means that the behavior of a system of many molecules is characterized by a many-body potential of the form

$$V=\sum_{i<j}\Phi_{ij}\ , \tag{14}$$

where Φ_{ij} is the interaction potential for an *isolated* system of two molecules. The calculations of Ree and Bender,[25] unfortunately, point to rather large non-pairwise-additive contributions to the interaction energy of a collection of H_2 molecules for intermolecular separations less than 4.5 bohrs. A many-body potential of the form given by Eq. (14) may still be adequate, but then one must replace Φ_{ij} by some *effective* pair potential Φ_{ij}^{eff}. Ree and Bender suggest on the basis of their calculations for a system of three H_2 molecules that triplet corrections to the "bare" pair potential Φ_{ij} may be adequate to give a Φ_{ij}^{eff} in fair agreement with the phenomenological potentials for intermolecular separations down to about 3 5 bohrs.

With an eye towards calculation of the properties of the solid, the unfortunate aspect of these results is that a rigorous theoretical determination of the short-range part of the pair potential appropriate to a solid is still to be accomplished, and is now considerably more complex. It does not appear that one can avoid performing *ab initio* calculations for three and perhaps more molecules For example, one might have expected that imposition of appropriate symmetry constraints on an H_4 calculation might improve matters. As an illustration, CI calculations for the linear H_4 system permit an imbalance in the weighting of ionic configurations for the inner with respect to the outer atoms. In a solid with inversion symmetry, these must have equal weight. However, agreement of the Heitler-London results with the minimal-basis full CI results for this geometry suggests that at least in this case the matter of symmetry is not important.

A comment should be made on the applicability of the spherically averaged potential to calculations for the solid. Because of the small molecular moment of inertia and the weak angular forces, it is well known that at atmospheric pressure, the H_2 molecules in solid hydrogen are essentially freely rotating.[51] As the solid is compressed, however, the size of the anisotropic component of the interaction energy continues to increase, until eventually the molecules undergo rotational oscillations about some preferred orientations. Since the low-lying eigenfunctions of

a free rotator are sizeable throughout much of the angular phase space, in contrast to the more localized eigenfunctions of a rotational oscillator, a spherical average over the angular variables of the interaction potential is expected to be a good approximation in this limit. A rough criterion for rotational behavior would be to require that the barrier to rotation U_0 be considerably smaller than, say, the $J=1\rightarrow3$ (orthohydrogen) level spacing of the free rotator,

$$U_0 < 10\hbar^2/2I = 0.003 \text{ hartree},$$

where J and I are, respectively, the angular momentum and moment of inertia of an H_2 molecule. The overall angular variation of the interaction potential as seen in Fig. 6 is already of this order for intermolecular separations of about 5 bohrs Detailed calculations by Raich and Etters[51] place the transition from rotation to rotational oscillation at densities corresponding to a near-neighbor separation of about 4.7 bohrs. These results are based on the exaggerated angular dependence of the de Boer potential, and so it is likely that rotational behavior persists for near-neighbor separations smaller than this. The molecular phase is likely to be stable for intermolecular separations as small as 3 5 bohrs,[10] and so the spherical average is probably not always an adequate approximation for *ground-state* energy calculations Ebner and Sung,[52] in particular, have stressed the importance of retaining the anisotropic interaction in such calculations. It is felt that the spherical average is justified for the high temperatures involved in the shock experiments.[10]

As mentioned in the Introduction, one source of interest in the short-range part of the H_2-H_2 interaction potential is the desire to accurately determine the molecular-to-atomic phase-transition pressure. Qualitative aspects of this problem are evident in even the simple linear versus equidistant H_4 systems, whose energies are plotted in Fig 3. In this figure, one can identify a zero-pressure atomic phase (interatomic distance $R/2=1$ 7 bohrs) that is unstable with respect to the corresponding zero-pressure molecular phase (intermolecular separation $R=6$ 5 bohrs; the van der Waals minimum is not visible on this scale). At sufficiently high pressure, the atomic phase becomes the more stable A common tangent construction even yields a reasonable transition pressure,

$$P=\Delta E/3R^2\Delta R = 3.3\times10^{-3} \text{ a u} \simeq 1 \text{ Mbar},$$

where E is the energy per molecule and R is the intermolecular separation Turning to serious calculations, we note that Neece *et al*,[8] Hoover *et al*,[10] and van Thiel *et al*[31] have all used the *same* atomic phase calculations[8] in their determination of the transition pressure. A glance at the corresponding choices for the H_2-H_2 interaction potential thus offers an idea as to the sensitivity of the transition pressure to this choice The molecular pair potentials used are (see Fig. 9) the dotted curve, the upper bound to the shaded region, and the dashed curve, respectively. The corresponding transition pressures are 0.84, 1.7, and 4.2 Mbar respectively. Trubitsyn[5] obtained a transition pressure of 4.6 Mbar using a molecular pair potential within 20% agreement of Eq. (11) (the dashed curve, Fig 9) over the range 3–8 bohrs. If the non-pairwise-additive effects are indeed as large as suggested by Ree and Bender, then there is moderate agreement between theory and experiment, pointing to a transition pressure in the neighborhood of 4 Mbar, or larger.[31]

VI. SUMMARY

Extended-basis CI calculations which include a diffuse $2p$ orbital appear to be capable of determining the total interaction energy between two hydrogen molecules for any separation. Consistent results among a number of such *ab initio* calculations suggests that the potential is known to better than 10% for intermolecular separations ranging from 2 5–5 bohrs. For slightly smaller separations, the composite H_2 bonds are likely to become unstable The angular variation of the interaction potential in the above range is about 15%, except for geometries approaching the linear arrangement, in which case the potential may increase by about 60%. There are not yet sufficient data to determine the analytic form of this dependence, although it appears to be of relatively high order Analytic forms for a spherical average over the angular degrees of freedom are readily obtained. As a function of intermolecular separation, such potentials fall off somewhat faster than an exponential

With respect to a pair potential suitable for use in highly compressed liquid or solid molecular hydrogen, the situation is somewhat more complex. It appears that three-body corrections must be added to the bare pair potential for intermolecular separations between 3.5 and 4.5 bohrs, and that at shorter separations even higher many-body corrections may be necessary. Such corrections lead to much improved agreement between the *ab initio* calculations and analyses of shock experiments, with the implication that the $T=0$ molecular-to-atomic phase transition in solid hydrogen occurs in the neighborhood of 4 Mbar.

ACKNOWLEDGMENTS

We have benefited from a conversation with Dr. F. H. Ree, and would like to thank Dr. P. J. Hay for help in evaluating some multicenter integrals. One of us (H.B.) thanks Erziehungsrat des Kantons Zurich, Switzerland, for financing his stay at Cornell University

*This work has been supported by the National Aeronautics and Space Administration, Contract No NGR-33-010-188

†Present Address Institut fur Theoretische Physik der Universitat, Schonberggasse 9, 8001 Zurich, Switzerland

[1]R Kronig, J. de Boer, and J. Korringa, Physica (Utrecht) 12, 245 (1946).
[2]A A. Abrikosov, Astron Zh 31, 112 (1954)
[3]W C. De Marcus, Astron J 63, 2 (1958).
[4]B J Alder and R. H Christian, Phys Rev. Lett 4, 450 (1960).
[5]V. P Trubitsyn, Fiz Tverd Tela 8, 862 (1966) [Sov Phys.-Solid State 8, 688 (1966)]
[6]J C Raich, J Chem Phys 45, 2673 (1966)
[7]T Schneider, Helv Phys Acta 42, 957 (1969)
[8]G. A. Neece, F J Rogers, and W. G Hoover, J. Compt Phys 7, 621 (1971)
[9]E G Brovman, Y. Kagan, and A Kholas, Zh. Eksp. Teor. Fiz 62, 1492 (1972) [Sov Phys.—JETP 35, 783 (1972)]
[10]W G Hoover, M Ross, C F Bender, F. J Rogers, and R. J. Olness, Phys. Earth Planet Inter 6, 60 (1972)
[11]A. Wilhams, MIT Solid State and Molecular Theory Group, Quarterly Progress Report No 46, 1962, p. 150 (unpublished)
[12]V Magnasco and G. F. Musso, J. Chem. Phys. 46, 4015 (1967).
[13]V. Magnasco and G. F Musso, J. Chem Phys. 47, 1723 (1967)
[14]V. Magnasco, G. F. Musso, and R. McWeeny, J Chem Phys 47, 4617 (1967)
[15]V Magnasco and G. F. Musso, J. Chem. Phys 47, 4629 (1967).
[16]C. W. Wilson, Jr. and W. A Goddard III, J. Chem. Phys 51, 716 (1969)
[17]C. W. Wilson, Jr. and W. A. Goddard III, J. Chem Phys. 56, 5913 (1972)
[18]M Rubinstein and I. Shavitt, J Chem. Phys 51, 2014 (1969)
[19]O Tapia, G. Bessis and S Bratož, Int. J. Quantum Chem. 4, 289 (1971).
[20]O Tapia, Chem Phys Lett 10, 613 (1971).
[21]O Tapia and G Bessis, Theoret Chim Acta 25, 130 (1972)
[22]C F Bender and H. F Schaefer III, J. Chem Phys 57, 217 (1972).
[23]D. M. Silver and R. M. Stevens, J. Chem Phys. 59, 3378 (1973)
[24]E. Kochanski, B Roos, P. Siegbahn, and M H Wood, report of work prior to publication
[25]F. H. Ree and C. F. Bender, Phys Rev. Lett 32, 85 (1974).
[26]J. W Stewart, Phys. Chem Solids 1, 146 (1956)
[27]M van Thiel and B. J Alder, Mol Phys 10, 427 (1966)
[28]Reported by G. I. Kerley, "A New Model of Fluids," Los Alamos Sci. Labs., Rept LA-4760, Los Alamos, New Mexico, 1971 (unpublished).
[29]M van Thiel, M Ross, W. G Hoover, B L. Hord, W H. Gust, A Mitchell, M. D'Addario, R. N Keeler, W M. Isbell, and K Boutwell, Bull. Am Phys Soc 17, 659 (1972)
[30]F. V Grigorev, S B Kormer, O. L Mikhailova, A P Tolochko, and V D. Urlin, Pis'ma Zh. Eksp. Teor Fiz 16, 286 (1972) [JETP Lett. 16, 201 (1972)].
[31]M. van Thiel, M. Ross, B L. Hord, A C Mitchell, W H Gust, M. J D'Addario, and R N. Keeler, Phys Rev Lett 31, 979 (1973)
[32]Atomic units are used throughout this paper In these units $e=1$, $\hbar=1$, and $m_e=1$. The units of energy and length are the hartree and the bohr, respectively 1 hartree = 27 21 eV = 627 51 kcal/mole = k_B(0 3158×10^6 °K), 1 bohr = 0.52917 Å
[33]Conroy and Malli present a nonvariational solution of the molecular Schrodinger equation H Conroy and G Malli, J Chem Phys 50, 5049 (1969)
[34]H. Margenau and N R Kestner, *Theory of Intermolecular Forces*, 2nd ed (Pergamon, Oxford, 1971) This book contains explicit expressions, and an excellent discussion of the Heitler-London calculation for four hydrogen atoms (Chap. 7)
[35]Gaussian-transform method: I Shavitt and M Karplus, J Chem. Phys. 43, 398 (1965), Zeta-function expansion M P Barnett and co-workers, see C. C J Roothaan and P S Bogus, *Methods in Computational Physics* (Academic, New York, 1963), Vol. 2, p 95.
[36]J de Boer, Physica (Utrecht) 9, 363 (1942)
[37]A E Evett and H Margenau, Phys Rev 90, 1021 (1953)
[38]E A Mason and J O. Hirschfelder, J. Chem. Phys 26, 756 (1957)
[39]H F Schaefer, *The Electronic Structure of Atoms and Molecules A Survey of Rigorous Quantum-Mechanical Results*, (Addison-Wesley, Reading, Mass , 1972)
[40]W Kolos and L Wolniewicz, J Chem Phys 49, 404 (1968).
[41]F H Ree (private communication)
[42]W Kolos and C C J Roothaan, Rev. Mod Phys. 32, 219 (1960).
[43]B P Stoicheff, Can J Phys 35, 730 (1957), Adv Spectrosc 1, 91 (1959)
[44]We have evaluated the Heitler-London expression for the rectangular and linear geometries, taking $\zeta=1.193$ and $r_0=1$ 4166 bohrs For the rectangle, the exact multicenter integrals tabulated by Magnasco and Musso (Refs. 12 and 13) were used In the linear case, the integrals were evaluated using a six Gaussian expansion of the Slater function [S Huzinaga, J. Chem. Phys 42, 1293 (1965)] This procedure was tested for the rectangular case The resultant values of the interac-

tion energy were accurate to better than ±0 0001 hartree

[45]In selecting the data for Figs 5 and 6, we have presumed the results of Kochanski *et al* (Ref 24) to be the most reliable in the region 4–5 bohrs They have used the largest basis (including a diffuse $2p$ orbital), obtained the lowest total energies, and their results are in fairly close agreement with those of Bender and Schaefer (Ref 22) for the linear case Accordingly, some data of Tapia and Bessis (Ref 21) and of Silver and Stevens (Ref 23) are not shown for the region 4–5 bohrs

[46]See Ref 34, Appendix A
[47]C F Bender, H. F Schaefer, and P. A. Kollman, Mol. Phys. 24, 235 (1972)
[48]J. Hirschfelder, C Curtiss, and R Bird, *Molecular Theory of Gases and Liquids* (Wiley, New York, 1957)
[49]H Margenau, Phys Rev. 64, 131 (1943)
[50]A Dalgarno, Adv Chem Phys 12, 143 (1967)
[51]J C Raich and R D. Etters, J. Low Temp Phys 6, 229 (1972).
[52]C Ebner and C C Sung, Solid State Comm. 11, 489 (1972), Phys. Earth Planet Inter 6 83 (1972)

Conduction in fully ionized liquid metals*

D J Stevenson
Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14850

N. W Ashcroft†
Institut für Festkorperforschung, Kernforschungsanlage, Julich, Germany

(Received 10 September 1973)

Electron transport is considered in high-density fully ionized liquid metals Ionic structure is described in terms of hard-sphere-correlation functions and the scattering is determined from self-consistently screened point ions Applications to the physical properties of the deep interior of Jupiter are briefly considered

I INTRODUCTION

We are concerned here with the problem of calculating the resistivity of dense conducting fluids consisting solely of massive point ions and a neutralizing gas of interacting electrons. Several systems of physical and astrophysical interest are included in a calculation assuming the following: (i) The density of the system is such that the electrons can be treated nonrelativistically. If n_e is the electron density, this restriction can be stated as $r_s \gg 10^{-2}$, where r_s is the usual linear measure of electron density

$$n_e = (\tfrac{4}{3}\pi r_s^3 a_0^3)^{-1} .$$

(ii) The electron gas is degenerate. This is an implied restriction on the temperature, namely

$$T \ll (6\times 10^5)/r_s^2 \ K .$$

(iii) The first Born approximation is adequate for the calculation of electron scattering cross sections from the ionic system This condition is satisfied for $r_s \lesssim 1/Z$ (where $+Ze$ is the charge on the point ion) and is discussed in detail in Appendix A. At lower densities (larger r_s), the validity of the results must be viewed with the caution normally attributed to low-order calculations in liquid metals.

(iv) The density-density-correlation function (static-structure factor) of the ionic system can be approximated reasonably well by regarding the ions as an assembly of impenetrable spheres. In the presence of an electron gas (and with due account for the effects of exchange, correlation, and the adiabatic response to ionic motion), the effective ion-ion interaction is characterized at short range by a steeply repulsive region, and at long range by a weak oscillatory tail.[1] At sufficiently high density ($r_s \ll 1$), the interaction between ions is expected to depart from the hard-core model and approach the simple screened interaction following from Thomas-Fermi theory (as used by Hubbard and Lampe[2]).

(v) The contribution to the resistivity from electron-electron collisions can be neglected. So long as the electron system is highly degenerate, this assumption is reasonable.

In Sec. II we outline the basis of the calculations for the conductivity, and in subsequent sections estimate the melting temperatures of these fully ionized systems. The extensions to alloys are also discussed, and insofar as they apply the results are considered in the context of the physical properties of the deep interior of Jupiter.

II CALCULATION

Within the adiabatic approximation we may write the resistivity of the dense ionized fluid of N ions in volume Ω as

$$\rho = m/n_e e^2 \tau , \tag{1}$$

where the transport relaxation time τ is given by

$$\frac{1}{\tau} = \langle\langle \frac{2\pi}{\hbar}\frac{1}{\Omega}\int \frac{d\vec{k}'}{(2\pi)^3} |V(\vec{k}-\vec{k}')\rho_{\vec{k}-\vec{k}'}|^2 (1-\cos\theta_{\vec{k}\vec{k}'})\delta(\epsilon_F - \epsilon_{k'})\rangle\rangle , \tag{2}$$

with

$$\epsilon_{k'} = \hbar^2 k'^2/2m, \quad \epsilon_F = \hbar^2 k_F^2/2m ,$$

and

$$k_F^3 = 3\pi^2 n_e .$$

Equation (1) represents the ensemble average of the resistivity calculated in Born approximation for elastic scattering from each configuration of the ions described by the density components

$$\rho_{\vec{k}-\vec{k}'} = \sum_{i=1}^{N} e^{-i(\vec{k}-\vec{k}')\cdot\vec{R}_i} , \tag{3}$$

where $\{\vec{R}_s\}$ is the instantaneous set of ionic positions. The matrix element of the (self-consistently screened) electron-ion scattering potential $V(r)$ is defined for plane-wave levels $|\vec{k}\rangle$ by

$$\Omega\langle\vec{k}|V|\vec{k}'\rangle = V(\vec{k}-\vec{k}') = \int_\Omega d\vec{r}\, e^{-i(\vec{k}-\vec{k}')\cdot\vec{r}} V(\vec{r}) \,. \quad (4)$$

If the scattering is sufficiently weak (Appendix A), Eqs. (1) and (2) reduce, as originally shown by Ziman,[3] to

$$\rho = \frac{a_0\hbar}{e^2}\frac{4\pi^3 Z}{a_0 k_F}\int_0^1 dy\, y^3 S(y) v^2(y) \,, \quad (5)$$

where $y = |\vec{k}-\vec{k}'|/2k_F$, and $v(y)$ is the electron-ion interaction scaled to its long-wavelength limit ($\frac{2}{3}\epsilon_F$). The quantity $a_0\hbar/e^2$ may be viewed as the atomic unit of resistivity and has the convenient practical value of 21.7 $\mu\Omega$cm. $S(\vec{k}-\vec{k}')$ is the liquid-structure factor defined by

$$S(\vec{q}) = (1/N)\langle\langle \rho_q \rho_{-q}\rangle\rangle - N\delta_{\vec{q},0} \,. \quad (6)$$

In the Percus-Yevick model[4] (for hard spheres of diameter σ), $S(\vec{q})$ is a function of the packing fraction η given by

$$\eta = \tfrac{1}{6}\pi n_{\text{ion}}\sigma^3 \,, \quad n_{\text{ion}} = N/\Omega \,. \quad (7)$$

For most classical fluids near their solidification points,[4,5] $\eta \approx 0.45$.

