

assume that there is no radiation present so that we shall not be concerned with radiation interaction terms. Neither of these neglected effects introduces appreciable errors, as long as we are interested only in the eigenvalues of the electronic system.

In addition to the foregoing assumptions, it is expedient to assume that all extranuclear electrons may be divided into two classes: (1) the inner electrons, which belong to closed shells that are rigidly attached to the nuclei and are not affected appreciably by changes in interatomic distances; (2) the outer, or valence, electrons, which are affected by changes in interatomic distance. The latter are responsible for most of the ordinary properties of solids. We shall assume that the effect of the rigidly bound electrons on the valence electrons may be described by means of a potential term of the same type as that which takes into account the effect of the nuclei. In other words, it will be assumed that the valence-electron wave function may be determined by use of a Hamiltonian operator in which the effect of the rigid-shell electrons is taken into account by the presence of an ordinary potential function. The validity of this method must be investigated for each solid and will be discussed in particular cases later. It will be seen that this procedure is usually satisfactory for simple substances.

According to these assumptions, the Hamiltonian operator for n valence electrons is

$$H = \sum_{i=1}^n \left(-\frac{\hbar^2}{2m} \Delta_i + V_i \right) + \frac{1}{2} \sum_{i,j=1}^n \frac{e^2}{r_{ij}} + I \quad (1)$$

where the indices i and j are to be summed over all n electrons, $-\hbar^2 \Delta_i / 2m$ is the kinetic energy operator of the i th electron, and V_i , which is the same for all valence electrons, is the potential energy of the i th electron in the field of the nuclei and bound electrons. e^2 / r_{ij} is the coulomb interaction potential between the i th and j th electrons, where e is the absolute value of the electronic charge and r_{ij} the distance between electrons. It should be noted that the cases $i = j$ are excluded in the last summation. Finally, I is a constant representing the interaction between the nuclei and between rigid-shell electrons on different atoms. Although this term does not enter into the determination of the valence-electron wave functions, it is included here to indicate that it may not be neglected when the cohesive energy of the solid is computed. It should be added that both V_i and I contain the internuclear distances parametrically.

Now, the stationary state of lowest energy minimizes the mean-value integral

$$E = \int \Psi^* H \Psi d\tau \quad (2)$$

according to the mean-value theorem. For this reason, the mean-value integral is of importance and merits considerable discussion. Let us break the Hamiltonian (1) into two parts, namely,

$$T = - \sum_i \frac{\hbar^2}{2m} \Delta_i, \quad (3a)$$

$$V = \sum_i V_i + \frac{1}{2} \sum_{i,j} \frac{e^2}{r_{ij}} + I, \quad (3b)$$

where T is the total kinetic energy operator and V is the total potential energy operator. In terms of T and V , Eq. (2) is

$$E = \int \Psi^* T \Psi d\tau + \int \Psi^* V \Psi d\tau. \quad (4)$$

We may write the first integral in the form

$$\frac{\hbar^2}{2m} \sum_i \int |\text{grad}_i \Psi|^2 d\tau \quad (5)$$

if we apply Green's theorem and assume that the surface integral vanishes, as is true in the cases in which we shall be interested. We may conclude that the mean value of T is a positive quantity since the integrand in (5) is positive. The second integral in (4) may be written in the form

$$\int |\Psi|^2 V d\tau, \quad (6)$$

for V is simply a function of the variables of integration. This integral may be regarded as the classical potential energy of the charge distribution

$$\rho = e|\Psi|^2 \quad (7)$$

in the potential field V/e , both the charge distribution and the potential function being static in the $3n$ -dimensional space of all electronic coordinates. Thus, *the mean value of the potential energy may be given a straightforward classical interpretation in terms of $3n$ -dimensional charge and potential distributions.*¹

Since the mean value of T is positive, it follows that ordinary atomic systems are energetically stable only because the mean value of V can be negative. Evidently the first term of (3b), which contains the energy of attractive nuclear-electron forces, is entirely responsible for the fact that (6) may be negative. At first sight, one might suppose that the

¹ Although Ψ is a function of the spin variables as well as the space variables, it is possible to average over spin without affecting this conclusion, since neither T nor V depends upon spin. Thus ρ in (7) may be taken as a simple function of the space variables.

wave function Ψ that minimizes E would distribute the electronic charge in such a way that the integral (6) should be as negative as possible. It is easy to see, however, that this wave function would have a very large kinetic energy associated with it. To minimize (6), it would be necessary to choose Ψ in such a way that ρ is small in regions where V is positive and large in regions where V is negative. This means, however, that Ψ would vary rapidly from point to point or that its gradient would be large. The mean value of (5) would then be large, and it is easy to show (cf. next paragraph) that E would not be as low as possible. On the other hand, suppose that we attempt to minimize E by making the mean kinetic energy as small as possible. In this case, we should choose Ψ to be a very slowly varying function, so that its gradient might be very small. A function of this type would not give a charge distribution ρ that preferentially localizes the electrons in the regions of negative V , so that E again would not be as small as possible. Thus we see that the mean values of T and V tend to counterbalance one another. If the mean value of V is large and negative, the mean value of T is large and positive; if the mean value of T is nearly zero, the mean value of V is also nearly zero. Hence, the function that minimizes E must effect a compromise between both terms.

This competition between T and V is indicated by the uncertainty relation

$$\Delta p_i \Delta q_i \geq \hbar.$$

Any increase in the localization of electrons implies that the Δq_i are decreased, a fact that in turn implies an increase in Δp_i . Since the mean value of T is roughly $\sum_i (\Delta p_i)^2 / 2m$, the kinetic energy increases as the

Δq_i decrease. Suppose, for example, that we have an electron moving in the potential field e/r of a proton. The potential energy is roughly $-e^2/\Delta r$, where Δr is the uncertainty in r , whereas the kinetic energy is roughly $\hbar^2/2m(\Delta r)^2$. The sum of these two energies does not have a minimum when Δr is zero or infinite, which would respectively minimize the two terms, but it has a minimum for $\Delta r = \hbar^2/me^2 \cong 0.5 \cdot 10^{-8}$ cm. Incidentally, this example shows that the scale of atomic size, namely, \hbar^2/me^2 , is determined by the compromise between kinetic and potential energy.

Although the competition between T and V tempers the wave function in such a way that neither of the terms is minimized, the wave function always favors both terms to some extent. Thus, it is found that the wave function for the lowest energy state is generally smooth and yet localizes the charge in suitable regions of potential. Since the potential (3b) becomes very large and positive in regions of the $3n$ -dimen-

sional coordinate space where the variables r_{ij} are very small, we may expect that ρ will generally exhibit a smooth minimum of the type shown in Fig. 1 in the neighborhood of such regions.

Upon these effects, which are related to the variational principle, are superimposed those which arise from the condition that Ψ must be antisymmetric. Roughly, the lowest energy members of the family of antisymmetric functions usually have a higher mean kinetic energy

than some of the functions that might be allowed were it not for the Pauli principle. The change in sign of Ψ under permutation of the variables of any two electrons implies that Ψ usually changes its sign in certain regions of configuration space. The word 'usually' is employed because in some simple exceptional cases, such as that of the lowest state of atomic helium, the antisymmetric condition affects only the

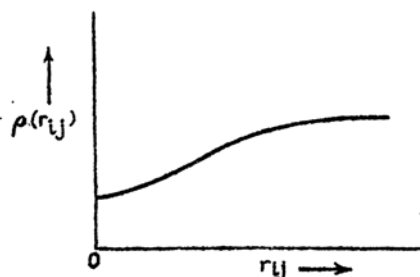


FIG. 1.—It may be expected that the electron density function possesses a smooth minimum of this type because of electron repulsion.

spin variables. The change in sign of Ψ implies in turn that there is a nonvanishing gradient in the same regions, and this implies a contribution to the kinetic energy (5). Unless this contribution is compensated by a drop in the mean potential energy, the best antisymmetric function does not give E an absolute minimum. Thus, the lowest states of all free atoms that involve more than two electrons (that is, atoms beyond helium) are raised because of the Pauli principle. On the other hand, we shall see that the potential energy of antisymmetric valence-electron wave functions in solids is usually lower than the energy of other functions and that the decrease in potential energy sometimes counterbalances the gain in kinetic energy. The origin of this decrease in V may be understood in the following way. As we remarked above, the antisymmetry of Ψ implies that the wave function may change its sign at regions of configuration space where r_{ij} is zero. Hence, the Pauli principle may imply that ρ has a minimum at regions where r_{ij} is zero. This condition, however, is just that which must be satisfied if the positive contribution to the potential from the e^2/r_{ij} terms is to be small.

43. The Helium-atom Problem.—We shall discuss only one of the important problems solved by the variational method, namely, that of helium. The Hamiltonian for this two-electron system is

$$H = -\frac{\hbar^2}{2\mu}\Delta_1 - \frac{\hbar^2}{2\mu}\Delta_2 + \frac{e^2}{r_{12}} - \frac{Ze^2}{r_1} - \frac{Ze^2}{r_2} \quad (1)$$

where r_1 and r_2 are the distances of the two electrons from the nucleus of

charge Ze ($Z = 2$ for He) and μ is the reduced electronic mass, namely,

$$\mu = \frac{mM}{m + M}$$

in which m is the electronic mass and M is the nuclear mass. The other quantities in (1) have been defined previously. In the approximation in which (1) is valid, the wave functions may be written in the form¹

$$\Psi = \Phi(\mathbf{r}_1, \mathbf{r}_2) \sigma(\zeta_1, \zeta_2) \quad (2)$$

where Φ is a function of the space variables, symbolized by \mathbf{r}_1 and \mathbf{r}_2 , and σ is a function of the spin variables. It may be shown from the symmetry of (1) under permutations of electrons that the only forms of Φ that lead to stationary values of $\int \Phi^* H \Phi d\tau$ are either symmetric or antisymmetric under interchange of \mathbf{r}_1 and \mathbf{r}_2 . This characteristic is peculiar to the two-electron problem. We shall designate a symmetrical Φ by Φ_s and an antisymmetrical one by Φ_A . According to the Pauli principle, σ must be antisymmetric when Φ is symmetric and symmetric when Φ is antisymmetric, in order that Ψ may be antisymmetric. There are three symmetric spin functions for two electrons, namely,²

$$\left. \begin{aligned} \sigma_s^1 &= \eta_1(1)\eta_2(1), \\ \sigma_s^0 &= \eta_1(1)\eta_2(-1) + \eta_1(-1)\eta_2(1), \\ \sigma_s^{-1} &= \eta_1(-1)\eta_2(-1), \end{aligned} \right\} \quad (3)$$

and there is one antisymmetric function, namely,

$$\sigma_A^0 = \eta_1(1)\eta_2(-1) - \eta_1(-1)\eta_2(1). \quad (4)$$

These functions are eigenfunctions of the spin operators

$$\Sigma^2 = (\sigma_{x_1} + \sigma_{x_2})^2 + (\sigma_{y_1} + \sigma_{y_2})^2 + (\sigma_{z_1} + \sigma_{z_2})^2 \quad (5)$$

and

$$\Sigma_z = (\sigma_{z_1} + \sigma_{z_2}). \quad (6)$$

The functions (3) correspond to the eigenvalue $2\hbar^2$ of (5) and to the eigenvalues \hbar , 0 and $-\hbar$ of (6), respectively, whereas (4) corresponds to the eigenvalue 0 of both (5) and (6). We may expect that a *singlet state* $\Phi_s \sigma_A^0$ has the lowest energy, since the presence of Φ_s implies a low kinetic energy (cf. Sec. 47).