We are dealing with point ions and the accuracy with which $v(y)$ can be specified is limited only by the uncertainties in the dielectric function $\epsilon(y)$. In the neighborhood of $y \sim 1$ [the regime dominating the integrand of (5)], $\epsilon(y)$ is quite well known and we take the interpolation form suggested by Hubbard,[6] so that

$$v(y) = -0.166 r_s/[y^2 + 0.166 r_s F(y)] \,, \quad (8)$$

where

$$F(y) = f(y)/[1 - 0.166 r_s f(y)(2y^2+g)^{-1}] \,,$$

$$g = (1 + 0.0262 r_s)^{-1} \,,$$

and

$$f(y) = \frac{1}{2} + \frac{1-y^2}{4y}\ln\left|\frac{1+y}{1-y}\right| .$$

In practice, the replacement of $F(y)$ in (8) by the Lindhard function $f(y)$ leads to the same resistivity (to within 2%), but the exchange and correlation corrections contained in $F(y)$ are important in calculations of quantities involving $[(1/\epsilon)-1]$, such as the effective pair interaction between ions.

Since $(r_s a_0 k_F)^3 = \frac{1}{4}(9\pi)$, we may rewrite (5) [using (8)] as

$$\rho/(r_s^3 Z) = 38.4 \times \int_0^1 dy\, y^3 S(y)[y^2 + 0.166 r_s F(y)]^{-2}\, \mu\Omega\text{cm} \,. \quad (9)$$

The utility of this expression is that the right-hand side is, for $r_s \lesssim 1$, a weak function of r_s and hence density. Figure 1 demonstrates this clearly. It is worth noting that the charge Z enters in the structure factor.[7]

To obtain the resistivity as a function of temperature, we require $T(\eta)$ at each density. This can

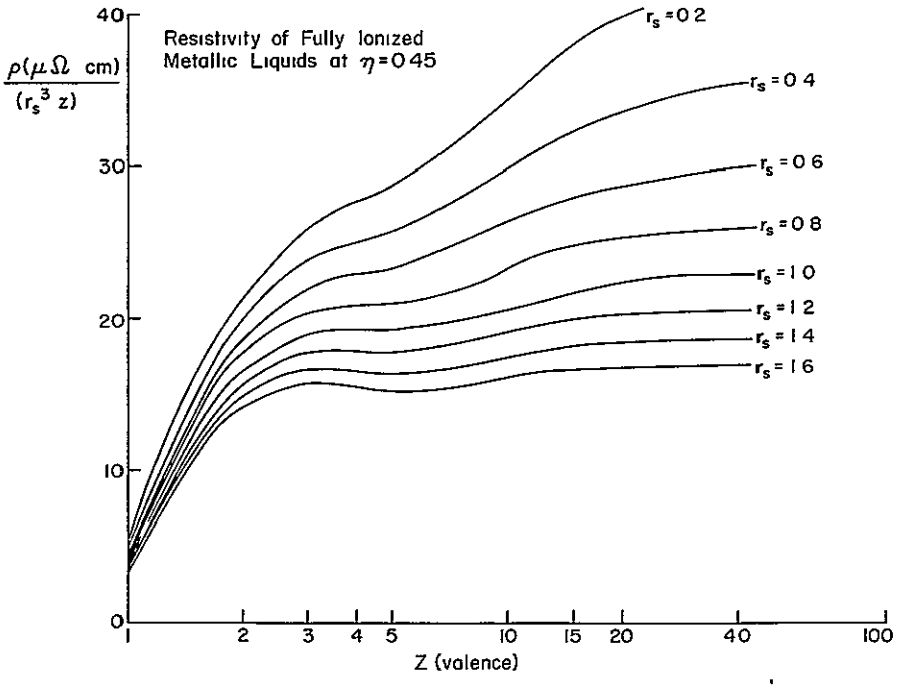


FIG. 1. Resistivity of fully-ionized liquids at $\eta = 0.45$.

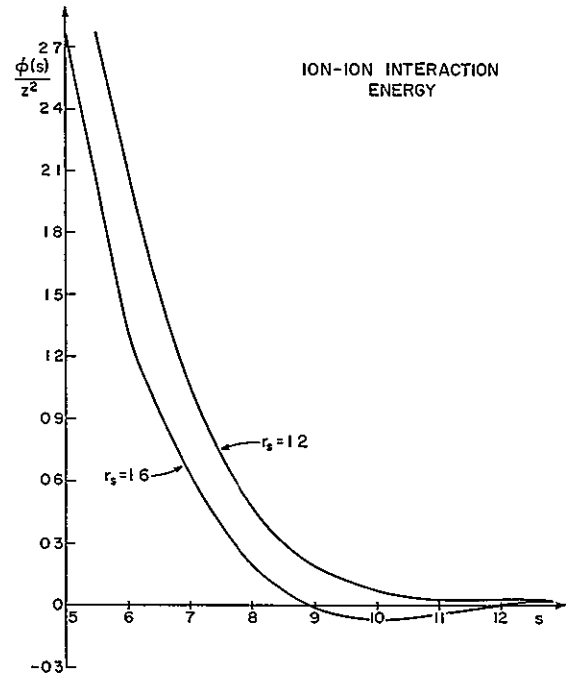


FIG. 2. Effective ion-ion interaction energy in units of $10^{-2}\epsilon_F$.

be obtained from a variational technique,[8] but the method is laborious and for the present purposes it is sufficient to use the approximate technique suggested by Ashcroft and Langreth.[1] We evaluate the pair interaction between point ions from

$$\phi(s) = (0.166 r_s)\frac{8Z^2}{s}\int_0^\infty \frac{x \sin sx}{x^2 + 0.166 r_s F(x)}dx, \tag{10}$$

which gives the pair energy at separation $r (r = s/2k_F)$ in units of ϵ_F (see Fig. 2). If $\phi_{\min}$ is the minimum value of $\phi(s)$, then the melting temperature T_M can be estimated from the relation

$$\phi(2k_F\sigma) - \phi_{\min} = \tfrac{3}{2}k_B T_M/\epsilon_F ,$$

provided $2k_F\sigma$ is evaluated at $\eta = 0.45$. It may be noted that this procedure gives T_M in sodium to within 10%. The same close agreement is not likely for fully ionized systems that have somewhat "softer" pair potentials (in reduced units) than that appropriate for sodium.[1] To find $d\eta/dT$, we evaluate the slope of $\phi(s)$

$$\frac{d\eta}{dT} = \left[9\eta/2s\cdot\left(\frac{d\phi}{ds}\right)T_F\right]_{s=2(18\pi Z\eta)^{1/3}}, \tag{11}$$

(where $T_F = (6\times10^5)/r_s^2$ K) and in this way obtain T_M (see Fig. 3) and the values of T appropriate to $\eta < 0.45$. An alternative method for obtaining T_M exploits the Lindemann rule (see Appendix B), but the simpler approach outlined above is no less accurate and is, in fact, more fundamental.

The results of our calculation for fully ionized H, He, and C are found summarized in Figs. 4, 5, and 6, respectively. We choose as a vertical axis the quantity (resistivity×density), since, as noted above, this combination, near T_M, is weak-

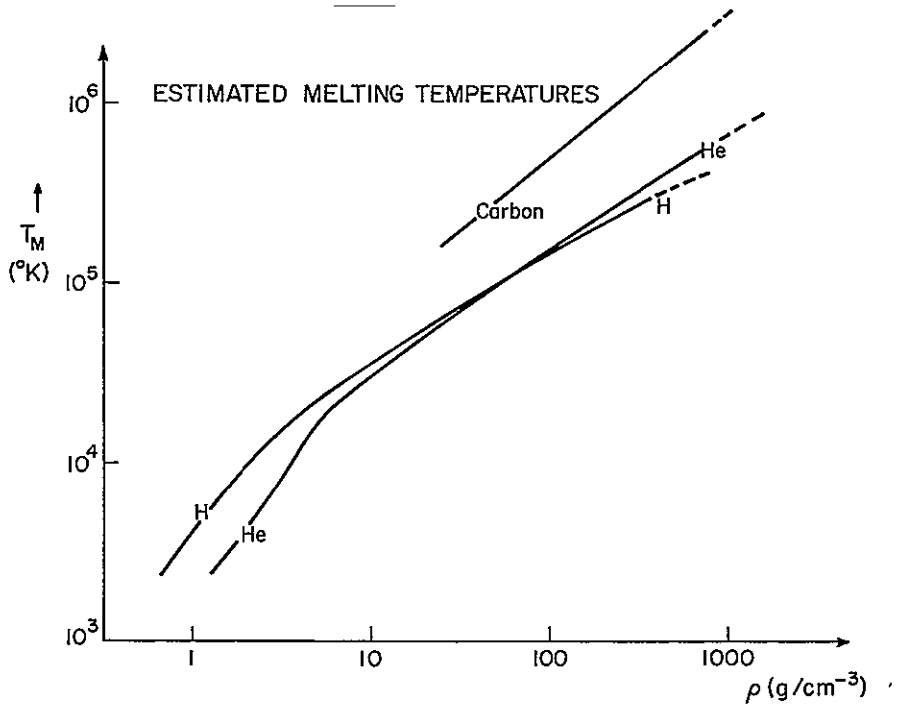


FIG. 3. Estimated melting temperatures.

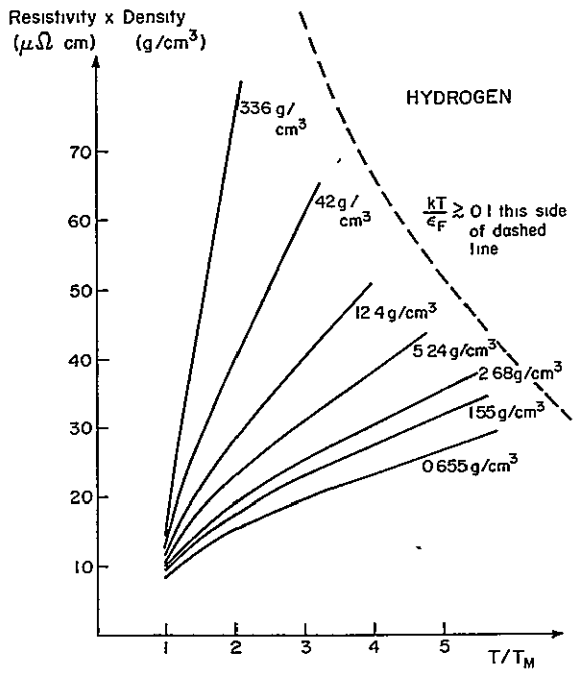


FIG. 4. Resistivity of hydrogen.

ly density dependent. It should be emphasized that if our estimates of T_M are incorrect, the form of the curves presented will remain substantially correct. We should also point out that at densities for which the element carbon is likely to be fully pressure ionized, the hard-sphere approximation to the ion-ion interaction may already depart substantially from reality.[9] Moreover, $kT_M/\epsilon_F \sim 0.05$, and this implies a significant nondegeneracy.

Figure 7 shows a comparison of our results with those of Hubbard and Lampe.[2] The quantity compared is the conductive opacity[10] as tabulated in Ref. 2. Our results are seen to be systematically lower, and the greatest difference occurs at low temperatures, where the crude approximation for $S(q)$ used in Ref. 2 is expected to be least accurate. We cannot, however, eliminate the possibility that the systematic discrepancy results from a disagreement in the temperature scale.

III EXTENSION TO ALLOYS

The extension to binary alloys is straightforward in principle.[1] The result equivalent to Eq. (5) can be written

$$\frac{\rho}{(r_s^3 Z^*)} = 38.4 \int_0^1 \frac{y^3\,dy}{[y^2 + 0.166\, r_s F(y)]^2} \times [x S_{22}(y) + 2x^{1/2}(1-x)^{1/2} S_{12}(y) + (1-x) S_{11}(y)]\ \mu\Omega\ \text{cm}, \qquad (12)$$

where x is the fractional number of ions of species 2, Z^* is the number of electrons per ion, and S_{22}, S_{12}, S_{11} are partial structure factors.[11] These structure factors not only depend on

$$\eta = \frac{\text{volume occupied by hard spheres}}{\text{total volume}}$$

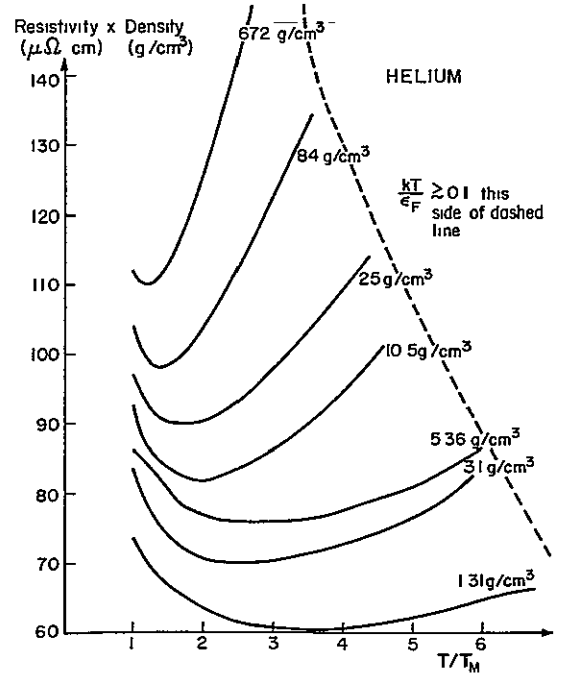


FIG. 5 Resistivity of helium.

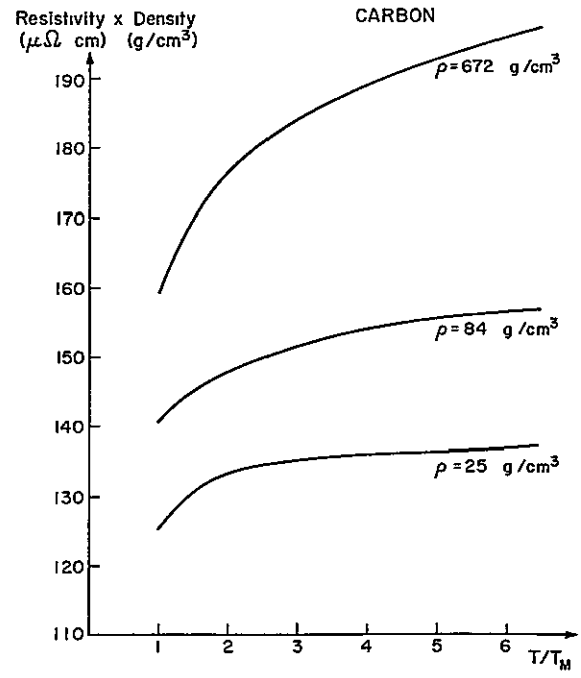


FIG. 6. Resistivity of carbon.

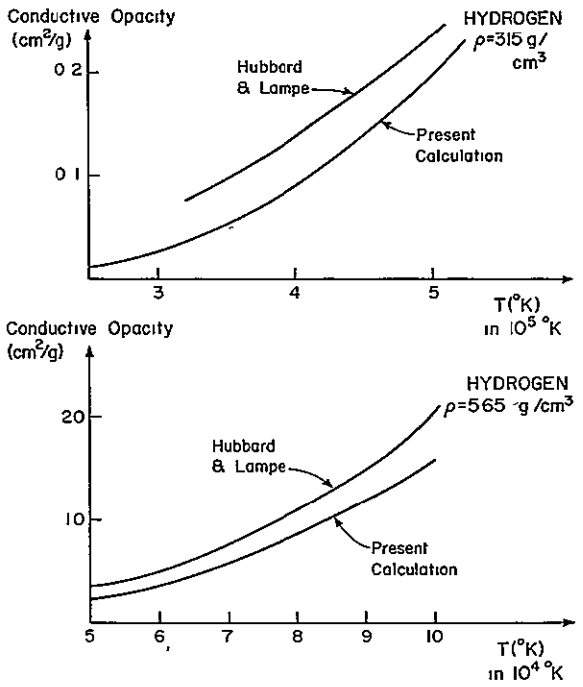


FIG. 7. Conductive opacity of hydrogen at two densities. A comparison of our results with those of Hubbard and Lampe (Ref. 2).

but also on

$$\alpha = \sigma_1/\sigma_2 ,$$

where σ_1 and σ_2 are the hard-sphere diameters of components 1 and 2, respectively.

If $\alpha = 1$, then Eq. (12) becomes identical to Eq. (9), except, of course, that Z^* is a function of x. In this special case, the results of Fig. 1 can be used to find the resistivity of any alloy[12] at the melting point.

Equation (10) shows that if $\varphi(s_0) = 0$ for the interaction between ions of species 1, then $\varphi(s_0) = 0$ for the ions of species 2. This suggests that α is near unity. However, the species with higher ionic charge is expected to have a "harder" core (for a given value of r_s). A detailed calculation[13] suggests that $\alpha = 0.75$ for a hydrogen–helium mixture; that is, the helium hard-sphere diameter is one-third larger than the hydrogen hard-sphere diameter. In Fig. 8, we show that this deviation from $\alpha = 1$ does not dramatically change the resistivity, and accordingly a reasonable approximation sets all hard-sphere diameters equal.

There is, however, no simple extension of our method for obtaining $d\eta/dT$ to the alloy problem. For $Z > 2$, the temperature dependence of the resistivity is sufficiently weak that it may be ignored in a first approximation (for $T_M < T \ll T_F$). For a hydrogen–helium alloy, a crude approximation simply interpolates between the temperature trends shown in Figs. 4 and 5.

IV SUMMARY AND APPLICATION

In the limited temperature and density range appropriate to Eq. (5) and the hard-sphere model, we find somewhat lower resistivities than those previously obtained[2] for fully ionized liquid metals. This is attributable to the use of a more accurate

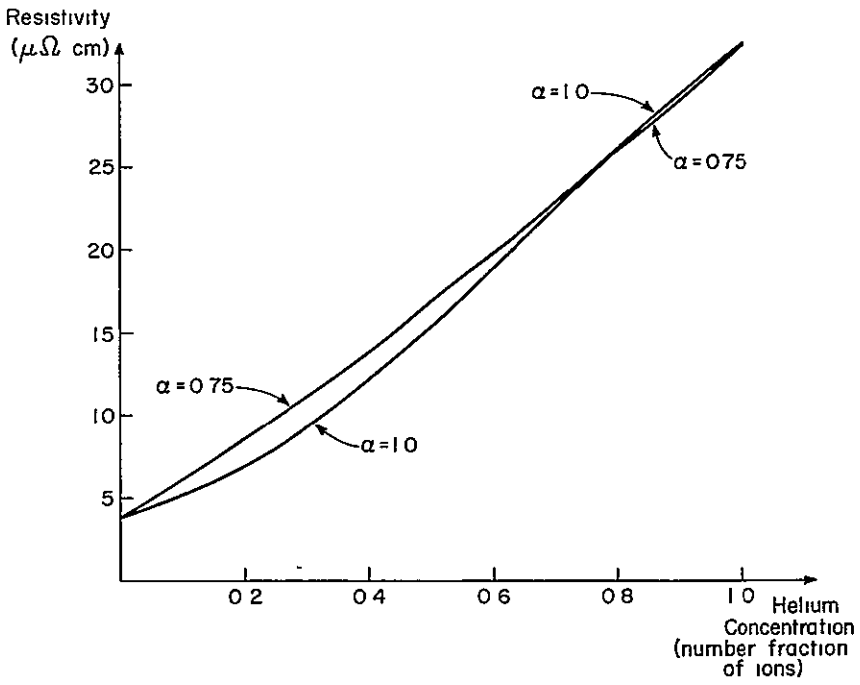


FIG. 8. Resistivity of an H-He alloy at $r_s = 1.0$ and $\eta = 0.45$. The effect of different hard-sphere diameters is shown.

electron-ion interaction and a more appropriate structure factor. A disadvantage of the present method is the need independently to estimate the temperature scale.