The most extensive and systematic investigation of the mean value

$$E = \int \Phi_s^* H \Phi_s d\tau \quad (7)$$

¹ For simplicity, we shall designate the spin variable of the i th electron by ζ_i instead of by ζ_{e_i} , as in the preceding chapter.

² The function $\eta_i(\pm 1)$ is identical with the function $\eta(\zeta_{e_i}, \pm 1)$ of the preceding chapter.

has been carried through by Hylleraas,¹ who used functions containing an increasing number of parameters in successive stages of approximation. In the case of helium, his method leads to a total binding energy that agrees with the experimental value to within a few hundredths of a per cent. We shall discuss the various types of function that he employed and shall attempt to interpret his results in terms of the general remarks of the preceding section. In this discussion, it is convenient to replace the six Cartesian coordinates of the electron by the variables

$$\left. \begin{aligned} s &= r_1 + r_2, \\ t &= r_1 - r_2, \\ u &= r_{12}, \end{aligned} \right\} \quad (8)$$

and by three other angular variables upon which the wave function of the lowest state does not depend.

a. First Approximation.—The simplest function used by Hylleraas has the form

$$\Phi_s = e^{-\alpha s} = e^{-\alpha r_1} e^{-\alpha r_2}, \quad (9)$$

which contains only one parameter, namely, α . The use of this function is equivalent to assuming that both electrons move in a coulomb field and have hydrogen-like wave functions corresponding to the 1s state. Since $\alpha = Z/a_h$ for one-electron atoms, where Z is the atomic number of the nucleus and a_h is the radius of the first Bohr orbit in hydrogen ($a_h = 0.531 \text{ \AA}$), the value of $\alpha a_h = Z'$ that gives $\int \Phi_s^* H \Phi_s d\tau$ its minimum value furnishes a value of the "effective nuclear charge" in which each electron moves. It is found that $Z' = Z - \frac{5}{16}$, which is $\frac{27}{16}$ for helium.

The mean value of H is then $-\frac{(Z - \frac{5}{16})^2 e^2}{a_h}$, or 5.695 Rydberg units (1 Rydberg unit equals 13.54 ev), which should be compared with the observed value of 5.810 Rydberg units.

A part of this difference of 0.7 ev per electron may be removed by using a better eigenfunction of the one-electron type. The Hartree-Fock procedure, which will be discussed more fully in the following sections, starts from the assumption

$$\Phi_s = \phi(r_1)\phi(r_2) \quad (10)$$

and determines ϕ in such a way as to minimize (7); that is, it determines the best Φ that may be written in the form (10). The energy associated² with the Hartree wave function is 5.734 Rydberg units, which leaves

¹ E. HYLLEAAS, *Z. Physik*, **54**, 347 (1929). See also H. BETHE, *Handbuch der Physik*, Vol. XXIV/1, p. 324.

² BETHE, *op. cit.*, p. 368.

about two-thirds of the difference, or about 0.45 ev per electron. This difference arises from the fact that electrons really do not move independently of one another but are *correlated*. The wave function not only is a function of the variables $s = r_1 + r_2$ and $t = r_1 - r_2$ but is also a function of $u = r_{12}$. This dependence has been discussed in Sec. 47 and is related to the fact that a potential term of the type e^2/u appears in the Hamiltonian.

b. Second Approximation.—In the next approximation, Hylleraas chose the function

$$\Phi_s = e^{-\alpha s}(1 + a_1 u + b_1 t^2) \quad (11)$$

and found

$$E = -5.805$$

for

$$\begin{aligned} \alpha &= \frac{1.82}{a_h} \\ a_1 &= \frac{0.29}{a_h} \\ b_1 &= \frac{0.13}{a_h^2} \end{aligned}$$

This leaves a difference of only 0.03 ev per electron between the computed and observed values. The function (11) shows the expected correlation between electrons since the variable u enters in it.

c. Higher Approximations.—Hylleraas extended these computations by using a power series in u , s , and t^2 as the coefficient of $e^{-\alpha s}$ and arrived at a limiting energy of 5.80749 Rydberg units which actually is greater than the observed value by about 0.0002 Rydberg unit. This discrepancy does not imply any flaw in the variation method but is due to the fact that the Hamiltonian (1) neglects relativistic effects which would introduce a correction of this order of magnitude.

One important fact that may be gained from this investigation is that the method of one-electron functions yields an energy which is in error by about 0.5 ev per electron, because it does not involve the necessary correlations. Since the binding energy of many solids is of the order of 1 ev per electron, we may expect that the cohesive energies derived from one-electron functions will often have a relatively high percentage of error.

49. The One-electron Approximation.¹—The one-electron scheme has proved to be the most fruitful of several approximate methods that

¹ General discussions of this topic may be found in the tract by L. Brillouin, *Actualités Scientifiques* iv (Hermann et Cie., Paris, 1934), and in E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (Cambridge University Press, 1935).

have been developed for obtaining qualitative and semiquantitative solutions of the Schrödinger equation when many electrons are present. As was mentioned in the introduction to the present chapter, this scheme is based upon a plan for constructing wave equations for an n -electron system from n one-electron wave functions.

It was assumed in the introductory development of the one-electron approximation that the total state function can be represented by the product of n one-electron functions. Thus,

$$\Psi(x_1, y_1, z_1, \dots, x_n, y_n, z_n) = \psi_1(x_1, y_1, z_1) \psi_2(x_2, y_2, z_2) \dots \psi_n(x_n, y_n, z_n). \quad (1)$$

Here the ψ_i are the normalized one-electron functions and x_i, y_i, z_i are the spatial coordinates of the i th electron. Spin variables were not explicitly introduced. In accordance with the Pauli principle, as it had been introduced into the Bohr theory, it was assumed that no more than two electrons can have identical functions and that any two electrons that do have identical ψ_i have opposite spins. The function (1) does not include electronic correlations explicitly since the state function for the i th electron is ψ_i , regardless of the position of the other electrons.