Systems for which the present calculations seem likely to apply include the interiors of the giant planets, in particular Jupiter. Most recent models of the Jovian interior postulate a central region of dense fluid. Its composition is predominantly metallic hydrogen, but is augmented by a small amount of helium (about 10% by number[14,15]). It is conceivable that the helium may not be completely ionized and if not, the electron–helium interaction may be more appropriate to that expected of neutral helium atoms.[16] We find that although it is possible for the resistivity to be enhanced *if* the helium remains un-ionized, this enhancement is mainly a consequence of the small increase in the value of r_s, rather than any substantial change in the scattering cross section from that expected for fully ionized atoms.

If we choose the central temperature[17] of Jupiter to be about 16 000 K, then we find that the resistivity of the fluid is expected to range from 4 $\mu\Omega$ cm at the center of Jupiter to about 8 $\mu\Omega$ cm at the boundary between metallic and molecular hydrogen. A conductivity characteristic of the deep interior of Jupiter is therefore

$$\sigma \sim 2\times 10^{17} \text{ esu} ,$$

a result somewhat larger than most previous estimates.[18]

Jupiter is observed to have a strong magnetic field, and in seeking internal mechanisms for its origin it is first of interest to decide whether the field could be primordial. If it were, then the quantity of central importance is the decay time T given in seconds by

$$T \sim 4\pi\sigma(L/c)^2 ,$$

where c is the velocity of light and L is a typical planetary dimension, which we take here as 5×10^9 cm. The result

$$T \sim 2\times 10^9 \text{ years}$$

may be seen to hinge not too seriously on the choice of L. Even if the value chosen is viewed as unreasonably large, the result for T remains such that the possibility of primordial origin is difficult to discount. In complete contrast to this, it is interesting to record that the high value of σ is likely to be favorable for a dynamo mechanism[19] underlying the generation of the magnetic field.

Finally, a straightforward application of the Wiedemann-Franz relation yields thermal conductivities for the interior of Jupiter ranging from (in erg/cm sec K) 9×10^8 at the center to 1×10^8 at the metallic boundary. Now the observed internal heat flux is very high,[20] but it is apparent that even conductivities of this magnitude are insufficient to maintain the measured flux *unless* we assume a much larger central temperature.[14,15] In a situation such as this, the system is unstable against convection, and the planet would rapidly cool. It would seem to follow that all but a small core of Jupiter must be convective. The size of this convective region is an open question.

ACKNOWLEDGMENTS

One of the authors (N. W A.) would like to thank Professor G Eilenberger and his colleagues at the KFA, Jülich, for their kind hospitality during the period when part of this work was completed.

APPENDIX A VALIDITY OF THE BORN APPROXIMATION

An elementary criterion for the validity of the Born approximation is that[2]

$$Ze^2/\epsilon_F \lesssim (\hbar^2/2m\epsilon_F)^{1/2} .$$

Here, the left-hand side is roughly the distance from the ion within which the interaction energy exceeds the Fermi energy. The right-hand side is of the order of the electron wavelength. It follows that

$$\epsilon_F \gtrsim 2Z^2 e^2/(\hbar^2/me^2) = 4Z^2 \text{ Ry},$$

whence $r_s \lesssim 1/Z$.

An alternative criterion is

$$\sigma_{\text{Born}}/4\pi a^2 \ll 1 ,$$

where $4\pi a^2$ is the "geometric" cross section. For a single ion

$$\sigma_{\text{Born}} = \frac{1}{2\pi k_F^2}\int_0^{2k_F} k\left(\frac{m}{\hbar^2}V(k)\right)^2 dk ,$$

where

$$V(k) = \frac{4\pi}{k}\int_0^\infty \sin kr\, V(r)\, r\, dr .$$

We calculate σ_{Born} approximately using Thomas-Fermi screening, i.e.,

$$V(r) = (Ze^2/r)e^{-q_s r},$$

so

$$V(k) = \frac{4\pi Ze^2}{k^2+q_s^2} = \frac{4\pi Ze^2}{x^2+0.166r_s}\,\frac{1}{(2k_F)^2} ,$$

where

$$x = k/2k_F .$$

Thus,

$$\sigma_{\text{Born}} = \frac{1}{2\pi k_F^2}\left(\frac{4\pi Z}{a_0}\right)^2 \frac{1}{(2k_F)^2}\int_0^1 \frac{x\,dx}{(x^2+0.166 r_s)^2}\,.$$

But,

$$a \simeq 1/q_s = 0.64\, r_s^{1/2} a_0\,,$$

and thus it follows that

$$\sigma_{\text{Born}}/4\pi a^2 \simeq (0.27 r_s^2 Z^2)/(1+0.166 r_s)\,.$$

Finally,

$$\sigma_{\text{Born}}/4\pi a^2 \ll 1 \text{ implies } r_s \lesssim 1/Z \text{ (as before).}$$

However, the Born cross section per ion in the condensed state is clearly different from that of a single isolated ion. We can calculate the "apparent" cross section, per ion, in the liquid by using the identity

$$n\sigma_a v_F \tau \equiv 1\,,$$

where τ is the "collision time" for an electron and n is the ion number density.

Since $\rho = m/n_{el} e^2 \tau$, we have, from Eq. (5),

$$\sigma_a = \frac{4\pi^3 Z^2}{k_F^2}\int_0^1 y^3 v^2(y) S(y)\,dy\,,$$

$$= Z(\rho(\mu\Omega\,\text{cm})/21.7)(r_s/1.92)a_0^2\,,$$

whence

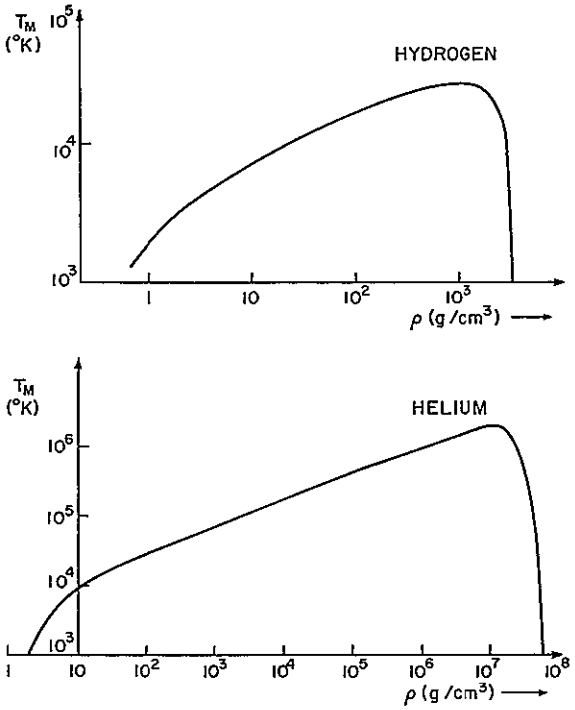


FIG. 9. The melting temperatures of metallic hydrogen and helium according to Lindemann's rule.

$$\sigma_a/4\pi a^2 \simeq 0.1Z(\rho(\mu\Omega\,\text{cm})/21.7)\,,$$

where ρ is calculated from the first Born approximation. (Note that this formula is valid for *any* simple liquid metal.)

For hydrogen at $r_s = 1.6$, $T = T_M$, we have $\sigma_a/4\pi a^2 \simeq 0.06$, and for helium at $r_s = 1.2$, $T = T_M$, we have $\sigma_a/4\pi a^2 \simeq 0.25$.

This suggests (but does not prove) that the Born approximation may be much better satisfied in the condensed state than for a single ion. Thus, our criterion $r_s \lesssim 1/Z$ may be too stringent. It is clear and expected, however, that the Born approximation is increasingly well satisfied as r_s becomes smaller.

APPENDIX B MELTING CRITERION

A commonly used criterion is Lindemann's rule. This can be written as[21]

$$\gamma = \sum_{k\lambda} \frac{\hbar(n_{k\lambda}+\frac{1}{2})}{\text{Mn}_{\text{ion}}\omega_{k\lambda}R_0^2}\,, \tag{B1}$$

where γ is the mean-square amplitude of the ions just below the melting point and is found, almost universally, to be about $\frac{1}{16}$. M is the ion mass, R_0 the interatomic spacing, $\omega_{k\lambda}$ a phonon frequency of wave vector k and polarization λ, and $n_{k\lambda}$ is the Bose-Einstein occupation factor.

For the high-density systems considered, Abrikosov[22] has shown that it is important to distinguish between the longitudinal and transverse modes, since the former are primarily determined by the bulk compressibility of the electron gas, whereas the latter are primarily determined by the Coulomb forces between ions.

We make a Debye approximation, but allow for the longitudinal and transverse "Debye" temperatures to be different Using the method outlined by Trubitsyn,[23] we obtain (in K)

$$\Theta_l \simeq \frac{2500 Z^{1/6}}{A^{1/2}}\left(\frac{22.1}{r_s^4} - \frac{3.66}{r_s^3} - \frac{7.17 Z^{2/3}}{r_s^3}\right)^{1/2}\,,$$

$$\Theta_t \simeq 8000(Z/\text{Ar}_s^3)^{1/2}\,;$$

The correlation energy of the electron gas is small and can be ignored. Equation (B1) can then be written

$$\frac{k_B\Theta_l}{MS_l^2}\left[1+4\left(\frac{T}{\Theta_l}\right)^2\int_0^{\Theta_l/T}\frac{x\,dx}{e^x-1}\right] + \frac{2k_B\Theta_t}{MS_t^2}\left[1+4\left(\frac{T}{\Theta_t}\right)^2\int_0^{\Theta_t/T}\frac{x\,dx}{e^x-1}\right] \simeq 0.47\,, \tag{B2}$$

where S_l, S_t are the appropriate sound velocities. We anticipate $T_M < \Theta_l, \Theta_t$ and so approximate Θ_l/T,

Θ_t/T by ∞ in the integrals. It is easy to show that this is valid provided

$$(\Theta/T)e^{-\Theta/T} \ll \tfrac{1}{6}\pi^2 \quad \text{for} \quad \Theta=\Theta_l \text{ and } \Theta_t \ ,$$

which is satisfied reasonably well for the cases studied. Equation (B2) can be written in numerical form, for low temperatures, as

$$\frac{0.22}{A^{1/2}Z^{5/6}[(22.1/r_s^4)-(3.66/r_s^3)-(7.17Z^{2/3}/r_s^3)]^{1/2}r_s^2}\left[1+\frac{2\pi^2}{3}\left(\frac{T}{\Theta_l}\right)^2\right]+\frac{0.13}{A^{1/2}Z^{7/6}r_s^{1/2}}\left[1+\frac{2\pi^2}{3}\left(\frac{T}{\Theta_t}\right)^2\right]\simeq 0.47$$

and is solved to obtain T_M (note that for $r_s \ll 1$, only the *transverse* modes are important in determining T_M). The results, shown in Fig. 9, give melting temperatures which differ by as much as a factor of 2 from those in Fig. 3. Similar results have been obtained by Pollack and Hansen.[24] The problem with Lindemann's rule is that an error in $\gamma(=\frac{1}{16}$ in the above calculation) propagates alarmingly through to the final calculation of T_M, in the case $T_M < \Theta_t, \Theta_l$. Typically, a 10% error in γ will give a 50% error in T_M. Moreover, our estimates of Θ_t, Θ_l are only approximate. (Our formula for Θ_l is, however, in excellent agreement with the Θ_D calculated by Neece, Rogers, and Hoover.[25]) Note that at sufficiently high densities, the zero-point motion alone will cause the lattice to melt. Lindemann's rule gives an estimate of the value of $r_s^{1/2}$ at which $T_M \to 0$. Since density varies as $(r_s^{1/2})^{-6}$, the density at which $T_M \to 0$ cannot be calculated to better than an order of magnitude using Lindemann's rule. (The pressure at which $T_M \to 0$ may be incorrect by almost two orders of magnitude.) As Abrikosov[15] observes, only hydrogen and helium will melt at absolute zero and sufficiently high densities This is because the densities required for heavier elements are such that the sizes of the *nuclei* become important.

*Supported in part by NASA under Contract No NGR-33-010-188, and by the National Science Foundation under Contract No. GH-36457.

†Permanent address: Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, N Y 14850.

[1]N. W. Ashcroft and D. C. Langreth, Phys. Rev. 159, 500 (1967).

[2]W. B. Hubbard and M. Lampe, Astrophys. J. Suppl. 18, 297 (1969).

[3]J. M. Ziman, Philos. Mag. 6, 1013 (1961).

[4]N. W. Ashcroft and J. Lekner, Phys. Rev. 145, 83 (1966).

[5]T. Wainwright and B. Alder, Nuovo Cimento Suppl. 9, 116 (1968).

[6]J. Hubbard, Proc. R. Soc. A 243, 336 (1957)

[7]S is a function only of the combination $q\sigma$. Since $q=2k_F y$, it follows from (7) that $q\sigma=2y(18\pi Z\eta)^{1/3}$ and the valence and packing fraction therefore enter in the combination $(Z\eta)^{1/3}$.

[8]D. Stroud and N. W. Ashcroft, Phys. Rev. B 5, 371 (1972).

[9]It is worth noting, however, that "softness" in the short-range interaction does not substantially alter the form of $S(k)$ [see D. Schiff and J P. Hansen, in *Proceedings of the Second International Conference on the Properties of Liquid Metals* (Taylor and Francis, London, 1973), p. 57]

[10]This follows from a simplistic application of the Wiedemann-Franz relation. The conductive opacity is proportional to the resistivity.

[11]N. W. Ashcroft and D. C. Langreth, Phys. Rev. 156, 685 (1967).

[12]Provided all spheres have the same diameters, we need not restrict ourselves to binary alloys.

[13]Following the methods of Ref. (8), the free energy of the alloy was minimized with respect to α and η, and the value of α obtained is consistent with a direct estimate from the form of the ion-ion interactions.

[14]W. B. Hubbard, Astrophys. J. 162, 687 (1970) (and other references given therein).

[15]V. P. Trubitsyn, Astron. Zh. 49, 420 (1972) [Sov. Astron.-AJ 16, 342 (1972)].

[16]Screening must be included [as in Eq. (4)]. For the unscreened interaction see, for example, J. E. Purcell, R. A. Berg, and A. E. S. Green, Phys. Rev. A 2, 107 (1970).

[17]This estimate of the temperature results from the use of an adiabatic model in which we equate the surface entropy (Ref. 18) to that appropriate to the planetary center The latter can be found by calculating the free energy of a hydrogen-helium liquid alloy using an extension of the method of Ref. (8)

[18]W. B. Hubbard, Astrophys. J. 152, 745 (1968).

[19]R. Hide, in *Magnetism and the Cosmos*, edited by W. R. Hindmarsh, F. J. Lowes, P. H. Roberts and S. K. Runcorn (American Elsevier, New York, 1965), p. 378.

[20]H. H Aumann, C. M. Gillespie, Jr., F J. Lowes, Astrophys. J. 157, L69 (1969).

[21]D. Pines, *Elementary Excitations in Solids* (Benjamin, New York, 1963), p. 19.

[22]A. A Abrikosov, Zh. Eksp. Teor. Fiz. 39, 1797 (1960) [Sov. Phys.—JETP 12, 1254 (1961)].

[23]V. Trubitsyn, Fiz Tverd. Tela 8, 862 (1966) [Sov. Phys.-Solid State 8, 688 (1966)].

[24]E. L. Pollock, J. P. Hansen (unpublished), J. P. Hansen, Physics Lett. A 41, 213 (1972).

[25]G. Neece, F. Rogers, and W. Hoover, J. Comput. Phys. 7, 621 (1971).

Ground-state energies of simple metals*

J Hammerberg and N W. Ashcroft
Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14850
(Received 11 April 1973)

A structural expansion for the static ground-state energy of a simple metal is derived Two methods are presented, one an approach based on single-particle band structure which treats the electron gas as a nonlinear dielectric, the other a more general many-particle analysis using finite-temperature perturbation theory The two methods are compared, and it is shown in detail how band-structure effects, Fermi-surface distortions, and chemical-potential shifts affect the total energy These are of special interest in corrections to the total energy beyond third order in the electron-ion interaction and hence to systems where differences in energies for various crystal structures are exceptionally small Preliminary calculations using these methods for the zero-temperature thermodynamic functions of atomic hydrogen are reported

I INTRODUCTION

Recent work in the theory of metallic phase stability has met with moderate success in accounting for the most stable crystalline structure, binding energy, and compressibility of a simple metal.[1,2] The theory depends upon a perturbation expansion of the ground-state energy ($T=0$ °K), usually to second order in the Fourier components of the pseudopotential evaluated at reciprocal-lattice vectors. In certain cases, however, the energy difference between structures is so small that it is essential to consider higher-order terms in a structural expansion for the energy. A case in point is atomic metallic hydrogen for which a second-order calculation of the ground-state energy per proton using a random-phase-approximation (RPA) dielectric function gives (static-lattice) energies of $-1.015\,32$, $-1.015\,97$, and $-1.015\,37$ Ry, respectively, for the sc, fcc, and bcc structures at a density ($r_s=1.6$) near the zero-pressure metastable equilibrium.

The procedures for constructing the perturbation expansion have been known since 1958 when Hubbard[3] developed a diagrammatic technique based upon solutions of a one-electron Hartree-like equation, a method which ultimately enabled him to express the energy in terms of the solutions to an integral equation. Later, self-consistent methods were proposed by Cohen[4] who treated the ground-state properties of a solid along the lines of the dielectric formulation of Nozières and Pines[5] for the electron gas. More recently, Brovman *et al.*[6] have used a modification of Hubbard's technique to calculate both binding energies and phonon spectra for simple metals. Lloyd and Sholl[7] have also presented explicit expressions for third-order corrections to the total energy using an analysis similar to that of Hohenberg and Kohn,[8] and Harrison[9] has discussed the interpretation of these contributions in terms of three-body interactions. What we present here is an explicit structural expansion which is convenient for calculation of ground-state energy as a function of density and which is simply related to the eigenvalues of the one-electron *band* Hamiltonian. We shall discuss its relation to a more complete solution given in terms of the $T=0$ °K limit of finite-temperature perturbation theory. Finally, we shall discuss certain differences between the present work and the previous theories mentioned above and apply these techniques to a calculation of the ground-state properties of atomic hydrogen. A comprehensive Bravais-lattice survey of the binding energy to third order in electron-ion interaction for this solid has been carried out by Brovman *et al.*[10] The purpose of our calculations is rather to study the magnitudes of higher-order corrections, in support of which we shall present numerical values for sc, fcc, and bcc lattices.