Hartree¹ suggested, on the basis of plausibility, that each one-electron function in (1) should satisfy a one-electron Schrödinger equation in which the potential includes a term that takes into account the coulomb field of the other electrons as well as the fields arising from nuclei and other charged particles. He chose this term as the classical electrostatic potential of the $n - 1$ normalized charge distributions $|\psi_i|^2$. In other words, his equation for ψ_i is²

$$-\frac{\hbar^2}{2m} \Delta \psi_i + \left[V_i(x_i, y_i, z_i) + \sum_j' e^2 \int \frac{|\psi_j|^2}{r_{ij}} d\tau_j \right] \psi_i = \epsilon_i \psi_i \quad (2)$$

where V_i is the field arising from nuclei, etc. Hartree developed a practical method, now known as the method of the self-consistent field, for solving the set of simultaneous equations, and applied this procedure to a number of atoms in a large-scale program of investigating the periodic chart. This program is still under way, although modified equations, which we shall discuss below, are now used in place of (2).

¹ D. R. HARTREE, *Cambridge Phil. Soc.*, **24**, 89 (1928).

² In the rest of this chapter, we shall designate the volume element for the i th electron by $d\tau_i$ when spin is not included and by $d\tau_i'$ when spin is included. In cases in which ambiguity may occur, however, we shall use the notation of the preceding chapter, namely, $d\tau(x_i, y_i, z_i)$, etc. Similarly, the volume element for the coordinates of the i th and j th electrons will sometimes be written as $d\tau_i d\tau_j$ and at other times as $d\tau_{ij}$.

The mean value of the Hamiltonian

$$H = \sum_i \left(-\frac{\hbar^2}{2m} \Delta_i + V_i \right) + \frac{1}{2} \sum_{i,j}' \frac{e^2}{r_{ij}} \quad (3)$$

for the function (1) is seen to be simply

$$E = \sum_{i=1}^n \left(-\frac{\hbar^2}{2m} \int \psi_i^* \Delta \psi_i d\tau_i + \int \psi_i^* V \psi_i d\tau_i \right) + \frac{1}{2} \sum_{i,j}' e^2 \int \int \frac{|\psi_i|^2 |\psi_j|^2}{r_{ij}} d\tau_i d\tau_j, \quad (4)$$

which reduces to

$$E = \sum_{i=1}^n \epsilon_i - \frac{1}{2} \sum_{i,j}' e^2 \int \int \frac{|\psi_i|^2 |\psi_j|^2}{r_{ij}} d\tau_i d\tau_j \quad (5)$$

when the ψ_i satisfy Hartree's equations.

There is only one equation of the type (2) for the normal state of helium since the two ψ are identical in this case. The energy (5) that was derived by using the solution of this equation was discussed in part *a* of the previous section.

It was recognized during the period after Hartree's first work that the Pauli principle has a more natural position in the new quantum theory than in the old. In accordance with the discussion of Chap. V, the principle is taken to imply that all wave functions must be antisymmetric under permutation of all electronic coordinates, including the spin variables. The function (1) is not a satisfactory wave function from this standpoint, for it is not antisymmetric. An allowable antisymmetric function may be constructed from the same set of one-electron functions, however, if each ψ_i is replaced by a function φ_i that is the product of ψ_i and a spin function $\eta_i(\zeta_i)$, thus:

$$\varphi_i(\mathbf{r}_i) = \psi_i(x_i, y_i, z_i) \eta_i(\zeta_i) \quad (6)$$

where \mathbf{r}_i in the left-hand side of (6) stands for the four variables x_i, y_i, z_i, ζ_i . We shall assume in the following discussion that the η_i are eigenfunctions of σ_z , the z component of spin. Hence, they may be labeled as $\eta_i(+1)$ or $\eta_i(-1)$, in the cases going with the eigenvalues $\hbar/2$ and $-\hbar/2$, respectively. The antisymmetric function that may be constructed from the φ_i is

$$\Psi = \sum_P (-1)^P P \cdot [\varphi_1(\mathbf{r}_1) \varphi_2(\mathbf{r}_2) \cdots \varphi_n(\mathbf{r}_n)] \quad (7)$$

where P runs over the $n!$ permutations of the n variables and p is the parity¹ of the p th permutation. This sum has the property that it may be written in the form of a determinant

$$\Psi = \begin{vmatrix} \varphi_1(\mathbf{r}_1)\varphi_1(\mathbf{r}_2) & \cdots & \varphi_1(\mathbf{r}_n) \\ \varphi_2(\mathbf{r}_1)\varphi_2(\mathbf{r}_2) & \cdots & \varphi_2(\mathbf{r}_n) \\ \vdots & & \vdots \\ \varphi_n(\mathbf{r}_1)\varphi_n(\mathbf{r}_2) & \cdots & \varphi_n(\mathbf{r}_n) \end{vmatrix}, \quad (8)$$

in which the φ_i are elements. The antisymmetric properties of Ψ are now evident from the properties of determinants, for the process of interchanging two columns, which reverses the sign of Ψ , corresponds to a permutation of the corresponding variables. Moreover, Ψ vanishes identically when two φ_i are equal. Hence, the Pauli principle, as it was employed in the Bohr theory, is automatically satisfied.

It may be shown from elementary principles of group theory² that (7) is the only antisymmetric combination of the φ_i . This function is not usually the only antisymmetric combination of the ψ_i , however, for it may be possible to assign spin functions to the ψ_i in more than one way. This possibility will be discussed further in the following sections.

The electrons do not move independently so long as there is more than one independent term in (7); that is, electronic motion is correlated in the antisymmetric wave function even though this function is composed of one-electron functions. These correlations are more or less accidental, for they arise from the Pauli principle rather than from the requirement that the electrons should keep away from one another. It turns out that these accidental correlations sometimes favor cohesion by keeping the electrons apart and sometimes hinder it by piling the electrons together. We shall have occasion to examine particular cases in detail in later chapters.

The function (8) is not normalized when the φ_i are normalized, so that (8) must be multiplied by a constant. Although this constant usually depends upon the choice of φ_i , it is simply $1/\sqrt{n!}$ in the particular case in which the φ_i are orthogonal, that is, when

$$\int \varphi_i^* \varphi_j d\tau(x, y, z, \zeta) = \delta_{ij}. \quad (9)$$

The integration in (9) implies a summation over the two values of spin

¹ The parity of a permutation is the number of interchanges that must be made in order to obtain the permutation from the standard arrangement. Thus, the parity of the permutation 2143 of the integers 1234 is 2, since the permutation may be obtained by interchanging 1 and 2, and 3 and 4, respectively.

² E. P. WIGNER, *Gruppentheorie* (Vieweg, Braunschweig, Germany, 1931).

variable ξ . If Ψ is given by (7),

$$\int \Psi^* \Psi d\tau' = \int \sum_P \sum_{P'} (-1)^{p+p'} P[\varphi_1^*(\mathbf{r}_1) \cdots \varphi_n^*(\mathbf{r}_n)] P'[\varphi_1(\mathbf{r}_1) \cdots \varphi_n(\mathbf{r}_n)] d\tau', \quad (10)$$

which is equal to

$$\sum_P \int P |\psi_1(\mathbf{r}_1)|^2 |\psi_2(\mathbf{r}_2)|^2 \cdots |\psi_n(\mathbf{r}_n)|^2 d\tau = n!$$

because of (9).

We shall assume hereafter that the condition (9) is satisfied. This does not place any important restriction upon the φ_i since those which have opposite spins are automatically orthogonal and those which have parallel spins may be made orthogonal by the Schmidt method.¹ This orthogonalization process does not affect (8), for a determinant remains unchanged if a constant multiple of the elements in one row is added to the corresponding elements in another, and the Schmidt method is equivalent to doing this.

Let us split the Hamiltonian into kinetic and potential energy parts, as in Sec. 47. The mean value of the former, namely,

$$T = -\frac{\hbar^2}{2m} \sum_{i=1}^n \Delta_i$$

is

$$\begin{aligned} & -\sum_{i=1}^n \frac{\hbar^2}{2m} \int \frac{1}{n!} \left[\sum_P (-1)^p P(\varphi_1^*(\mathbf{r}_1) \varphi_2^*(\mathbf{r}_2) \cdots \varphi_n^*(\mathbf{r}_n)) \right] \Delta_i \cdot \\ & \quad \left[\sum_{P'} (-1)^{p'} P'(\varphi_1(\mathbf{r}_1) \cdots \varphi_n(\mathbf{r}_n)) \right] d\tau(x_1, \cdots, z_n, \xi_1, \cdots, \xi_n) \\ & = -\sum_{i=1}^n \frac{\hbar^2}{2m} \int \varphi_i^* \Delta_i \varphi d\tau(x, y, z, \xi) = -\sum_{i=1}^n \frac{\hbar^2}{2m} \int \psi_i^* \Delta_i \psi d\tau(x, y, z) \quad (11) \end{aligned}$$

when the eigenfunction is (7). This result is exactly the same as that derived by use of (1). Hence, the mean value of T is unaffected by the use of a determinantal eigenfunction. The same statement is true for

that part of the potential energy which can be written in the form $\sum_{i=1}^n V_i$,

¹ See, for example, R. COURANT and D. HILBERT, *Methoden der mathematische Physik* (Julius Springer, Berlin, 1924).

where V_i depends only upon the coordinates of the i th electron and is the same function of these as V_j is of the coordinates of the j th electron, as can be seen from the fact that the form of (11), as a sum of one-electron integrals, depends only upon the property that T is a sum of one-electron operators. Hence,

$$\sum_i \bar{V}_i = \sum_i \int \varphi_i^* V(\mathbf{r}) \varphi_i d\tau(x, y, z, \zeta). \quad (12)$$

If V is independent of spin, this is

$$\sum_i \int \psi_i^* V(\mathbf{r}) \psi_i d\tau(x, y, z). \quad (13)$$

The mean value of those terms of the Hamiltonian, such as

$$\frac{1}{2} \sum_{i,j}' \frac{e^2}{r_{ij}}, \quad (14)$$

in which the elements of summation depend upon the coordinates of two electrons, is affected by the use of an antisymmetric function. The mean value of a typical term of (14), e^2/r_{12} , is

$$\begin{aligned} \frac{1}{n!} \int \left[\sum_P (-1)^P P(\varphi_1^*(\mathbf{r}_1) \cdots \varphi_n^*(\mathbf{r}_n)) \right] \frac{e^2}{r_{12}} \cdot \\ \left[\sum_{P'} (-1)^{P'} P'(\varphi_1(\mathbf{r}_1) \cdots \varphi_n(\mathbf{r}_n)) \right] d\tau' = \\ \frac{1}{n!} \sum_{P, P'} (-1)^{P+P'} \int P[\varphi_1^*(\mathbf{r}_1) \cdots \varphi_n^*(\mathbf{r}_n)] \frac{e^2}{r_{12}} P'[\varphi_1(\mathbf{r}_1) \cdots \varphi_n(\mathbf{r}_n)] d\tau'. \quad (15) \end{aligned}$$