II FORMULATION OF THE PROBLEM

We consider in this section the problem of computing the total energy of a system of N interacting electrons in a static periodic one-body potential. Later[11] we shall relax this restriction and consider the modifications arising from dynamic effects. To begin with we shall restrict our considerations to $T=0$ °K and subsequently extend the analysis to nonzero temperatures.

The Hamiltonian for our system thus restricted may be written

$$H=H_{ee}+H_{ei}+H_{ii}\,, \tag{1}$$

where H_{ee} describes the kinetic and interaction energy of a system of coupled electrons, i.e.,

$$H_{ee}=\frac{1}{2m}\sum_i p_i^2+\frac{1}{2}\sum_{i,j}{}'\frac{e^2}{|\vec{r}_i-\vec{r}_j|}\;; \tag{2}$$

H_{ii} describes the interaction energy of the rigid lattice of ions of valence Z, i.e.,

$$H_{ii}=\frac{1}{2}\sum_{\alpha,\beta}{}'W(\vec{R}_\alpha,\vec{R}_\beta)\;; \tag{3}$$

and H_{ei} describes the interaction of electrons with the lattice, i.e.,

$$H_{ei}=\sum_i V(\vec{r}_i)\ . \tag{4}$$

In Eq. (3), $W(\vec{R}_1, \vec{R}_2)$ is the bare ion-ion interaction, and in Eq. (4), $V(\vec{r})$ is the periodic one-body potential. We may express H_{ee} in terms of second-quantized operators, i.e.,

$$\frac{1}{2m}\sum_i \hat{p}_i^2 \to \frac{\hbar^2}{2m}\sum_{\vec{k}} k^2 c_{\vec{k}}^\dagger c_{\vec{k}}\ , \tag{5}$$

$$\frac{1}{2}\sum_{i,j}{}' \frac{e^2}{|\vec{r}_i-\vec{r}_j|} \to \frac{1}{2}\sum_{\vec{k},\vec{k}',\vec{q}} w(\vec{q})\, c_{\vec{k}+\vec{q}}^\dagger c_{\vec{k}'-\vec{q}}^\dagger c_{\vec{k}'} c_{\vec{k}}\ , \tag{6}$$

where

$$\Omega w(q)=\int_\Omega e^{-i\vec{q}\cdot\vec{r}}\frac{e^2}{r}\,d^3r=\frac{4\pi e^2}{q^2} \tag{7}$$

is the Fourier transform of the bare Coulomb interaction (Ω being the volume of the system[12]). There is the usual problem of handling the $q=0$ term. To resolve it we carry out the following series of manipulations (the thermodynamic limit being taken as the ultimate step). First, we subtract from H_{ee} the term

$$\Omega\tfrac{1}{2}e^2\left(\frac{N}{\Omega}\right)^2\int_\Omega \frac{1}{r}\,d^3r\ , \tag{8}$$

that is, the interaction term in H_{ee} is replaced by

$$\frac{1}{2}\sum_{i,j}{}'\left(\frac{e^2}{|\vec{r}_i-\vec{r}_j|}-\frac{e^2}{\Omega}\int_\Omega \frac{1}{r}\,d^3r\right), \tag{9}$$

and H_{ee} accordingly becomes

$$H_{ee}\to\frac{1}{2m}\sum_{\vec{k}}\hbar^2k^2 c_{\vec{k}}^\dagger c_{\vec{k}} + \frac{1}{2}\sum_{\vec{k},\vec{k}',\vec{q}\neq 0} w(\vec{q})\, c_{\vec{k}+\vec{q}}^\dagger c_{\vec{k}'-\vec{q}}^\dagger c_{\vec{k}'} c_{\vec{k}}\ , \tag{10}$$

which is the familiar electron-gas Hamiltonian, denoted in what follows by H_{eg}. We now add Eq. (8) to H_{ii}. Thus,

$$H_{ii}\to\frac{1}{2}\sum_{\alpha\beta}{}' W(\vec{R}_\alpha,\vec{R}_\beta) + \frac{1}{2}\int_\Omega\frac{\rho_0^2}{|r_1-r_2|}\,d^3r_1\,d^3r_2\ , \tag{11}$$

where $\rho_0=Ne/\Omega$. The term which has been added is the self-energy of a uniform background of negative (or positive) charge. To Eq. (11) we add the interaction energy of the ions with this negative background so that H_{ii} becomes

$$H_{ii}\to\frac{1}{2}\sum_{\alpha\beta}{}' W(\vec{R}_\alpha,\vec{R}_\beta)+\frac{1}{2}\int_\Omega\frac{\rho_0^2}{r_{12}}\,d^3r_1\,d^3r_2 - \sum_\alpha\int_\Omega\frac{\rho_0}{e}W(\vec{r},\vec{R}_\alpha)\,d^3r\ , \tag{12}$$

which, if we neglect Born-Mayer terms, is just the Madelung energy for the assembly of ions. Finally, we subtract the same interaction energy from the last term in H, H_{ei}, obtaining

$$H_{ei}\to\sum_i V(\vec{r}_i)+\sum_\alpha\int_\Omega\frac{\rho_0}{e}W(\vec{r},\vec{R}_\alpha)\,d^3r\ . \tag{13}$$

The original Hamiltonian has now been separated into three well-defined parts. Taking its average over the ground state, we have as the expression for the total energy per electron

$$\frac{E}{N}=\frac{1}{N}\left(\langle H_{eg}\rangle+\langle H_{ei}\rangle\right)+E_M\ , \tag{14}$$

where E_M is the Madelung energy, i.e., the energy per electron of a lattice of positive ions in a uniform background of negative charge. Note that the first term in Eq (14) is not the energy per electron of the interacting electron gas since the ground-state wave function is that appropriate to an electron gas in which a periodic array of ions is immersed.

Let us consider the second term in more detail. For even a *simple* metal, the interaction potential $V(\vec{r})$ is not known in general from first principles. From the point of view of band theory, however, it may be well represented by a weak pseudopotential, at least for the valence states. (We set aside in this discussion questions of core-level shifts and their effect on the total energy.) If we make this pseudopotential approximation and furthermore consider a local approximation in which the periodic potential is a simple superposition of bare pseudopotentials at each lattice site, then Eq. (13) becomes

$$H_{ei}=\sum_{i\alpha}v(\vec{r}_i-\vec{R}_\alpha)+\sum_\alpha\int_\Omega\frac{Ze\rho_0}{|\vec{r}-\vec{R}_\alpha|}\,d^3r\ , \tag{15}$$

or in terms of the Fourier transform of v,

$$H_{ei}=\sum_{\vec{k},i}\rho_{\vec{k}}^{(i)}v(\vec{k})\,e^{i\vec{k}\cdot\vec{r}_i} + \sum_\alpha\int_\Omega\frac{Ze\rho_0}{r}\,d^3r\ , \tag{16}$$

where

$$\Omega v(\vec{k})=\int_\Omega v(\vec{r})e^{-i\vec{k}\cdot\vec{r}}\,d^3r \tag{17}$$

and

$$\rho_{\vec{k}}^{(i)}=\sum_\alpha e^{-i\vec{k}\cdot\vec{R}_\alpha} \tag{18}$$

In particular, the $\vec{k}=0$ term is given by

$$\lim_{\vec{k}\to 0}\left(N_iNv(\vec{k})+N_iZe\rho_0\int_\Omega\frac{d^3r}{r}\right), \tag{19}$$

where N_i is the number of ions, $N=ZN_i$. As an example, for a potential which is Coulombic beyond a certain "core" radius r_c[13]

$$\Omega \lim_{\vec{k}\to 0} v(\vec{k}) = -Ze^2 \int_{r_c}^{\infty} \frac{d^3r}{r} + \int_0^{r_c} v(\vec{r})\, d^3r \,. \quad (20)$$

Hence the long-range parts in (19) cancel and we are left with

$$H_{ei} = \sum_{\vec{k}\neq 0,\, i} \rho_{\vec{k}}^{(i)} v(\vec{k})\, e^{i\vec{k}\cdot\vec{r}_i} + N\frac{N_i}{\Omega}\int_0^{r_c} v(\vec{r})\, d^3r + NZe\rho_0 \int_0^{r_c} \frac{d^3r}{r} \,, \quad (21)$$

which we rewrite

$$H_{ei} = \sum_{i,\vec{k}}{}' \rho_{\vec{k}}^{(i)} v(\vec{k})\, e^{i\vec{k}\cdot\vec{r}_i} + NE_c \,, \quad (22)$$

where the "core" contribution[14] E_c is independent of structure, and the prime means that $\vec{k}=0$ is excluded from the summation. Thus, the ground-state energy ($T=0\,^\circ$K) can be written in the form

$$\frac{E}{N} = \frac{1}{N}\langle H_{eg}\rangle + \frac{1}{N}\left\langle \sum_{\vec{k}}{}' \rho_{\vec{k}}^{(i)} v(\vec{k})\, \rho_{-\vec{k}} \right\rangle + E_c + E_M \,, \quad (23)$$

where $\rho_{-\vec{k}} = \sum_i e^{i\vec{k}\cdot\vec{r}_i}$. For a lattice of bare protons, we note that Eq. (23) is exact with $E_c = 0$ and $v(\vec{k}) = w(\vec{k})$. However, for the general case it is approximate since it is not clear that a single-particle equation describing the band structure with a local $v(\vec{r})$ can be derived from H as given in Eq. (23) with the *same* $v(\vec{r})$. Moreover, it is not strictly correct to write the lattice potential as a simple superposition. With these reservations, we may address ourselves to the task of computing the average

$$\frac{1}{N}\left\langle H_{eg} + \sum_{\vec{k}}{}' \rho_{\vec{k}}^{(i)} v(\vec{k})\, \rho_{-\vec{k}} \right\rangle \,. \quad (24)$$

If we treat

$$H_1 \equiv \sum_{\vec{k}}{}' \rho_{\vec{k}}^{(i)} v(\vec{k})\, \rho_{-\vec{k}}$$

as a perturbation, then the unperturbed problem *is* the interacting electron gas. Indeed, the problem is that of a dense distribution of identical impurities in the electron gas except that for a crystal, the impurities are arrayed in a definite order. Alternatively, one may simultaneously treat both electron-electron and electron-ion interactions as perturbations and carry out the usual double-perturbation expansion. In the following sections we present two methods for computing the energy shift due to H_1, one closely related to a single-particle picture, the other a more general many-particle method.

III BAND APPROACH

In this section we consider the calculation of the ground-state energy from a single-particle point of view. The physical picture is the following. We have a system of electrons whose interactions with the static lattice are described by a pseudopotential. The electron gas may be viewed as a nonlinear dielectric and the pseudoions as the source of external potentials which induce charge density responses in it. The energy associated with this induction process is given by the well-known expression[15]

$$\delta W = \int \delta V(\vec{r})\, \rho(\vec{r})\, d^3r \,. \quad (25)$$

This is the work which an external contrivance must do in changing the potentials from some value V to $V+\delta V$. In terms of Fourier transformed quantities this becomes

$$\delta W = \Omega \sum_{\vec{k}} \delta V(-\vec{k})\, \rho(\vec{k}) \,. \quad (26)$$

The contribution of the electron-ion interaction to the total energy is then given by

$$W = \int_0^V \delta W \,. \quad (27)$$

In general, we may write the averaged number density $\rho(\vec{k})$ as

$$\rho(\vec{k}) = \chi_1(\vec{k})V(\vec{k}) + \sum_{\vec{q}} \chi_2(\vec{k},\vec{q})\, V(\vec{k}+\vec{q})\, V(-\vec{q}) + \sum_{\vec{q}_1,\vec{q}_2} \chi_3(\vec{k},\vec{q}_1,\vec{q}_2)\, V(\vec{k}+\vec{q}_1+\vec{q}_2) \times V(-\vec{q}_1)V(-\vec{q}_2) + \cdots \,, \quad (28)$$

the first term of which is the usual linear-response expression. It is easy to show that this leads to an expression for the change in energy given by

$$\frac{N}{\Omega}E_b \equiv \frac{W}{\Omega} = \frac{1}{2}\sum_{\vec{k}} \chi_1(\vec{k})V(-\vec{k})V(\vec{k}) + \frac{1}{3}\sum_{\vec{k},\vec{q}} \chi_2(\vec{k},\vec{q})V(\vec{k}+\vec{q})V(-\vec{q})V(-\vec{k}) + \frac{1}{4}\sum_{\vec{k},\vec{q}_1,\vec{q}_2} \chi_3(\vec{k},\vec{q}_1,\vec{q}_2)V(\vec{k}+\vec{q}_1+\vec{q}_2) \times V(-\vec{q}_1)V(-\vec{q}_2)V(-\vec{k}) + \cdots \,, \quad (29)$$

which we shall refer to as the *band-structure energy*[16] and which is determined from the induced charge density through Eqs. (28). Note that $V(0)$ is to be excluded from the summation (a requirement of charge neutrality as discussed in Sec. II). Equation (29) thus presents us with a well-defined method for calculating E_b in terms of the charge density.

From the point of view of single-particle band theory we calculate the charge density from the Bloch wave function of an electron in a periodic potential $\tilde{V}$. In terms of plane waves

$$\langle \vec{r}\,|\vec{k}-\vec{K}\rangle = \Omega^{-1/2} e^{i(\vec{k}-\vec{K})\cdot\vec{r}} \,, \quad (30)$$

the wave function is written (we assume a Bravais lattice)

$$|\psi_{\vec{k}}\rangle = \sum_{\vec{K}} c_{\vec{k}-\vec{K}} |\vec{k}-\vec{K}\rangle \equiv \sum_i c_i |i\rangle , \tag{31}$$

where the coefficients c_i satisfy the equations

$$\begin{aligned}
&(\mathcal{E}_0 - E + \tilde{V}_{00})c_0 + \tilde{V}_{01}c_1 + \sum_{i\neq 0,1} \tilde{V}_{0i}c_i = 0 ,\\
&\tilde{V}_{10}c_0 + (\mathcal{E}_1 - E + \tilde{V}_{11})c_1 + \sum_{i\neq 0,1} \tilde{V}_{1i}c_i = 0 ,\\
&\cdots \qquad\qquad \cdots \\
&\tilde{V}_{i0}c_0 + \tilde{V}_{i1}c_1 + (\mathcal{E}_i - E + \tilde{V}_{ii})c_i \\
&\qquad\qquad + \sum_{j\neq 0,1,i} \tilde{V}_{ij}c_j = 0 ,\\
&\cdots \qquad\qquad \cdots ,
\end{aligned} \tag{32}$$

with $\mathcal{E}_i = (\hbar^2/2m)(\vec{k}-\vec{K}_i)^2$ and $\tilde{V}_{ij} = \langle \vec{k}-\vec{K}_i | \tilde{V} | \vec{k}-\vec{K}_j \rangle$. An iterative solution of Eqs. (32) yields a Brillouin-Wigner expansion for the c_i, namely,

$$\begin{aligned}
c_i = c_0 &\left(\frac{\tilde{V}_{i0}}{E-\mathcal{E}_i} + \sum_j{}' \frac{\tilde{V}_{ij}\tilde{V}_{j0}}{(E-\mathcal{E}_i)(E-\mathcal{E}_j)} \right.\\
&\left. + \sum_{jk}{}' \frac{\tilde{V}_{ij}\tilde{V}_{jk}\tilde{V}_{k0}}{(E-\mathcal{E}_i)(E-\mathcal{E}_j)(E-\mathcal{E}_k)} + \cdots \right)\\
+ c_1 &\left(\frac{\tilde{V}_{i1}}{E-\mathcal{E}_i} + \sum_j{}' \frac{\tilde{V}_{ij}\tilde{V}_{j1}}{(E-\mathcal{E}_i)(E-\mathcal{E}_j)} \right.\\
&\left. + \sum_{jk}{}' \frac{\tilde{V}_{ij}\tilde{V}_{jk}V_{k1}}{(E-\mathcal{E}_i)(E-\mathcal{E}_j)(E-\mathcal{E}_k)} + \cdots \right),
\end{aligned} \tag{33}$$

where the prime excludes 0, 1. Equation (33) leads to folded secular equations

$$\begin{aligned}
&(\mathcal{E}_0 - E + U_{00})c_0 + U_{01}c_1 = 0 ,\\
&U_{10}c_0 + (\mathcal{E}_1 - E + U_{11})c_1 = 0 ,
\end{aligned} \tag{34}$$

with the U's defined by

$$U_{lm} = \tilde{V}_{lm} + \sum_i{}' \frac{\tilde{V}_{li}\tilde{V}_{im}}{E-\mathcal{E}_i}$$

$$+ \sum_{ij}{}' \frac{\tilde{V}_{li}\tilde{V}_{ij}V_{jm}}{(E-\mathcal{E}_i)(E-\mathcal{E}_j)} + \cdots . \tag{35}$$

In Eq. (35) the prime excludes l, m from the summations. Note that although $U^*_{lm} = U_{ml}$, $U_{l-m,0} \neq U_{lm}$. The folding transformation is valid for any l, m and accordingly,

$$\begin{aligned}
&(\mathcal{E}_l + U_{ll} - E)c_l + U_{lm}c_m = 0 ,\\
&U_{ml}c_l + (\mathcal{E}_m + U_{mm} - E)c_m = 0 .
\end{aligned} \tag{36}$$

These equations define a two-band (upper denoted by superscript $^{(+)}$ and lower by $^{(-)}$) situation for which the solution for $\vec{K}_1 \Rightarrow 0$ is

$$\begin{aligned}
E_{(-)} &= \mathcal{E}_0 + U^{(-)}_{00}\\
&\quad + |U^{(-)}_{0m}| \{\gamma^{(-)}_m - [1+(\gamma^{(-)}_m)^2]^{1/2}\} ,\\
\gamma^{(-)}_m &\equiv (2|U^{(-)}_{0m}|)^{-1}(\mathcal{E}_m - \mathcal{E}_0 + U^{(-)}_{mm} - U^{(-)}_{00}) ,
\end{aligned} \tag{37}$$

$$c^{(-)}_m = \frac{U^{(-)}_{m0}c^{(-)}_0}{\mathcal{E}_m + U^{(-)}_{mm} - E_{(-)}} = -\frac{\mathcal{E}_0 + U^{(-)}_{00} - E_{(-)}}{U^{(-)}_{0m}} c^{(-)}_0 ,$$

or

$$c^{(-)}_m = \frac{|U^{(-)}_{0m}|}{U^{(-)}_{0m}} \{\gamma^{(-)}_m - [1+(\gamma^{(-)}_m)^2]^{1/2}\} c^{(-)}_0 . \tag{38}$$

A similar expression holds for the upper band with $(-) \rightarrow (+)$ and $\{\gamma^{(-)}_m - [1+(\gamma^{(-)}_m)^2]^{1/2}\} \rightarrow \{\gamma^{(+)}_m + [1+(\gamma^{(+)}_m)^2]^{1/2}\}$. We may use these results to calculate a number density, i.e.,

$$\rho(\vec{r}) = 2\sum_{[\vec{k}]} \sum_{ij} c^*_i c_j \langle i|\vec{r}\rangle\langle\vec{r}|j\rangle , \tag{39}$$

where $\sum_{[\vec{k}]}$ denotes a summation restricted to *occupied* levels. The Fourier transform of Eq. (39) gives

$$\rho_l = \frac{2}{\Omega} \sum_{[\vec{k}]} \sum_i c^*_i c_{i-l} , \tag{40}$$

which for a single occupied band reads

$$\rho_l = \frac{2}{\Omega} \sum_{[\vec{k}]} \sum_i [\gamma_i - (1+\gamma_i^2)^{1/2}][\gamma_{i-l} - (1+\gamma^2_{i-l})^{1/2}] \Big/ \left(1 + \sum_{j\neq 0} [\gamma_j - (1+\gamma_j^2)^{1/2}]^2\right) . \tag{41}$$

Alternatively, this may be rewritten using Eq. (37) as

$$\rho_l = \frac{2}{\Omega} \sum_{[\vec{k}]} \left(\frac{2U_{0l}}{E_{\vec{k}} - \mathcal{E}_l - U_{ll}} + \sum_{i\neq 0,l} \frac{U_{i-l,0}U^*_{i,0}}{(E_{\vec{k}} - \mathcal{E}_{i-l} - U_{i-l,i-l})(E_{\vec{k}} - \mathcal{E}_i - U_{ii})} \right) \Big/ \left(1 + \sum_{j\neq 0} \frac{|U_{j0}|^2}{(E_{\vec{k}} - \mathcal{E}_j - U_{jj})^2}\right) . \tag{42}$$

These last two expressions are easily generalized if two bands are occupied. If more than two bands are occupied it is necessary to begin with the folded secular equation appropriate to that number. We note again that in Eq (41) the $\vec{k}$ summation is only over occupied levels. Thus we are summing up to the *true* Fermi surface rather than within a Fermi sphere (the more common situation in perturbation theory).