If a given permutation P sends \mathbf{r}_i into \mathbf{r}_1 and \mathbf{r}_j into \mathbf{r}_2 and the remaining $n-2$ variables into $\mathbf{r}_3, \dots, \mathbf{r}_n$ in some particular way, the integral in (15) that contains this P vanishes unless the P' in the same integral sends the same set into $\mathbf{r}_3, \dots, \mathbf{r}_n$ and either \mathbf{r}_i into \mathbf{r}_1 and \mathbf{r}_j into \mathbf{r}_2 or \mathbf{r}_i into \mathbf{r}_2 and \mathbf{r}_j into \mathbf{r}_1 . When P' does satisfy this condition, the integration over the variables $\mathbf{r}_3, \dots, \mathbf{r}_n$ reduces (15) to

$$\frac{e^2}{n(n-1)} \sum_{i,j} \left[\int \frac{\varphi_i^*(\mathbf{r}_1) \varphi_j^*(\mathbf{r}_2) \varphi_i(\mathbf{r}_1) \varphi_j(\mathbf{r}_2)}{r_{12}} d\tau'_{12} - \int \frac{\varphi_i^*(\mathbf{r}_1) \varphi_j^*(\mathbf{r}_2) \varphi_i(\mathbf{r}_2) \varphi_j(\mathbf{r}_1)}{r_{12}} d\tau'_{12} \right]$$

because of (9). Since the result is the same for the other $n(n-1)$ terms in the summation (14), the mean value of this quantity is

$$\frac{e^2}{2} \sum_{i,j}' \left[\int \frac{|\varphi_i(\mathbf{r}_1)|^2 |\varphi_j(\mathbf{r}_2)|^2}{r_{12}} d\tau'_{12} - \int \frac{\varphi_i^*(\mathbf{r}_1) \varphi_j^*(\mathbf{r}_2) \varphi_i(\mathbf{r}_2) \varphi_j(\mathbf{r}_1)}{r_{12}} d\tau'_{12} \right]. \quad (16)$$

The summation over the spins in the first terms may be performed at once, giving

$$\frac{e^2}{2} \sum_{i,j}' \int \frac{|\psi_i(\mathbf{r}_1)|^2 |\psi_j(\mathbf{r}_2)|^2}{r_{12}} d\tau(x_1, \dots, z_n). \quad (17)$$

The corresponding sum in the second term is zero whenever $\eta_i \neq \eta_j$; however, we shall leave this result in the form given in (16) for the present.

The expression (17) is exactly the same as the last term in (4), which was derived by taking the mean value of (14) for the simple product function (1). Hence, the only effect on the mean value of H of using the determinantal form of Ψ is the introduction of the *exchange energy*

$$-\frac{e^2}{2} \sum_{i,j}' \int \frac{\varphi_i^*(\mathbf{r}_1) \varphi_j^*(\mathbf{r}_2) \varphi_i(\mathbf{r}_2) \varphi_j(\mathbf{r}_1)}{r_{12}} d\tau'_{12}. \quad (18)$$

This term results, not from any unusual nonclassical force between electrons, but because we have used the determinantal eigenfunction instead of (1). As we said previously, the product function (1) contains no interelectronic correlations, whereas the determinantal function (7) does contain them. The exchange term (18) is simply the contribution that these correlations make to the energy. If the exchange energy is negative, the charge distribution $e|\Psi|^2$, corresponding to (7), has a lower self-potential than the charge distribution corresponding to (1), for the accidental correlations in (7) keep the electrons apart. On the other hand, if (18) is positive, the accidental correlations in (7) raise the self-energy of the charge distribution, because they push electrons together. We cannot predict the sign of (18) without knowing the form of the φ_i ; each case must be investigated separately.

In order to illustrate the ways in which exchange terms are related to correlations, we shall derive the expression for the probability density of two electrons in a set of n . Let \mathbf{r}_1 and \mathbf{r}_2 be the coordinates of the electrons and $P(\mathbf{r}_1, \mathbf{r}_2)$ the probability density. To find $P(\mathbf{r}_1, \mathbf{r}_2)$ we must integrate $|\Psi|^2$ over all variables except \mathbf{r}_1 and \mathbf{r}_2 . The result, which may be derived easily by the methods that were used in obtaining (16), is

$$P(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{n(n-1)} \left[\sum_{i,j}' |\varphi_i(\mathbf{r}_1)|^2 |\varphi_j(\mathbf{r}_2)|^2 - \sum_{\substack{i,j \\ \parallel \text{ spins}}} \varphi_i^*(\mathbf{r}_1) \varphi_j^*(\mathbf{r}_2) \varphi_i(\mathbf{r}_2) \varphi_j(\mathbf{r}_1) \right] \quad (19)$$

where the second sum extends only over pairs of φ_i with parallel spin. It follows from the manner of deriving $P(\mathbf{r}_1, \mathbf{r}_2)$ that the coulomb and exchange energy is simply $n(n-1)/2$ times the integral

$$e^2 \int \frac{P(\mathbf{r}_1, \mathbf{r}_2)}{r_{12}} d\tau_1 d\tau_2, \quad (20)$$

which is the self-energy of the charge distribution $eP(\mathbf{r}_1, \mathbf{r}_2)$.

Let us evaluate (19) for the special case of perfectly free electrons in a large cubical box. If we use periodic boundary conditions, the wave functions are

$$\psi_{\mathbf{k}} = \frac{1}{\sqrt{V}} e^{2\pi i \mathbf{k} \cdot \mathbf{r}}. \quad (21)$$

Here, V is the volume of the box and \mathbf{k} is the electronic wave-number vector, the components of which take the discrete values

$$k_x = \frac{n_x}{L}, \quad k_y = \frac{n_y}{L}, \quad k_z = \frac{n_z}{L} \quad (22)$$

where n_x, n_y, n_z are arbitrary integers and L is the length of an edge of the box. We shall assume that all values of \mathbf{k} that lie within a sphere of radius k_0 appear in (19) with both spins. This system of free electrons evidently is equivalent to the system used in the simple Sommerfeld model of a metal.