The above expressions, although formally exact within the one-electron approximation, are difficult to use in practice. If we knew the analytic dependence of the U's on $\vec{V}$, we could perform the integration in Eq. (27) (for example, by associat-

ing with $\tilde{V}$ a coupling constant over which we ultimately integrate) and finally carry out the sum on $\vec{k}$. However, only in the extreme approximation of retaining a single γ is this analytically tractable. We can, on the other hand, expand the expression for ρ_I in powers of $\tilde{V}$. If we then assume that $\tilde{V}(\vec{k}) = V(\vec{k})/\epsilon(\vec{k})$, where $\epsilon(\vec{k})$ is the static limit of the electron-gas dielectric function, it is possible to calculate the energy shift from Eq. (29). The results of such an expansion are given in Appendix B. In Sec. IV we shall derive an expression for the energy shift from a more complete theory and see that the simple theory above must be only slightly modified.

IV FINITE-TEMPERATURE RESULTS

In this section we calculate the total energy of the system of electrons and ions using the techniques of finite-temperature perturbation theory. If we choose as the unperturbed system one having a spherical Fermi surface (e g., noninteracting *or* interacting electron gas), it is, in fact, *necessary* to use this method, a consequence of the fact that for interacting electrons in a periodic potential, the adiabatically generated state of the zero-temperature method is not the true ground state, no matter how weak the lattice potential. The state generated adiabatically from a spherical ground state can never depart from a state with a spherical Fermi surface and cannot produce the crossing of levels[17,18] resulting from the imposition of a periodic potential. In the finite-temperature theory, however, the mean occupation number of a given quantum state is no longer restricted to be either 0 or 1. Thus the Fermi surface of the unperturbed system is permitted to distort in such a way that the thermodynamic potential is minimized subject to the constraint of fixed overall density and in consequence the *true* Fermi surface is attained at each stage of the calculation.

The temperature formalism is most simply stated in terms of Green's functions. We shall follow the exposition of Martin and Schwinger[19] and define the single-particle Green's function as

$$G_{\alpha\beta}(\vec{r}_1, t_1; \vec{r}_2, t_2) = (-i)\langle T\psi_\alpha(\vec{r}_1, t_1)\psi_\beta^\dagger(\vec{r}_2, t_2)\rangle , \quad (43)$$

where the angular brackets denote the grand canonical ensemble average

$$\langle O\rangle = \mathrm{Tr}e^{-\beta(H-\mu N)}O/\mathrm{Tr}e^{-\beta(H-\mu N)} , \quad (44)$$

ψ, $\psi^\dagger$ are Heisenberg field operators, α, β are spin indices; and T is the time ordering operator for real t (and the it ordering operator for imaginary times). We Fourier transform G and write the result itself as a Fourier series

$$G_{\alpha\beta}(\vec{p}_1, \vec{p}_2; t) = \frac{1}{\Omega^2}\int_\Omega d^3r_1 d^3r_3 \times e^{-i\vec{p}_1\cdot\vec{r}_1}e^{i\vec{p}_2\cdot\vec{r}_2}G_{\alpha\beta}(\vec{r}_1, \vec{r}_2, t) , \quad (45)$$

$$G_{\alpha\beta}(\vec{p}_1, \vec{p}_2; t) = \frac{1}{-i\beta} \times \sum_\nu e^{-i\omega_\nu t}G_{\alpha\beta}(\vec{p}_1, \vec{p}_2, \omega_\nu) \quad (0 \le it \le \beta) , \quad (46)$$

where $\omega_\nu = (\pi/-i\beta)(2\nu+1) + \mu$, $\nu = 0, \pm 1, \ldots$, so that

$$G_{\alpha\beta}(\vec{p}_1, \vec{p}_2; \omega_\nu) = \int_0^{(-i\beta)} G_{\alpha\beta}(\vec{p}_1, \vec{p}_2, t)dt . \quad (47)$$

These results are consequences of the boundary condition satisfied by G for imaginary times. The average value of a one-body operator is given in terms of the Green's function by

$$\langle V\rangle = \frac{1}{\beta}\sum_{\alpha,\vec{p},\vec{k},\nu} V(-\vec{p})\Omega G_{\alpha\alpha}(\vec{k}, \vec{k}-\vec{p}, \omega_\nu)e^{\omega_\nu 0^+} . \quad (48)$$

In order to compute the ground-state energy we use the statistical mechanical theorem[20] which states that for any parameter λ in the Hamiltonian,

$$\frac{\partial\Xi}{\partial\lambda} = \left\langle\frac{\partial H(\lambda)}{\partial\lambda}\right\rangle , \quad (49)$$

where Ξ is the thermodynamic potential, the differentiation is at fixed T, Ω, μ; and the average is that defined in Eq. (44). For the Hamiltonian we take that given in Eq. (24). If we associate a coupling constant λ with the bare interaction V, we then have, upon integration,

$$\Xi(\mu) = \Xi_0(\mu) + \int_0^1 \frac{d\lambda}{\lambda}\langle\lambda V\rangle_\lambda . \quad (50)$$

To calculate the ground-state energy we take the $T = 0$ limit of $\Xi(\mu) + \mu\bar{N}$, i.e.,

$$\bar{N}E' = \lim_{T\to 0}\left([\Xi_0(\mu) + \mu\bar{N}] + \int_0^1 \frac{d\lambda}{\lambda}\langle\lambda V\rangle_\lambda\right) \quad (51)$$

which we write

$$E' = E_0 + E_b ,$$
$$E_0 = \frac{1}{\bar{N}}\lim_{T\to 0}[\Xi_0(\mu) + \mu\bar{N}] , \quad (52)$$
$$E_b = \frac{1}{\bar{N}}\lim_{T\to 0}\int_0^1 \frac{d\lambda}{\lambda}\langle\lambda V\rangle_\lambda .$$

The ground-state energy is then

$$E = E_0 + E_b + E_M + E_c . \quad (53)$$

We note that E_0 is not the ground-state energy of the electron gas at density $\bar{N}/\Omega$ since the chemical potential μ is that appropriate to the complete system, namely, electrons *and* ions. But E_b has the same form as that derived in Sec. III, for we may expand $G(\vec{k}, \vec{q}, \omega_\nu)$ in a Laurent series:

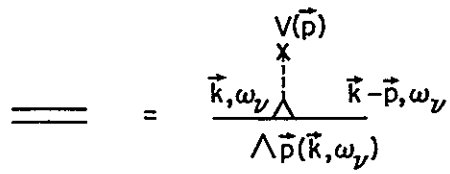


FIG. 1. First-order correction to the Green's function. The solid line represents the electron-gas Green's function, the dashed line is the bare external potential, and the triangle is the vertex function of the electron gas.

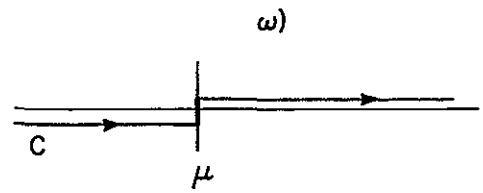


FIG. 2. Integration contour for Eq. (59).

$$\Omega G(\vec{k}, \vec{q}; \omega_\nu) = \sum_{n=0}^{\infty} \lambda^n G^{(n)}(\vec{k}, \vec{q}, \omega_\nu) , \tag{54}$$

so that, using Eq. (48),

$$\langle \lambda V \rangle_\lambda = \frac{2}{\beta} \sum_{\vec{p},\vec{k},\nu} \sum_{n=0}^{\infty} V(-\vec{p}) \lambda^{n+1} \times G^{(n)}(\vec{k}, \vec{k}-\vec{p}, \omega_\nu) e^{\omega_\nu 0^+} \tag{55}$$

and the expression for E_b now reads[21]

$$E_b = \lim_{T\to 0} \frac{1}{N} \sum_{n=0}^{\infty} \frac{1}{n+2} \frac{2}{\beta} \sum_{\vec{p},\vec{k},\nu} V(-\vec{p}) \times G^{(n+1)}(\vec{k}, \vec{k}-\vec{p}; \omega_\nu) e^{\omega_\nu 0^+} , \tag{56}$$

which is of the same form as Eq (29) and constitutes a more formal derivation of it.

In order to calculate E_b we need explicit expressions for the quantities $G^{(n+1)}$. Considering the lowest-order term we note that $G^{(1)}$ can be calculated in terms of known electron-gas quantities. We have

$$G^{(1)}(\vec{k}, \vec{k}-\vec{p}, \omega_\nu) = G^{(0)}(\vec{k}, \omega_\nu) V(\vec{p}) \times \Lambda_{\vec{p}}(\vec{k}, \omega_\nu) G^{(0)}(\vec{k}-\vec{p}, \omega_\nu) , \tag{57}$$

which is shown graphically in Fig. 1. Here $G^{(0)}(\vec{k}, \omega_\nu)$ is the Green's function of the interacting electron gas and $\Lambda_{\vec{p}}(\vec{k}, \omega_\nu)$ is the zero-frequency vertex function.[22] The second-order term in the band-structure energy is then from Eq. (56),

$$E_b^{(2)} = \frac{1}{N} \lim_{T\to 0} \frac{1}{\beta} \sum_{\vec{k},\vec{p},\nu} |V(\vec{p})|^2 \Lambda_{\vec{p}}(\vec{k}, \omega_\nu) \times G^{(0)}(\vec{k}, \omega_\nu) G^{(0)}(\vec{k}-\vec{p}, \omega_\nu) , \tag{58}$$

which, upon transforming to a contour integral, gives

$$E_b^{(2)} = \frac{1}{N} \sum_{\vec{p},\vec{k}} |V(\vec{p})|^2 \int_C \Lambda_{\vec{p}}(\vec{k}, \omega) \times G^{(0)}(\vec{k}, \omega) G^{(0)}(\vec{k}-\vec{p}, \omega) d\omega , \tag{59}$$

where C is the contour of Fig. 2. From the *definition* of the zero frequency dielectric function of the electron gas[22] we therefore have

$$E_b^{(2)} = \frac{1}{2} \frac{1}{N} \sum_{\vec{p}} |V(\vec{p})|^2 \frac{1}{w(\vec{p})} \left(\frac{1}{\epsilon(\vec{p}, 0; \mu)} - 1 \right) , \tag{60}$$

with $w(\vec{p})$ defined in Eq. (7) and μ being the *exact* chemical potential

The higher-order terms in the expansion of G are, on the other hand, not well known, and the analogues of $\Lambda_{\vec{p}}(\vec{k}, \omega)$ must be approximated.

We illustrate our approximation by recalculating Eq. (58). Using the spectral resolution of $G^{(0)}(\vec{p}, \omega)$, i.e.,

$$G^{(0)}(\vec{p}, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega' \frac{A(\vec{p}, \omega')}{\omega - \omega'} , \tag{61}$$

we have

$$E_b^{(2)} = \frac{1}{N} \lim_{T\to 0} \frac{1}{\beta} \sum_{\vec{p},\vec{k},\nu} \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} |V(\vec{p})|^2 \times \Lambda_{\vec{p}}(\vec{k}, \omega_\nu) \frac{A(\vec{k}, \omega_1)}{\omega_\nu - \omega_1} \frac{A(\vec{k}-\vec{p}, \omega_2)}{\omega_\nu - \omega_2} , \tag{62}$$

which, exploiting a further transformation of the ν sum to a contour integral gives two contributions from the simple poles·

$$E_b^{(2)} = \frac{1}{N} \lim_{T\to 0} \sum_{\vec{p},\vec{k}} \int \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} |V(\vec{p})|^2 A(\vec{k}, \omega_1) A(\vec{k}-\vec{p}, \omega_2) \left(\frac{\Lambda_{\vec{p}}(k, \omega_1)}{\omega_1 - \omega_2} n(\omega_1) + \frac{\Lambda_{\vec{p}}(k, \omega_2)}{\omega_2 - \omega_1} n(\omega_2) \right), \tag{63}$$

where $n(\omega) = (e^{\beta(\omega-\mu)} + 1)^{-1}$. Our first approximation is to make an undamped quasiparticle *Ansatz* for the spectral function, i.e.,

$$A(\vec{p}, \omega) = 2\pi \delta(\omega - \mathcal{E}_0(\vec{p}) - \Sigma_1(\vec{p})) , \tag{64}$$

with $\Sigma_1(\vec{p})$ defined to be the real part of the self-energy satisfying Dyson's equation $\mathcal{E}(\vec{p}) = \mathcal{E}_0(p) + \Sigma_1(\vec{p}, \mathcal{E}(\vec{p}))$. Then the right-hand side of Eq. (63) becomes

$$\frac{1}{\bar{N}}\sum_{\vec{p},\vec{k}}|V(\vec{p})|^2\left(\frac{\Lambda_{\vec{p}}(\vec{k},\,\mathcal{E}_0(\vec{k})+\Sigma_1(\vec{k}))\,\theta\,(\mu-\mathcal{E}_0(\vec{k})-\Sigma_1(\vec{k}))}{\mathcal{E}_0(\vec{k})+\Sigma_1(\vec{k})-\mathcal{E}_0(\vec{k}-\vec{p})-\Sigma_1(k-p)}\right.$$

$$\left.-\frac{\Lambda_{\vec{p}}(\vec{k},\,\mathcal{E}_0(\vec{k}-\vec{p})+\Sigma_1(\vec{k}-\vec{p}))\,\theta(\mu-\mathcal{E}_0(\vec{k}-\vec{p})-\Sigma_1(\vec{k}-\vec{p}))}{\mathcal{E}_0(\vec{k})+\Sigma_1(\vec{k})-\mathcal{E}_0(\vec{k}-\vec{p})-\Sigma_1(\vec{k}-\vec{p})}\right). \tag{65}$$

Our second approximation is to neglect vertex corrections and replace $\Lambda_{\vec{p}}(\vec{k},\,\omega)$ in these expressions by $\epsilon^{-1}(\vec{p},\,0,\,\mu)$.[23] We then have

$$\frac{1}{\bar{N}}\sum_{\vec{p},\vec{k}}|V(\vec{p})|^2\,\frac{1}{\epsilon(\vec{p},\,0,\,\mu)}\left(\frac{\theta(\mu-\mathcal{E}_0(\vec{k})-\Sigma_1(\vec{k}))-\theta(\mu-\mathcal{E}_0(\vec{k}-\vec{p})-\Sigma_1(\vec{k}-\vec{p}))}{[\mathcal{E}_0(\vec{k})-\mathcal{E}_0(\vec{k}-\vec{p})]+[\Sigma_1(\vec{k})-\Sigma_1(\vec{k}-\vec{p})]}\right). \tag{66}$$

Furthermore, we write the chemical potential as

$$\mu=\mu_{eg}+\delta\mu^b\,, \tag{67}$$

where μ_{eg} is the chemical potential of an electron gas of density $\bar{N}/\Omega$, i.e.,

$$\mu_{eg}=\mathcal{E}_F+\Sigma_1(k_F\,,\,\mu_{eg})\,, \tag{68}$$

with $k_F^3=3\pi^2\bar{N}/\Omega$ and $\mathcal{E}_F=(\hbar^2/2m)k_F^2$ so that (66) becomes

$$\frac{1}{\bar{N}}\sum_{\vec{p},\vec{k}}|V(\vec{p})|^2\,\frac{1}{\epsilon(\vec{p},\,0,\,\mu)}\{\theta(\mathcal{E}_F+\delta\mu^b-\mathcal{E}_0(\vec{k})-[\Sigma_1(\vec{k},\,\mathcal{E}(\vec{k}))-\Sigma_1(k_F\,,\,\mu_{eg})])$$

$$-\theta(\mathcal{E}_F+\delta\mu^b-\mathcal{E}_0(\vec{k}-\vec{p})-[\Sigma_1(\vec{k}-\vec{p},\,\mathcal{E}(\vec{k}-\vec{p}))-\Sigma_1(k_F,\,\mu_{eg})]\}\,\frac{1}{[\mathcal{E}_0(\vec{k})-\mathcal{E}_0(\vec{k}-\vec{p})]+[\Sigma_1(\vec{k})-\Sigma_1(\vec{k}-\vec{p})]}\,. \tag{69}$$

The final approximation is to neglect differences in self-energies. For an electron gas at metallic densities this approximation is fairly well satisfied.[24] Thus the final approximate expression is

$$\frac{1}{\bar{N}}\sum_{\vec{p},\vec{k}}|V(\vec{p})|^2\,\frac{1}{\epsilon(\vec{p},\,0,\,\mu)}\,\frac{\theta(\mathcal{E}_F+\delta\mu^b-\mathcal{E}_0(\vec{k}))-\theta(\mathcal{E}_F+\delta\mu^b-\mathcal{E}_0(\vec{k}-\vec{p}))}{\mathcal{E}_0(k)-\mathcal{E}_0(\vec{k}-\vec{p})}\,. \tag{70}$$

For higher-order terms we proceed in the same manner. Denoting the above approximation to $V(\vec{p})\Lambda_{\vec{p}}$ by a double broken line and by a double wavy line the analogous approximation for the electron-electron interaction, we include the class of diagrams given in Figs. 3 and 4. It can be shown that these correspond to a random-phase approximation in the sense described by Cohen and Ehrenreich[25] *provided* one takes $\epsilon(\vec{p},\,0,\,\mu)$ to be the Lindhard dielectric function.