Since $|\psi_{\mathbf{k}}|^2 = 1/V$, the first sum in (19) is equal to the constant value $1/V^2$. The second sum may be written

$$\sum_{\mathbf{k}_1, \mathbf{k}_2} e^{2\pi i (\mathbf{k}_1 - \mathbf{k}_2) \cdot (\mathbf{r}_1 - \mathbf{r}_2)} \quad (23)$$

where \mathbf{k}_1 and \mathbf{k}_2 are to be summed over all values in the sphere of radius k_0 . If the number of electrons is sufficiently large and the levels are sufficiently dense, this sum may be replaced¹ by an integral and then reduced to

$$\frac{9}{2} \left[\frac{2\pi k_0 r \cos 2\pi k_0 r - \sin 2\pi k_0 r}{(2\pi k_0 r)^3} \right]^2$$

where $r = |\mathbf{r}_1 - \mathbf{r}_2|$. Hence,

$$P(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{V^2} \left\{ 1 - \frac{9}{2} \left[\frac{2\pi k_0 r \cos 2\pi k_0 r - \sin 2\pi k_0 r}{(2\pi k_0 r)^3} \right]^2 \right\}.$$

The coefficient of $1/V^2$ is plotted in Fig. 2 as a function of r . The

¹ E. P. WIGNER and F. SEITZ, *Phys. Rev.*, **43**, 804 (1933).

probability that the electrons will be at the same point is just half the probability that they will be far apart and varies smoothly in between. The second term in (19) is entirely responsible for the correlation term which, in turn, gives rise to the exchange energy when (20) is computed. In cases in which the electrons are not entirely free, some correlation effect is provided by the first terms in (19).

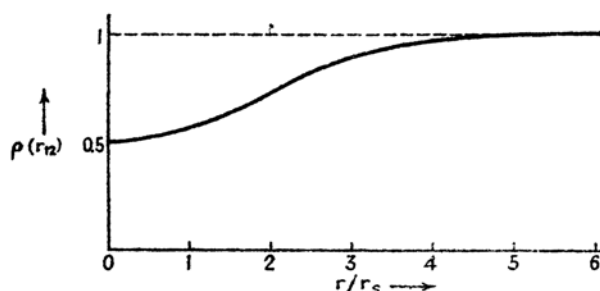


FIG. 2.—The relative probability of finding two perfectly free electrons at a distance r from one another. The parameter r_s is the radius of the sphere whose volume is equal to the mean volume per electron. In this case the correlation effect corresponding to the minimum near $r = 0$ arises from exchange.

The introduction of antisymmetric wave functions detracts considerably from the plausibility of Hartree's equations (2) since they do not take into account the correlations that give rise to the exchange energy. For this reason, Fock¹ and Slater² suggested independently that the variational theorem should be used to derive a set of equations for the best one-electron functions. We shall discuss these equations in Sec. 51, after investigating the question of multiplicity in the next section. We shall see that Hartree's equations are satisfactory when Ψ has the form (1) but that additional terms must be added to these when the wave function has the form (7).

50. Eigenfunctions of Definite Multiplicity.—It was pointed out in the discussion of Sec. 41, Chap. V, that the antisymmetric stationary states of any system can be chosen as eigenfunctions of both Σ^2 and Σ_z , as well as of H , as long as the Hamiltonian is independent of spin. It was also pointed out that the functions that have the same eigenvalue $\hbar^2 S(S+1)$ of Σ^2 can be divided into groups of $2S+1$ which have the same energy. This $(2S+1)$ -fold degenerate level is said to have multiplicity $2S+1$. The constituent states may be chosen as eigenfunctions of Σ_z , and there is then one state for each of the $2S+1$ possible eigenvalues of Σ_z , which range from S to $-S$ with integer differences. States of different multiplicity usually have different energies except in the special case in which accidental degeneracy occurs. For these

¹ V. FOCK, *Z. Physik*, **61**, 126 (1930). See also footnote 1, p. 234.

² J. C. SLATER, *Phys. Rev.*, **35**, 210 (1930).

reasons, the eigenfunctions of H , that is, the functions that minimize the integral

$$\int \Psi^* H \Psi d\tau'$$

are usually eigenfunctions of Σ^2 . The eigenfunctions may be chosen as states of definite multiplicity, however, even when there is accidental degeneracy. Hence, we shall consider this to be a general condition on Ψ .

The antisymmetric function (7) of the preceding section usually is not an eigenfunction of Σ^2 when the ψ_i are different and the η_i are selected at random. The determinant has unit multiplicity (that is, $S = 0$), however, in the special case in which the ψ_i are equal in pairs and the spins of the members of equal pairs are opposite.¹ Hence, the determinant is a satisfactory function from the standpoint of multiplicity in this particular case. Fortunately, this case is an important one for all simple solids, since they have unit multiplicity in the normal state. As a result, the form of Fock's equations, which is discussed in the next section and is based upon a determinantal wave function, is valid for the normal state of simple solids. The exceptional solids are ferromagnetic and strongly paramagnetic substances.

51. Fock's Equations.—The results that are obtained by following Fock and Slater's plan for determining the best one-electron functions from the variational theorem are derived in Appendix I. It is found that these results depend upon the initial choice of the complete wave function. If the function has the form (1), Sec. 49, the equations that the ψ_i must satisfy turn out to be Hartree's equations (2). If, in addition, we specify that the different ψ are to be orthogonal or that

$$\int \psi_i^* \psi_j d\tau = \delta_{ij}, \quad (1)$$

the best equations are

$$-\frac{\hbar^2}{2m} \Delta \psi_i + \left(V_i + \sum_j' e^2 \int \frac{|\psi_j|^2}{r_{ij}} d\tau_j \right) \psi_i = \epsilon_i \psi_i + \sum_j' \lambda_{ij} \psi_j. \quad (2)$$

These equations have the same form as Hartree's except for the presence of the term

$$\sum_j' \lambda_{ij} \psi_j,$$

which arises from the condition (1). The λ_{ij} are the Lagrangian parameters associated with the orthogonality stipulation.