We next examine certain complications which appear in fourth and higher orders and which are illustrated by the fourth-order diagram of Fig. 5. This gives a contribution to the band-structure energy

$$\frac{2}{\beta}\,\frac{1}{4}\sum_{\vec{k},\vec{p},\vec{q}_1,\vec{q}_2,\nu}V(-\vec{k})\Lambda_{-\vec{k}}(\vec{p}+\vec{k},\,\omega_\nu)G^{(0)}(\vec{p},\,\omega_\nu)V(\vec{q}_1)\Lambda_{\vec{q}_1}(\vec{p},\,\omega_\nu)G^{(0)}(\vec{p}+\vec{q}_1\,,\,\omega_\nu)V(\vec{q}_2)$$

$$\times\Lambda_{\vec{q}_2}(\vec{p}+\vec{q}_1\,,\,\omega_\nu)G^{(0)}(\vec{p}+\vec{q}_1+\vec{q}_2\,,\,\omega_\nu)V(\vec{k}-\vec{q}_1-\vec{q}_2)\Lambda_{\vec{k}-\vec{q}_1-\vec{q}_2}(\vec{p}+\vec{q}_1+\vec{q}_2\,,\,\omega_\nu)G^{(0)}(\vec{p}+\vec{k},\,\omega_\nu). \tag{71}$$

In evaluating the ν sum, we perform a contour integration and the possibility of double poles is evident (see Fig. 6). The double pole contribution gives rise from differentiation of the factor $(e^{\beta(\omega-\mu)}+1)^{-1}$, to a δ-function contribution in the $T=0$ limit, i.e.,

$$\Delta E=-\sum_{\vec{k},\vec{p},\vec{q}}\left|\frac{V(\vec{k})}{\epsilon(\vec{k})}\right|^2\left|\frac{V(\vec{q})}{\epsilon(\vec{q})}\right|^2$$

$$\times\,\frac{\delta(\mathcal{E}_F+\delta\mu^b-\mathcal{E}_0(\vec{p}))}{[\mathcal{E}_0(\vec{p})-\mathcal{E}_0(\vec{p}+\vec{k})][\mathcal{E}_0(\vec{p})-\mathcal{E}_0(\vec{p}+\vec{q})]}\,. \tag{72}$$

From Eqs. (A5) and (A8), the origin of this term is clear. It arises from an expansion of $\theta(E_F-E(\vec{k}))$, where $E(\vec{k})$ is the eigenvalue of the single-electron band-structure Hamiltonian. It is important to note that this expansion is invalid when $\vec{k}$ is too near a zone plane: in fact, ΔE of Eq. (72) diverges quadratically there. Although the behavior of these anomalous[26] contributions is general, we can ignore them provided the θ functions occurring in the other expressions are modified from $\theta(\mu^b-\mathcal{E}_0(\vec{k}))$ to $\theta(\mu^b-E(\vec{k}))$, where $\mu^b=\mathcal{E}_F+\delta\mu^b=\mathcal{E}_F$ and is the chemical potential one computes in a band-structure calculation from[27,28]

$$\bar{N}=2\sum_{\vec{k}}\theta(E_F-E(\vec{k}))\,. \tag{73}$$

The contributions from (71) not involving δ functions may be shown to give the first three terms of Eq. (A8). The first term of this expression is well defined; however, the second and third terms, owing to the squared denominator, are di-

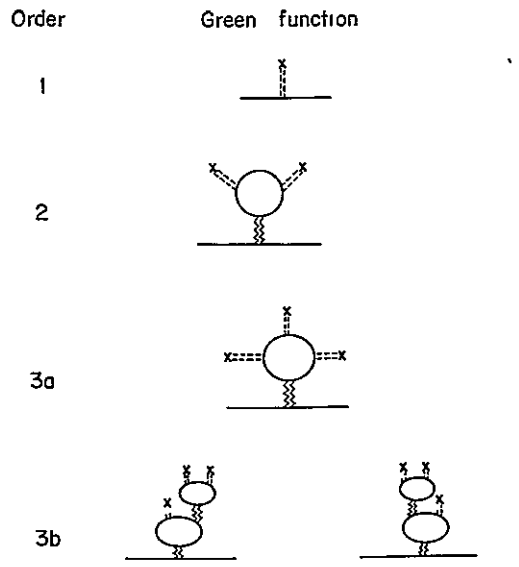
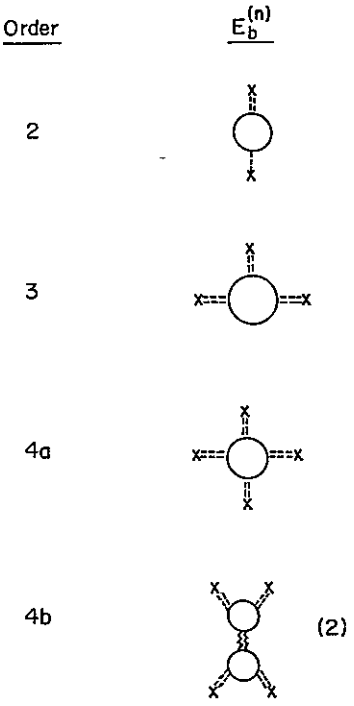


FIG. 3. Corrections to the Green's function The double-dashed line and double-wavy line represent the dielectric approximation described in the text



(2)

FIG. 4. Contributions to the band-structure energy.

vergent when the Fermi sphere is near a zone plane. This divergence is an artifact of the asymptotic nature of the expansion (A1). In Appendix B we show that a resummation of diagrams leads to a finite result.

Finally, we make a remark concerning the electron-gas term $E_0(\mu)$. This can be calculated from approximate expressions for $\Xi(\mu)$ (e.g., the Nozières-Pines formula). However, to gain some physical insight, we expand $\Xi_0(\mu)+\mu\bar{N}$ about $\mu_0 = \mu_{eg}$,

$$\Xi_0(\mu)+\mu\bar{N}=\bar{N}E_0(\mu_0)+\frac{1}{2}\left(\frac{\partial^2\Xi_0}{\partial\mu^2}\right)(\delta\mu^b)^2+\cdots, \quad (74)$$

and noting

$$\langle(\Delta N)^2\rangle = k_B T\left(\frac{\partial N}{\partial\mu}\right)_{T,V}, \quad (75)$$

we see that the right-hand side of Eq. (74) becomes

$$\bar{N}E_0(\mu_0)-(1/2k_BT)\langle(\Delta N\delta\mu^b)^2\rangle+\cdots, \quad (76)$$

so that the change in electron-gas energy *lowers* the total energy and is clearly related to the distortion of the spherical Fermi surface of the electron gas into the lattice symmetric Fermi surface of the periodic system. We may also observe that if Eq. (50) is written

$$\Xi(\mu)=\Xi_0(\mu)+\Xi_b(\mu) \quad (77)$$

and expanded to fourth order in the external potential, the following expressions are obtained for internal energy, chemical potential, and pressure:

$$E=[E_0(\mu_0)+E_b^{(2)}(\mu_0)+E_b^{(3)}(\mu_0)+E_b^{(4)}(\mu_0)]+\tfrac{1}{2}(1/\Omega_0)K_T(\delta\mu_2)^2+O(V^5), \quad (78)$$

$$\mu=\mu_0+\delta\mu_2+\delta\mu_3+\delta\mu_4+O(V^5), \quad (79)$$

where

$$\delta\mu_2=\left(E_b^{(2)}(\mu_0)-\frac{r_s}{3}\frac{dE_b^{(2)}(\mu_0)}{dr_s}\right), \quad (80)$$

$$\delta\mu_3=\left(E_b^{(3)}(\mu_0)-\frac{r_s}{3}\frac{dE_b^{(3)}(\mu_0)}{dr_s}\right), \quad (81)$$

$$\delta\mu_4=\left(E_b^{(4)}(\mu_0)-\frac{r_s}{3}\frac{dE_b^{(4)}(\mu_0)}{dr_s}\right), +\frac{1}{\Omega_0}K_T\delta\mu_2\frac{1}{9}\left(r_s^2\frac{d^2E_b^{(2)}}{dr_s^2}-2r_s\frac{dE_b^{(2)}}{dr_s}\right)$$

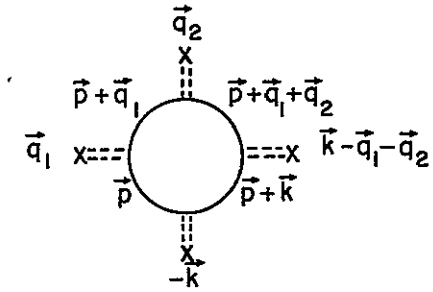


FIG. 5. Fourth-order contribution to the band-structure energy given by term 4a of Fig. 4

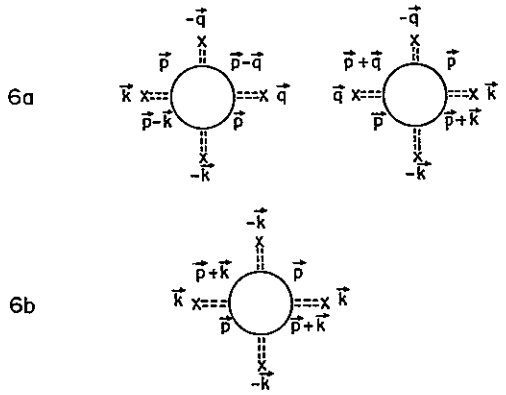


FIG. 6. Divergent fourth-order diagrams.

$$+\frac{1}{\Omega_0}K_T(\delta\mu_2)^2\left(1+\frac{r_s}{6B_T}\frac{dB_T}{dr_s}\right), \qquad (82)$$

and

$$p=[p_0(\mu_0)+p^{(2)}+p^{(3)}+p^{(4)}]$$
$$+\frac{1}{\Omega_0^2}K_T\delta\mu_2\frac{1}{9}\left(r_s^2\frac{d^2E_b^{(2)}}{dr_s^2}-2r_s\frac{dE_b^{(2)}}{dr_s}\right)$$
$$+\frac{1}{2}\frac{1}{\Omega_0^2}K_T(\delta\mu_2)^2\left(1+\frac{1}{3}\frac{r_s}{B_T}\frac{dB_T}{dr_s}\right)+O(V^5), \qquad (83)$$

where

$$p^{(m)}=-\frac{1}{4\pi r_s^2}\frac{dE_b^{(m)}(\mu_0)}{dr_s}. \qquad (84)$$

The quantity $K_T=1/B_T$ is the isothermal compressibility of the interacting electron gas and, as before, μ_0 is the chemical potential of the interacting electron gas, both evaluated at density Ω_0^{-1}. (The bracketed terms are to be expected from zero-temperature perturbation theory.) We note that the two methods agree to third order but in fourth order differ for the physical reason outlined above (i.e., Fermi-surface distortion). These differences although small are not always negligible as will be shown in Sec V.

Recapitulating to this point, we have seen that the theory presented in Sec. III must be modified in several ways. First, the electron-gas term in the total energy must be corrected to take into account the shift in chemical potential due to the ions. Second, the expressions of Sec. II for χ_n, except for the first, must be multiplied by an additional factor of $\epsilon^{-1}(\vec{k}, 0, \mu)$. Third, terms such as 4b of Fig. 4 must be included in a self-consistent calculation. (These are essentially Hubbard's[3] H diagrams which from his point of view are connected with double counting.) We now turn to a discussion of the magnitude of these various corrections for the particular case of a solid composed of massive protons arrayed on a Bravais crystal lattice.

V ATOMIC HYDROGEN

In this section we present the results of calculations for zero-temperature thermodynamic properties of three atomic hydrogen lattices, simple cubic (sc), face-centered cubic (fcc), and body-centered cubic (bcc). This choice was made partly for convenience of computation, but more importantly because of the relatively large difference in Madelung constant between sc and the other two structures. We shall use expressions (78)–(83) and proceed order by order.

A Electron gas

We have taken the Nozières-Pines interpolation formula for the ground-state energy of the interacting electron gas[29].

$$E_0(\mu_0)=\tfrac{3}{5}(\tfrac{9}{4}\pi)^{2/3}r_s^{-2}-(3/2\pi)(\tfrac{9}{4}\pi)^{1/3}r_s^{-1}$$
$$+(-0.115+0.031\ln r_s). \qquad (85)$$

In a comparison of structures, the magnitude of the structure-independent contribution plays no role so that a better approximation is not necessary. In any case, the Nozières-Pines expression compares very well with more recent forms.[30]

B Madelung energy

The Madelung energy may be written in the form

$$E_M=-A_M/r_s, \qquad (86)$$

where the Madelung constant A_M for the three structures is given by[31] sc, 1.760 122; fcc, 1.791 749; and bcc, 1 791861.

C Second-order band-structure energy

We take the Lindhard expression for the dielectric function in the calculation of the terms in the band-structure energy:

$$\epsilon(\eta;\mu_0)=1+(1/2\pi)(4/9\pi)^{1/3}r_s g(\eta),$$
$$g(\eta)=\frac{2}{\eta^2}\left(\frac{1-\eta^2}{4\eta}\ln\left|\frac{\eta+1}{\eta-1}\right|+\frac{1}{2}\right), \qquad (87)$$

with $\eta\equiv k/2k_F$. Then the second-order band-structure energy may be written[1,2]

$$E_b^{(2)}(\mu_0)=-\frac{1}{6\pi^2}$$
$$\times\sum_{\eta\neq 0}\eta^{-2}\frac{g(\eta)}{1+(1/2\pi)(4/9\pi)^{1/3}r_s g(\eta)}. \qquad (88)$$

D Third-order band-structure energy

This contribution is given by Eq. (A7) and corresponds to diagram 3 of Fig. 4. It may be written in the following form:

TABLE I. Parameters in expansion (91) of third-order band-structure energy.

Lattice (real space)	b_0	b_1	b_2	b_3	c_4
sc	0.08202	0.1195	0.1506	0.1748	−0.00310
bcc	0.06483	0.06591	0.05467	0.04050	−0.00275
fcc	0.06663	0.06945	0.05933	0.04555	−0.00260

$$E_b^{(3)}(\mu_0)=-\frac{16}{9\pi}\left(\frac{4}{9\pi}\right)^{1/3} r_s \times \sum_{\vec{\eta},\vec{\eta}_1}{}' \tilde{w}(-\vec{\eta})\tilde{w}(\vec{\eta}-\vec{\eta}_1)\tilde{w}(\vec{\eta}_1) H^{(3)}(\vec{\eta},\vec{\eta}_1) , \tag{89}$$

where $\tilde{w}(\vec{\eta})=1/\eta^2\epsilon(\vec{\eta},\mu_0)$, $\eta_1=k,\ 2k_F$, and

$$H^{(3)}(\vec{\eta},\vec{\eta}_1)=k_F\frac{1}{\Omega}\times\sum_{\vec{q}}\frac{n_0(\vec{q})}{[\mathcal{E}_0(\vec{q})-\mathcal{E}_0(\vec{q}+\vec{k})][\mathcal{E}_0(\vec{q})-\mathcal{E}_0(\vec{q}+\vec{k}_1)]} . \tag{90}$$

The complete expression for $H^{(3)}(\vec{\eta},\vec{\eta}_1)$ is given in Appendix C. The third-order contribution thus depends linearly on r_s apart from a weak dependence contained in the dielectric functions. The function $H^{(3)}(\eta,\eta_1)$ in this approximation is independent of r_s and depends purely on the structure. It is everywhere finite but has discontinuous derivatives for certain values of $\vec{\eta}$, $\vec{\eta}_1$ as discussed by Lloyd and Sholl.[7] We have expanded $E_b^{(3)}(\mu_0)$ as a power series in the parameter $cr_s=-(1/2\pi)(4/9\pi)^{1/3}r_s$, which occurs in the Lindhard function. Thus

$$E_b^{(3)}(\mu_0)=ar_s[b_0+cr_sb_1+(cr_s)^2b_2+(cr_s)^3b_3+\cdots] , \tag{91}$$

where $a=-(16/9\pi)(4/9\pi)^{1/3}$. The values of these structural constants are given in Table I.

E Fourth-order band-structure energy

There are several distinct contributions in this order. First we consider the most divergent parts of the last two terms in Eq. (A8), namely,

$$-\frac{2}{N}\sum_{\vec{k}} n_0(\vec{k})\sum_{l\neq 0}\left|\frac{V(\vec{K}_l)}{\epsilon(\vec{K}_l)}\right|^4\frac{1}{(\mathcal{E}_0-\mathcal{E}_l)^3} \tag{92}$$

and

$$-\frac{1}{N}\sum_{\vec{k}}\sum_{l\neq 0}\left|\frac{V(\vec{K}_l)}{\epsilon(\vec{K}_l)}\right|^4\frac{\delta(E_f-\mathcal{E}_0)}{(\mathcal{E}_0-\mathcal{E}_l)^2} , \tag{93}$$

which we write

$$\tilde{E}_4=\sum_l[E_1^{(4)}(\vec{K}_l)+E_2^{(4)}(\vec{K}_l)] . \tag{94}$$

In Fig. 7 we show $E_1^{(4)}(\eta)/E_2^{(0)}(\eta)$ and $E_2^{(4)}(\eta)/E_2^{(0)}(\eta)$ as functions of η [where $E_b^{(2)}=\sum_\eta' E_2^{(0)}(\eta)$] along with the resummed expression given in Appendix B. Note that $E_2^{(4)}(\eta)$ is part of the anomalous contribution as discussed in Sec. IV and that it must be included at finite order to give the appropriate limiting agreement with the resummed diagrams. Furthermore, we note from the positions of the first reciprocal-lattice vectors that the contribution of this term will be small. The behavior exhibited in this term is representative of the nature of any spurious divergences introduced by zone planes and illustrates the interconnection between band-structure effects and the methods (finite T and $T=0$) of perturbation theory.

Second, we consider contributions from diagram 4b of Fig. 4. This term may be written

$$E_2^{(4b)}=\frac{64}{27\pi}\left(\frac{4}{9\pi}\right)r_s^3\sum_{\eta_1}{}'\frac{1}{\eta_1^2\epsilon(\eta_1)}G^2(\eta_1) , \tag{95}$$

where

$$G(\eta_1)=\sum_{\eta_2}{}'\frac{1}{\eta_2^2\epsilon(\eta_2)}\frac{1}{(\vec{\eta}_1-\vec{\eta}_2)^2\epsilon(\vec{\eta}_1-\vec{\eta}_2)}\times[2H^{(3)}(\vec{\eta}_1,\vec{\eta}_2)+H^{(3)}(\vec{\eta}_2,\vec{\eta}_2-\vec{\eta}_1)] \tag{96}$$

and can be calculated readily since the expressions for $H^{(3)}(\vec{\eta}_1,\vec{\eta}_2)$ are known. Furthermore, apart from the weak r_s dependence of ϵ this term is proportional to r_s^3. Numerical results for two representative values of r_s are given in Table II.

Next we consider the correction which arises as a consequence of the chemical potential shift, namely, the last term in Eq. (78):

$$E_1^{(4b)}=\tfrac{1}{2}(1/\Omega_0)K_T(\delta\mu_2)^2 . \tag{97}$$

This is known from the expressions for the compressibility of the electron gas and the second-order value of the chemical potential. In fact, as a consequence of the compressibility sum rule, it may be shown that this term is precisely given by the diagram for $E_2^{(4b)}$ in the limit that the momentum transferred by the internal Coulomb line approaches zero.

Finally, we consider contributions due to diagrams of the form labeled 4a in Fig. 4. There are two contributions apart from those already discussed in the first part of this section and are given in Eq. (A8). One is an off-diagonal part

$$2\Omega_0\sum_{\substack{l\neq 0,I\\ j\neq 0,I\\ I\neq 0}}\frac{V(\vec{K}_l)}{\epsilon(\vec{K}_l)}\frac{V(-\vec{K}_l)}{\epsilon(\vec{K}_l)}\frac{V(\vec{K}_j-\vec{K}_l)}{\epsilon(\vec{K}_j-\vec{K}_l)}\frac{V(\vec{K}_l-\vec{K}_l)}{\epsilon(\vec{K}_j-\vec{K}_l)}$$

TABLE II Contributions to fourth order in electron-ion interaction to fre energy. $E_2^{(4b)}$—Fig. 4(b), E^{4a}—Fig. 4(a), $\tilde{E}_4$—Fig. 8(c). $E_1^{(4b)}$—chemical potential correction—see text of Sec. IV.

	$r_s=1.6$			$r_s=1.36$		
	sc	fcc	bcc	sc	fcc	bcc
E_0	0.190106	. .	..	0 415590	...	. .
E_M	−1.100076	−1 119843	−1.119913	−1.294207	−1.317462	−1.317545
E_2	−0.105351	−0 086230	−0.085549	−0.106694	−0.086949	−0.086237
E_3	−0.03227	−0.02753	−0 02687	−0.02815	−0.02385	−0.02327
E_1^{4b}	0.00844	0 00555	0.005459	0.00587	0.003832	0.003765
E_2^{4b}	0.00108	0.00076	0.000762	0 000696	0 000482	0.000485
$\tilde{E}_4$	−0.00187	−0.000454	−0.000385	−0.00170	−0.000339	−0.000287
E^{4a}	−0.0077	−0.0067	−0.0044	−0.0055	−0.0048	−0 0037

$$\times\frac{1}{\Omega}\sum_{\vec{k}}\frac{n_0(\vec{k})}{(\mathcal{E}_0-\mathcal{E}_s)(\mathcal{E}_0-\mathcal{E}_j)(\mathcal{E}_0-\mathcal{E}_i)} \tag{98}$$

and the other has diagonal parts

$$2\Omega_0\sum_{\substack{i\neq 0,\,l\\ l\neq 0}}\left|\frac{V(\vec{K}_l)}{\epsilon(\vec{K}_l)}\right|^2\left|\frac{V(\vec{K}_i)}{\epsilon(\vec{K}_i)}\right|^2$$
$$\times\frac{1}{\Omega}\sum_{\vec{k}}n_0(\vec{k})\left(\frac{1}{(\mathcal{E}_0-\mathcal{E}_l)^2(\mathcal{E}_0-\mathcal{E}_{l-l})}\right.$$
$$\left.-\frac{1}{(\mathcal{E}_0-\mathcal{E}_l)^2(\mathcal{E}_0-\mathcal{E}_s)}\right) \tag{99a}$$

and

$$-\frac{1}{N}\sum_{\substack{i\neq 0,\,l\\ l\neq 0}}\left|\frac{V(\vec{K}_l)}{\epsilon(\vec{K}_l)}\right|^2\left|\frac{V(\vec{K}_i)}{\epsilon(\vec{K}_i)}\right|^2$$
$$\times\sum_{\vec{k}}\frac{\delta(E_F-\mathcal{E}_0)}{(\mathcal{E}_0-\mathcal{E}_l)(\mathcal{E}_0-\mathcal{E}_i)}\,. \tag{99b}$$

Equation (99b) is an anomalous contribution, which disappears along with the singularities from the double poles if the resummation of Appendix B is used. These terms are awkward to handle in nu-

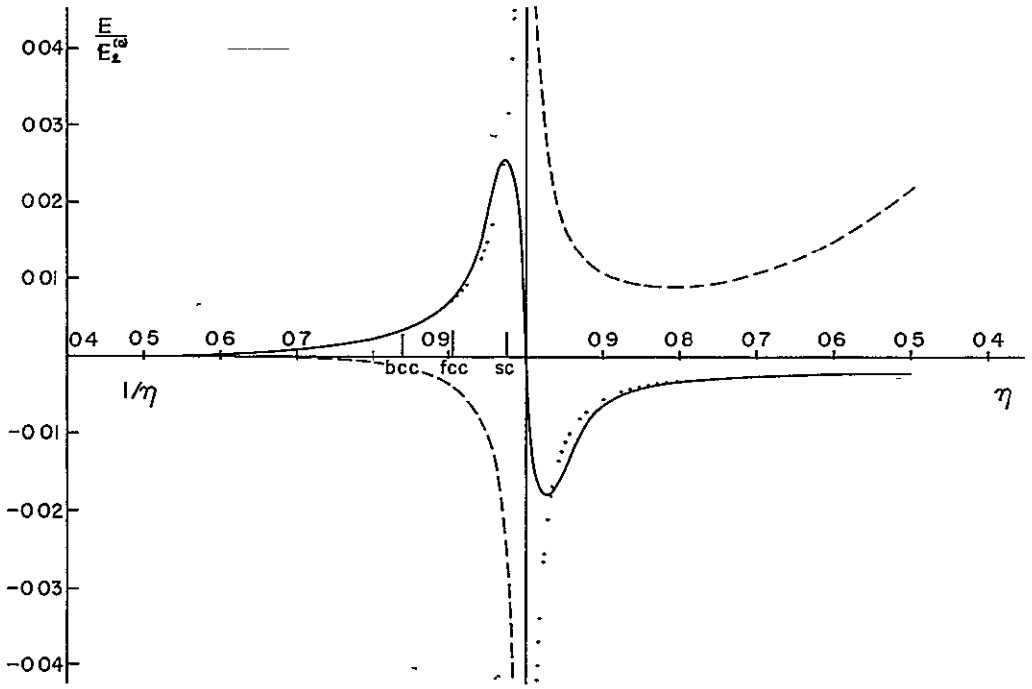


FIG. 7. Solid line, $[E_4(\eta)-E_4^{(0)}(\eta)]/E_2^{(0)}(\eta)$ (cf. Appendix B), dashed line, $E_1^{(4)}(\eta)/E_2^{(0)}(\eta)$; and dotted line, $[E_1^{(4)}(\eta)+E_2^{(4)}(\eta)]/E_2^{(0)}(\eta)$ [cf. Eqs. (92) and (93) and Appendix B]. Note the left-hand axis is $1/\eta$; right-hand axis is η. Vertical bars represent shortest reciprocal lattice vectors for the structures indicated.

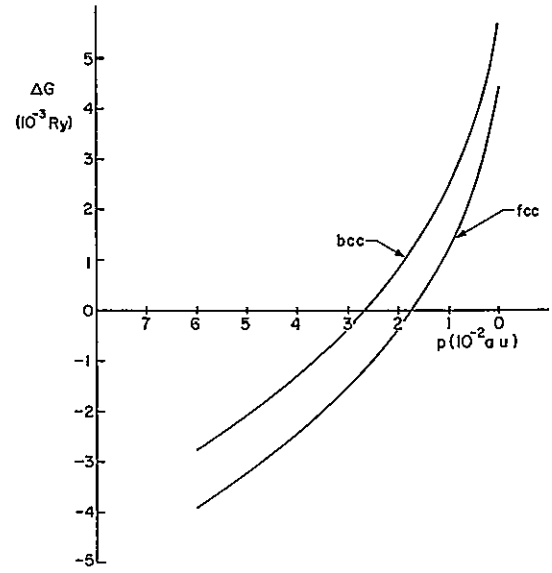


FIG 8 Gibbs free-energy difference relative to the simple cubic lattice for fcc and bcc metallic hydrogen.

merical work, although in principle there is no difficulty. [One problem is the time needed to calculate a nine-dimensional sum. Another is that the kernel

$$K(\vec{\eta}_1,\vec{\eta}_2,\vec{\eta}_3)=\frac{1}{\Omega}\sum_{\vec{k}}{}'\frac{n_0(\vec{k})}{(\mathcal{E}_0-\mathcal{E}_1)(\mathcal{E}_0-\mathcal{E}_2)(\mathcal{E}_0-\mathcal{E}_3)} \tag{100}$$

has, as yet, no analytic representation. We have been able to reduce it to a two-dimensional integral. It has an asymptotic expansion which gives

$$K(\vec{\eta}_1,\vec{\eta}_2,\vec{\eta}_3)\sim-\frac{1}{48\pi^2}\frac{1}{(2k_F)^3}\frac{1}{\eta_1^2}\frac{1}{\eta_2^2}\frac{1}{\eta_3^2} \times(\eta_1,\eta_2,\eta_3\gtrsim 1.5)\ . \tag{101}$$

We calculated these terms [Eqs. (98) and (99)] by taking as an approximation for $K(\vec{\eta}_1,\vec{\eta}_2,\vec{\eta}_3)$, its large η expansion, and by setting $1/\epsilon(\eta)=1$. The former is an underestimate but note that for the structures we consider η is always >1. The latter is an overestimate. The form is then

$$E^{(4a)}\cong-\frac{1}{4(3\pi)^2}\left(\frac{4}{9\pi}\right)^{2/3}r_s^2 \times\sum_{\eta_1}\eta_1^2\left(\sum_{\eta_2}{}'\frac{1}{\eta_2^4}\frac{1}{(\vec{\eta}_1-\vec{\eta}_2)^2}\right)^2\equiv r_s^2 c_4\ , \tag{102}$$

which is proportional to r_s^2 and is probably an underestimate overall. The values for the factor c_4 are given in Table I.]

In Tables III–V we give the thermodynamic functions p, E, G, at $T=0$ °K calculated to third order in the electron-ion interaction. In Table II we list the explicit contributions to fourth order at $r_s=1.6$ and $r_s=1.36$ corresponding to low pressure and 1.9 Mbar, respectively.[32] The contribution $E^{(4a)}$ is an estimate as noted above. Note the approximate cancellation in the fourth order, and further that at high pressures the sc lattice is predicted to be unstable relative to fcc and bcc (see Fig. 8).

VI DISCUSSION AND CONCLUSIONS

We have given a procedure for calculating the ground-state energy of a simple metal and have shown that there are basically four contributions involved, viz., electron gas, static dielectric energy, Madelung, and core exclusion. Furthermore, we have seen that the shift in chemical potential from that of a uniform electron gas must be taken into account in calculations going beyond second order. In particular, we have emphasized that $T=0$ time-dependent perturbation theory does not give the true ground state when the unperturbed system is taken to have a spherical Fermi surface (a fact first noted by Kohn and Luttinger[17]) and have shown the relationship of this to the deformation of the unperturbed Fermi surface. We have observed that if one expands the free energy uniformly in powers of electron-ion interaction, differences between finite- and zero-temperature perturbation theory appear only in fourth and higher orders, and furthermore, that certain divergences at zone planes can be resolved by resummations.

The preliminary calculations reported here for atomic hydrogen seem to indicate that a happy cancellation may occur in the fourth order, at least for the sc, fcc, and bcc structures, although more detailed calculations are required to be certain of

TABLE III. $T=0$ °K equation of state for atomic hydrogen (to third order). Note that these results are appropriate to a static lattice and do not, therefore, include phonon contributions to the equation of state. Note also that one atomic unit of pressure = 147 15 Mbar

	Pressure		
r_s	sc	fcc	bcc
1.65	-2.03×10^{-4}	-5.16×10^{-4}	-5.23×10^{-4}
1.60	7.89	4.31	4.24
1.55	2.13×10^{-3}	1.72×10^{-3}	1.71×10^{-3}
1.50	3.92	3 45	3.44
1.45	6.32	5.78	5.77
1 40	9.54	8.91	8.90
1 35	1.38×10^{-2}	1.31×10^{-2}	1.31×10^{-2}
1.30	1.96	1.88	1 87
1.25	2 74	2 64	2 64
1.20	3.79	3 67	3 67
1.15	5.22	5.08	5.08
1 10	7.19	7.02	7.02
1.05	9.92	9.71	9.71
1 00	1.37×10^{-1}	1.35×10^{-1}	$1\ 35\times10^{-1}$

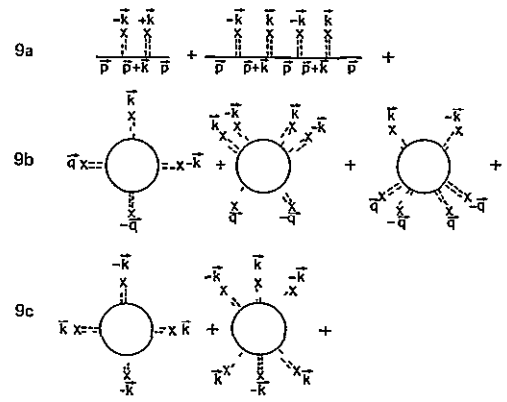


FIG. 9 (a) Partial summation of Green's function. (b) Partial summation for the diagrams of 6a. (c) Partial summation for the diagrams of 6b.

this. The calculations reported have been done using the Lindhard dielectric function. In third and higher orders this is a very good approximation since the dielectric function occurs as $1/\epsilon$. However, in the second order, $\epsilon - 1$ appears. A better choice of ϵ acts to change the magnitude of the second-order contribution slightly but does not affect the energy differences between sc and the two other cubic structures. The use of the Lindhard function, as noted in Sec. III, corresponds to a self-consistent Hartree (RPA) approximation. We remark that the zero pressure density of the structures studied will be extremely sensitive to the exact fourth order corrections due to the weakness of the minimum in the free energy as seen in Table IV. Also, a third-order calculation predicts an instability of the sc structure relative to the two close packed lattices at a pressure of ~2–3 Mbar (see Fig. 8). The exact transition pressure is again sensitive to the magnitude of the fourth-order corrections. It is clear, however, that such a transition must appear at some pressure, for the band-structure corrections depend upon positive powers of r_s, whereas the Madelung term depends inversely upon r_s. Thus eventually, the static lattice having the lowest Madelung energy should be most stable.

Brovman *et al.*[10] have computed ground-state energies for atomic hydrogen at zero pressure by using the $T=0$ expansion to third order in the electron-ion interaction and found an interesting class of low-energy anisotropic structures. We regard the effect of fourth-order corrections to these calculations as an open question, but one that can be settled using the above expressions. It is also important to point out that whereas including higher-order band-structure effects in χ_1 has negligible effect (see Fig. 7) this may not be so in higher orders for certain directions in reciprocal space corresponding to Fermi sphere tangency to zone

TABLE IV. Free energy at $T=0$ °K for atomic hydrogen vs r_s (to third order).

	Free energy		
r_s	sc	fcc	bcc
1.65	−1.048 03	−1.043 38	−1.042 09
1 64	−1.048 07	−1.043 53	−1.042 24
1.63	−1.048 05	−1.043 61	−1 042 33
1.62	−1 047 96	−1 043 63	−1.042 36
1.61	−1.047 81	−1 043 60	−1.042 33
1.60	−1 047 59	−1.043 45	−1.042 22
1.55	−1.045 38	−1.041 88	−1.040 62
1.50	−1 041 04	−1 038 18	−1.036 93
1.45	−1 034 14	−1.031 97	−1 030 73
1.40	−1.024 14	−1.022 72	−1.021 49
1.35	−1.010 42	−1.009 79	−1.008 58
1.30	−0.992 17	−0.992 42	−0.991 22
1.25	−0.968 42	−0.969 61	−0 968 43
1.20	−0.937 96	−0.940 19	−0.939 02
1.15	−0.899 28	−0.902 62	−0.901 47
1.10	−0.850 45	−0.855 02	−0.853 88
1.05	−0.789 03	−0.794 95	−0.793 83
1.00	−0.711 88	−0.719 29	−0.718 19

planes (see Appendix C). Finally, we again emphasize that we have treated the lattice as static and that it will be necessary to consider lattice zero point energy in a complete determination of structural stability since the zero point energy is of the magnitude $E_b^{(3)}$. Calculations of such phonon effects are in progress.[33]

ACKNOWLEDGMENTS

We would like to thank Dr B. Nickel and Dr. A. B. Bringer for numerous helpful discussions.

APPENDIX A

To derive an expansion of Eq. (41) we write the band energy as

TABLE V. Gibbs free energy at $T=0$ °K vs pressure for atomic hydrogen (to third order).

	Gibbs free energy		
Pressure	sc	fcc	bcc
0.0	−1.0481	−1.0436	−1.0424
5.0×10^{-4}	−1.0390	−1 0349	−1.0336
1.0×10^{-3}	−1.0305	−1.0266	−1.0253
5.0	−0.9707	−0.9683	−0.9670
1.0×10^{-2}	−0.9092	−0.9080	−0.9068
2 0	−0.8081	−0.8085	−0 8073
3 0	−0 7233	−0.7248	−0 7237
5 0	−0 5809	−0.5841	−0 5829
1.0×10^{-1}	−0.3019	−0.3085	−0.3075
5 0	−0.9707	−0.9683	−0.9670
1.0	1.8572	1.8377	1.8387
5.0	5.6614	5.6273	5.6282

$$E(\vec{k}) = \mathcal{E}_0(\vec{k}) + \sum_i \frac{\tilde{V}_{0i}\tilde{V}_{i0}}{\mathcal{E}_0 - \mathcal{E}_i} + \cdots \tag{A1}$$

and the occupation number as

$$n(\vec{k}) = n_0(\vec{k}) - \delta(E_F - \mathcal{E}_0(\vec{k})) \sum_i \frac{\tilde{V}_{0i}\tilde{V}_{i0}}{\mathcal{E}_0 - \mathcal{E}_i} + \cdots , \tag{A2}$$

where $n_0(\vec{k}) = \theta(E_F - \mathcal{E}_0(\vec{k}))$ and $n(\vec{k}) = \theta(E_F - E(\vec{k}))$. Clearly, when $\vec{k}$ is near a zone plane, these must be viewed as asymptotic. We find the following expressions for the Fourier components of the density:

$$\rho_l^{(1)} = \frac{2}{\Omega}\sum_{\vec{k}} n_0(\vec{k}) \frac{\tilde{V}_{-l0}}{\mathcal{E}_0 - \mathcal{E}_l} , \tag{A3}$$

$$\rho_l^{(2)} = \frac{2}{\Omega}\sum_{\vec{k}} n_0(\vec{k}) \left(2 \sum_{i\neq 0,l} \frac{\tilde{V}_{0i}\tilde{V}_{il}}{(\mathcal{E}_0 - \mathcal{E}_i)(\mathcal{E}_0 - \mathcal{E}_l)} + \sum_{i\neq 0,l} \frac{\tilde{V}_{il}\tilde{V}_{0l}}{(\mathcal{E}_0 - \mathcal{E}_{i-l})(\mathcal{E}_0 - \mathcal{E}_i)} \right) , \tag{A4}$$

$$\begin{aligned} \rho_l^{(3)} = \frac{2}{\Omega}\sum_{\vec{k}} n_0(\vec{k}) \Bigg(& 2 \sum_{i,j\neq 0,l} \frac{\tilde{V}_{-li}\tilde{V}_{ij}\tilde{V}_{j0}}{(\mathcal{E}_0 - \mathcal{E}_i)(\mathcal{E}_0 - \mathcal{E}_j)(\mathcal{E}_0 - \mathcal{E}_l)} + 2 \sum_{\substack{i\neq 0,l \\ j\neq 0,i}} \frac{\tilde{V}_{il}\tilde{V}_{0j}\tilde{V}_{jl}}{(\mathcal{E}_0 - \mathcal{E}_{i-l})(\mathcal{E}_0 - \mathcal{E}_i)(\mathcal{E}_0 - \mathcal{E}_j)} \\ & + 2\frac{\tilde{V}_{0l}}{(\mathcal{E}_0 - \mathcal{E}_l)^2} \sum_{i\neq l,0} \frac{\tilde{V}_{li}\tilde{V}_{il}}{\mathcal{E}_0 - \mathcal{E}_i} - 2\frac{\tilde{V}_{0l}}{(\mathcal{E}_0 - \mathcal{E}_l)^2} \sum_{i\neq 0} \frac{\tilde{V}_{0i}\tilde{V}_{i0}}{\mathcal{E}_0 - \mathcal{E}_i} - 2\frac{\tilde{V}_{0l}}{\mathcal{E}_0 - \mathcal{E}_l} \sum_{i\neq 0} \frac{\tilde{V}_{0i}\tilde{V}_{i0}}{(\mathcal{E}_0 - \mathcal{E}_i)^2} \Bigg) \\ & - \frac{2}{\Omega}\sum_{\vec{k}} \delta(E_F - \mathcal{E}_0(\vec{k})) \frac{2\tilde{V}_{0l}}{\mathcal{E}_0 - \mathcal{E}_l} \sum_{i\neq 0} \frac{\tilde{V}_{0i}\tilde{V}_{i0}}{\mathcal{E}_0 - \mathcal{E}_i} . \end{aligned} \tag{A5}$$

Using Eq. (29) and supplying the extra factor of $\epsilon^{-1}(\vec{k})$ in the third and higher orders, we find, for the energy,

$$E_b^{(2)} = \frac{1}{N}\sum_{\vec{k},l} |V(\vec{K}_l)|^2 \frac{1}{\epsilon(\vec{K}_l)} \frac{2n_0(\vec{k})}{\mathcal{E}_0 - \mathcal{E}_l} , \tag{A6}$$

$$E_b^{(3)} = \frac{2}{N}\sum_{\vec{k}} \sum_{\substack{i\neq 0,l \\ l\neq 0}} n_0(\vec{k}) \frac{V(-\vec{K}_l)}{\epsilon(-\vec{K}_l)} \frac{V(+\vec{K}_i - \vec{K}_l)}{\epsilon(+\vec{K}_i - \vec{K}_l)} \frac{V(\vec{K}_l)}{\epsilon(\vec{K}_i)} \frac{1}{(\mathcal{E}_0 - \mathcal{E}_i)(\mathcal{E}_0 - \mathcal{E}_l)} , \tag{A7}$$

$$\begin{aligned} E_b^{(4a)} = \frac{2}{N}\sum_{\vec{k}} n_0(\vec{k}) \Bigg(& \sum_{\substack{i\neq 0,l \\ j\neq 0,l \\ l\neq 0}} \frac{V(\vec{K}_l)}{\epsilon(\vec{K}_l)} \frac{V(-\vec{K}_i)}{\epsilon(-\vec{K}_i)} \frac{V(\vec{K}_i - \vec{K}_j)}{\epsilon(\vec{K}_i - \vec{K}_j)} \frac{V(\vec{K}_j - \vec{K}_l)}{\epsilon(\vec{K}_j - \vec{K}_l)} \frac{1}{(\mathcal{E}_0 - \mathcal{E}_l)(\mathcal{E}_0 - \mathcal{E}_i)(\mathcal{E}_0 - \mathcal{E}_j)} + \sum_{\substack{i\neq 0 \\ i\neq 0,l}} \left|\frac{V(\vec{K}_l)}{\epsilon(\vec{K}_l)}\right|^2 \left|\frac{V(\vec{K}_i - \vec{K}_l)}{\epsilon(\vec{K}_i - \vec{K}_l)}\right|^2 \\ & \times \frac{1}{(\mathcal{E}_0 - \mathcal{E}_l)^2(\mathcal{E}_0 - \mathcal{E}_i)} - \sum_{\substack{l\neq 0 \\ i\neq 0}} \left|\frac{V(\vec{K}_l)}{\epsilon(\vec{K}_l)}\right|^2 \left|\frac{V(\vec{K}_i)}{\epsilon(\vec{K}_i)}\right|^2 \frac{1}{(\mathcal{E}_0 - \mathcal{E}_l)^2(\mathcal{E}_0 - \mathcal{E}_i)} \Bigg) - \frac{1}{N}\sum_{\vec{k}} \sum_{\substack{i\neq 0 \\ l\neq 0}} \left|\frac{V(\vec{K}_l)}{\epsilon(\vec{K}_l)}\right|^2 \left|\frac{V(\vec{K}_i)}{\epsilon(\vec{K}_i)}\right|^2 \frac{\delta(E_F - \mathcal{E}_0)}{(\mathcal{E}_0 - \mathcal{E}_i)(\mathcal{E}_0 - \mathcal{E}_i)} . \end{aligned} \tag{A8}$$

APPENDIX B

The diagrams which correspond to the second and third terms of (A8) are shown in Fig. 6. The two diagrams of 6(a) are equal in magnitude when summed over $\vec{k}, \vec{q}$ so we need only calculate one and multiply the result by a factor of 2. We now observe that the series of Fig. 9(a) may be summed, i.e.,

$$G_0(\vec{p}, \omega_\nu) \sum_{m=1}^{\infty} [|\lambda \tilde{V}(\vec{k})|^2 G_0(\vec{p}+\vec{k}, \omega_\nu) G_0(\vec{p}, \omega_\nu)]^m = G_0(p, \omega_\nu) \left(\frac{1}{1 - \lambda^2 |\tilde{V}(\vec{k})|^2 G_0(\vec{p}+\vec{k}, \omega_\nu) G_0(p, \omega_\nu)} - 1 \right) \tag{B1}$$

Hence the series of Fig. 9(b) can also be summed, and supplying the factor of 2, the resummation gives, for Fig. 6(a),

$$\frac{4}{N}\sum_{\substack{\vec{k}+\vec{q} \\ \vec{p},\nu}} |\tilde{V}(\vec{k})|^2 |\tilde{V}(\vec{q})|^2 \int_0^1 \lambda^3 d\lambda \frac{1}{G_0^{-1}(\vec{p}, \omega_\nu) G_0^{-1}(\vec{p}+\vec{k}, \omega_\nu) - \lambda^2 |\tilde{V}(\vec{k})|^2} \frac{1}{G_0^{-1}(\vec{p}, \omega_\nu) G_0^{-1}(\vec{p}+\vec{q}, \omega_\nu) - \lambda^2 |\tilde{V}(\vec{q})|^2} , \tag{B2}$$

which no longer has double poles and hence is always finite. Similarly, the contribution of Fig. 6(b) is the first term in the series of Fig. 9(c), which may be summed to give

$$\frac{1}{N}\sum_{\vec{p},\vec{k},\nu} |\tilde{V}(\vec{k})|^2 \int_0^1 d(\lambda^2) \left(\frac{1}{G_0^{-1}(\vec{p}, \omega_\nu) G_0^{-1}(\vec{p}+\vec{k}, \omega_\nu) - \lambda^2 |\tilde{V}(\vec{k})|^2} - G_0(\vec{p}, \omega_\nu) G_0(\vec{p}+\vec{k}, \omega_\nu) \right) . \tag{B3}$$

This again has only simple poles, and moreover is seen to be a correction to $E_b^{(2)}$ rather than $E_b^{(4)}$. In fact, the integrals appearing in Eq. (B3) can be done analytically.

APPENDIX C

The principal-value integral for Eq. (90) is given by

$$H^{(3)}(\vec{\eta}_1,\vec{\eta}_2)=\frac{1}{64\pi^2}\frac{1}{\eta_1\eta_2\sin^2\theta}\mathrm{Re}\Big[(\eta_1-\eta_2\cos\theta)\ln\Big(\frac{\eta_2+1}{\eta_2-1}\Big)+(\eta_2-\eta_1\cos\theta)\ln\Big(\frac{\eta_1+1}{\eta_1-1}\Big)$$
$$+\theta\ln[(\eta_1^2-1)(\eta_2^2-1)]-\theta\ln[(\eta_1^2\eta_2^2+\eta_1^2+\eta_2^2-4\eta_1\eta_2\cos\theta+\cos2\theta)+2\theta(\eta_1\eta_2-\cos\theta)]\Big]\,, \tag{C1}$$

where

$$\theta=(\eta_1^2+\eta_2^2-2\eta_1\eta_2\cos\theta-\sin^2\theta)^{1/2}\,.$$

When $\theta=-i\theta'$, i.e., when $\vec{\eta}_1$, $\vec{\eta}_2$, $\vec{\eta}_1-\vec{\eta}_2$ form a triangle which can be inscribed in a circle of diameter <1, this function becomes

$$H^{(3)}(\vec{\eta}_1,\vec{\eta}_2)=\frac{1}{64\pi^2}\frac{1}{\eta_1\eta_2\sin^2\theta}\Big((\eta_1-\eta_2\cos\theta)\ln\Big|\frac{\eta_2+1}{\eta_2-1}\Big|+(\eta_2-\eta_1\cos\theta)\ln\Big|\frac{\eta_1+1}{\eta_1-1}\Big|$$
$$+\theta'\arg[(\eta_1^2\eta_2^2+\eta_1^2+\eta_2^2-4\eta_1\eta_2\cos\theta+\cos2\theta)-2i\theta'(\eta_1\eta_2-\cos\theta)]\Big)\,, \tag{C2}$$

with

$$\theta'=(\sin^2\theta-\eta_1^2-\eta_2^2+2\eta_1\eta_2\cos\theta)^{1/2}\,,$$

and the argument function is the principal branch with the branch cut along the positive real axis. When the Fermi sphere is contained within the first Brillouin zone (the cases we have considered), it is sufficient to use the principal-value integral. However, when this is not the case, one must use the symmetric form which occurs in Eq. (29).

$$H^{(3)}(\vec{\eta}_1,\vec{\eta}_2)\to\tfrac{1}{3}[\tilde{H}^{(3)}(\vec{\eta}_1,\vec{\eta}_2)$$
$$+H^{(3)}(\vec{\eta}_1-\vec{\eta}_2,-\vec{\eta}_2)+H^{(3)}(\vec{\eta}_2-\vec{\eta}_1,-\vec{\eta}_1)]$$
$$\equiv(1/48\pi^2)\overline{\Lambda}_0^{(3)}(\vec{\eta}_1,\vec{\eta}_2)\,, \tag{C3}$$

where, when $H^{(3)}$ is as given in (C2), the tilde over the first term means that 2π must be subtracted from the argument function.[34] (This ensures that the proper small-η limit obtains.) Moreover, in the region of $\vec{\eta}_1,\vec{\eta}_2$ space for which $\theta\approx0$, it is necessary to include detailed band structure in energy denominators to avoid anomalously large val-

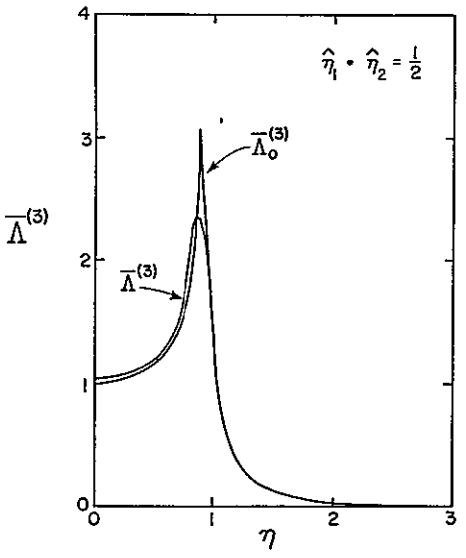


FIG. 10. Normalized susceptibilities $\overline{\Lambda}_0^{(3)}(\vec{\eta}_1,\vec{\eta}_2)$ and $\overline{\Lambda}^{(3)}(\vec{\eta}_1,\vec{\eta}_2)$ vs η for $|\vec{\eta}_1|=|\vec{\eta}_2|=\eta$ and $\hat{\eta}_1\,\hat{\eta}_2=\frac{1}{2}$. $\overline{\Lambda}^{(3)}$ includes band structure, $\overline{\Lambda}_0^{(3)}$ does not.

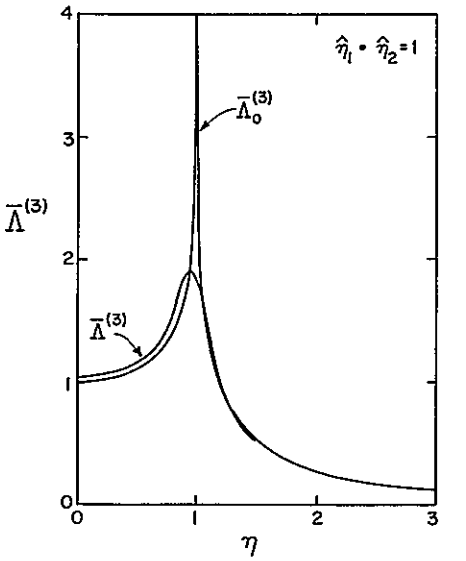


FIG. 11. Normalized susceptibilities $\overline{\Lambda}_0^{(3)}(\vec{\eta}_1,\vec{\eta}_2)$ and $\overline{\Lambda}^{(3)}(\vec{\eta}_1,\vec{\eta}_2)$ vs η for $|\vec{\eta}_1|=|\vec{\eta}_2|=\eta$ and $\hat{\eta}_1\cdot\hat{\eta}_2=1$. $\overline{\Lambda}^{(3)}$ includes band structure, $\overline{\Lambda}_0^{(3)}$ does not.

ues. For example, in Figs. 10 and 11 we show $\bar{\Lambda}_0^{(3)}(\vec{\eta}_1, \vec{\eta}_2)$ for $|\vec{\eta}_1| = |\vec{\eta}_2|$ and two values of $\hat{\eta}_1 \cdot \hat{\eta}_2$ as a function of η compared with $\bar{\Lambda}^{(3)}(\vec{\eta}_1, \vec{\eta}_2)$, the same function modified by band structure in the manner of Appendix B. In the region about $\hat{\eta}_1 \cdot \hat{\eta}_2 = 1$, the reduction can be substantial.

---

*Work supported in part by NASA, contract No NGR-33-010-188, and by the National Science Foundation, contract No GH-33637, through the facilities of the Materials Science Center at Cornell University, Report No 1943

[1]N W Ashcroft and D C Langreth, Phys Rev **155**, 682 (1967)
[2]V Heine and D Weaire, Solid State Phys **24**, 249 (1970)
[3]J Hubbard, Proc R Soc A **243**, 336 (1958), Proc R Soc A **244**, 199 (1958)
[4]M H Cohen, Phys Rev **130**, 1301 (1963)
[5]P Nozières and D Pines, Nuovo Cimento **9**, 470 (1958)
[6]E G Brovman and Yu Kagan, Zh Eksp Teor Fiz 57, 1329 (1969) [Sov Phys -JETP **30**, 721 (1970)], and references therein
[7]P Lloyd and C A Sholl, J Phys C **1**, 1620 (1968)
[8]P Hohenberg and W Kohn, Phys Rev **136**, B864 (1964)
[9]W A Harrison, Phys Rev B **7**, 2408 (1973)
[10]E G Brovman, Yu Kagan, and A Kholas, Zh Eksp Teor Fiz **61**, 2429 (1971) [Sov. Phys -JETP **34**, 1300 (1972)]
[11]N W Ashcroft and J Hammerberg (unpublished)
[12]The arguments below are to be interpreted in the sense of an implied thermodynamic limit, that is, $N, \Omega \to \infty$, $N/\Omega = \text{const}$ Thus, for example, we may replace $N(N-1)/\Omega^2$ by $(N/\Omega)^2$
[13]Many pseudopotentials may be so characterized, see, e g , N W Ashcroft, Phys Lett 23, 48 (1966), J Phys C 1, 232 (1968)
[14]By "core" we mean to allude to the deviation of the pseudopotential from a pure-Coulombic form in the core region due to the pseudopotential transformation and not to imply any other effect of core levels on the energy
[15]See, for example, J D Jackson, *Classical Electrodynamics* (Wiley, New York, 1962), p 123 We note that the induction is understood to be at fixed total charge so that (25) is the appropriate expression
[16]This terminology is frequently reserved for the first member of the sum
[17]W Kohn and J M Luttinger, Phys Rev **118**, 41 (1960)
[18]J M Luttinger and J C Ward, Phys Rev **118**, 1417 (1960)
[19]P C Martin and J Schwinger, Phys Rev **115**, 1342 (1959)
[20]This follows directly from differentiation of $\Xi = -(1/\beta)\ln \mathrm{Tr}\, e^{-\beta[H(\lambda)-\mu N]}$ See also L D Landau and E M Lifshitz, *Statistical Physics* (Addison-Wesley, Reading, Mass , 1969), p 46
[21]The linear term vanishes since $G^{(0)}(\vec{k}, \vec{k}-\vec{p}, \omega_\nu) = G^{(0)}(\vec{k}, \vec{k}-\vec{p}, \omega_\nu)\, \delta(\vec{k}-\vec{k}+\vec{p})$ and $V(0) \equiv 0$
[22]See D C Langreth, Phys Rev **181**, 753 (1969)
[23]For a discussion of this approximation see H Yasuhara and M Watabe, Prog Theor Phys **49**, 1785 (1973)
[24]See, e g , L Hedin and B I Lundqvist, J Phys C **4**, 2064 (1971)
[25]H Ehrenreich and M H Cohen, Phys Rev **115**, 786 (1959)
[26]See Ref 17
[27]This follows by assuming the exact quasiparticle energies to be $\mathcal{E}(\vec{k}) = E(\vec{k}) + \Sigma_1(\vec{k}, \mathcal{E}(\vec{k}))$ and using the Luttinger (Ref 28) formula, $\bar{N} = 2\sum_{\vec{k}} \theta(\mu - \mathcal{E}(\vec{k}))$ together with the above approximations for the self-energy
[28]J M Luttinger, Phys Rev **119**, 1153 (1960)
[29]We use atomic units $\hbar = e^2/2 = 2m = 1$
[30]P Vashishta and K S Singwi, Phys Rev B **6**, 875 (1972)
[31]C A Sholl, Proc Phys Soc Lond **92**, 434 (1967)
[32]For comparison we have calculated the energy to second order for the liquid metal from $E = E_{eg} + (4k_F^0/\pi)\int_0^\infty dy\,[S_L(y)/\epsilon_H(y) - 1]$, where S_L is the hard-sphere Percus-Yevick structure factor and ϵ_H is the Hubbard dielectric function modified to satisfy the compressibility sum rule For $r_s = 1\ 6$ and $\bar{\eta}$ = packing fraction = 0 45, we find $E = -1\ 016$, these results being very insensitive to $\bar{\eta}$ ($\sim 2\%$ for $\bar{\eta} \sim 0\ 7 \to 0\ 3$) Most of this change can be traced to the Madelung energy
[33]A correct evaluation of the phonon spectrum is essential in determining the zero-pressure density for a given structure, as well as in assessing dynamic stability Futhermore, only when the presence of phonons is taken into proper account can the virial theorem ($E = -K + 3p\Omega$, where K is the kinetic energy of electrons *and* ions) be satisfied
[34]For details, see J Hammerberg, thesis (Cornell University, 1973) (unpublished)