The Pauli principle is not properly included in either of the total wave functions on which Hartree's equations and the equations (2) are based. For reasons discussed in Chap. V, it is necessary to use an anti-

¹ See footnote 2, p. 210.

symmetric function of the electron coordinates. When this is done by employing a determinantal function of the type (7) of Sec. 49 with the condition

$$\int \varphi_i \varphi_j d\tau(x, y, z, \zeta) = \delta_{ij}, \quad (3)$$

the best equations are found to be (cf. Appendix I)

$$-\frac{\hbar^2}{2m} \Delta \varphi_i(\mathbf{r}_1) + \left[V(\mathbf{r}_1) + \sum_j e^2 \int \frac{|\varphi_j(\mathbf{r}_2)|^2}{r_{12}} d\tau_2' \right] \varphi_i(\mathbf{r}_1) - \sum_j \left[e^2 \int \frac{\varphi_i(\mathbf{r}_2) \varphi_j^*(\mathbf{r}_2)}{r_{12}} d\tau_2' \right] \varphi_j(\mathbf{r}_1) = \epsilon_i \varphi_i(\mathbf{r}_1) + \sum_j \lambda_{ij} \varphi_j(\mathbf{r}_1). \quad (4)$$

The parameters λ have the same significance as those in (2), whereas the essentially new terms, namely,

$$- \sum_j \left[e^2 \int \frac{\varphi_i(\mathbf{r}_2) \varphi_j^*(\mathbf{r}_2)}{r_{12}} d\tau_2' \right] \varphi_j(\mathbf{r}_1), \quad (5)$$

which we shall call "exchange terms," arise from the use of a determinantal eigenfunction. It should be noted that the integrals in (5) are functions of \mathbf{r}_1 . The exchange terms may be regarded as "non-conventional" potential integrals that take into account the accidental correlations of the determinantal wave function, just as the exchange integrals of Sec. 49 may be regarded as taking into account the change in self-energy caused by these correlations.

Equations (4), which we shall henceforth call "Fock's equations," have many symmetrical properties not possessed by Hartree's equations. If we add the expression

$$\left[e^2 \int \frac{|\varphi_i(\mathbf{r}_2)|^2}{r_{12}} d\tau_2' \right] \varphi_i(\mathbf{r}_1)$$

to the first summation on the left-hand side of (4) and subtract the same expression from the second summation, we obtain

$$-\frac{\hbar^2}{2m} \Delta \varphi_i(\mathbf{r}_1) + \left[V + \sum_j e^2 \int \frac{|\varphi_j(\mathbf{r}_2)|^2}{r_{12}} d\tau_2' \right] \varphi_i(\mathbf{r}_1) - \sum_j \left[e^2 \int \frac{\varphi_j^*(\mathbf{r}_2) \varphi_i(\mathbf{r}_2)}{r_{12}} d\tau_2' \right] \varphi_j(\mathbf{r}_1) = \sum_j \lambda_{ij} \varphi_j(\mathbf{r}_1), \quad (6)$$

in which $\lambda_{ii} = \epsilon_i$ and none of the sums is primed. Next, let us set

$$\rho(\mathbf{r}_1, \mathbf{r}_2) = \sum_j \varphi_j^*(\mathbf{r}_1) \varphi_j(\mathbf{r}_2), \quad (7)$$

which is known as "Dirac's density matrix." In terms of this function, the first summation in (6) is simply

$$\left[e^2 \int \frac{\rho(\mathbf{r}_2, \mathbf{r}_2)}{r_{12}} d\tau'_2 \right] \varphi_i(\mathbf{r}_1), \quad (8)$$

and the second is

$$-e^2 \int \frac{\rho(\mathbf{r}_2, \mathbf{r}_1) \varphi_i(\mathbf{r}_2)}{r_{12}} d\tau'_2. \quad (9)$$

Following Dirac,¹ we shall define an operator A in terms of (9) by the relation

$$A \varphi_i(\mathbf{r}_1) = -e^2 \int \frac{\rho(\mathbf{r}_2, \mathbf{r}_1) \varphi_i(\mathbf{r}_2)}{r_{12}} d\tau'_2. \quad (10)$$

Thus, the operator A , acting upon $\varphi_i(\mathbf{r}_1)$, multiplies this function by $\rho(\mathbf{r}_2, \mathbf{r}_1)$, changes the variable \mathbf{r}_1 to \mathbf{r}_2 , and integrates over \mathbf{r}_2 . It may be readily verified that A is both linear and Hermitian. The integral that appears in (8) we shall designate by U ; it is simply the coulomb potential of the three-dimensional charge distribution. If we use the foregoing terminology, (6) reduces to

$$H^F \varphi_i(\mathbf{r}_1) = \sum_j \lambda_{ij} \varphi_j(\mathbf{r}_1) \quad (11)$$

where H^F is the *Fock Hamiltonian operator*, namely,

$$H^F = -\frac{\hbar^2}{2m} \Delta + V + U + A, \quad (12)$$

which is the same for all electrons in the system.

The parameters λ_{ij} in (11), aside from λ_{ii} , should be selected in such a way as to ensure the orthogonality of the φ_i . A possible choice of these parameters when the spectrum is non-degenerate is $\lambda_{ij} = 0 (i \neq j)$, since functions that satisfy the equation

$$H^F \varphi_i = \epsilon_i \varphi_i$$

are orthogonal because H^F is Hermitian. This is not the only possible choice of λ_{ij} , however, and others have been used on occasion; but we shall find that this choice is a convenient one in a large part of the following discussion.

The difficulties encountered in solving Hartree's equations also arise with Fock's equations, for it is usually necessary to obtain a solution by some method of successive approximations, such as Hartree's method of the self-consistent field. This procedure is more difficult to apply to Fock's equations because the exchange terms introduce many complications.

¹ P. A. M. DIRAC, *Cambridge Phil. Soc.*, **26**, 376 (1930).

In the next section, we shall give a brief summary of the solutions of Hartree's and Fock's equations for free atoms. The actual technique does not interest us so much as the results and their deviation from experimental results, for these give us an estimate of the error that may be expected when the equations are solved for solids.

52. The Solutions of Hartree's and Fock's Equations for Single Atoms.—The simplest nontrivial problem to which the methods of Hartree and Fock-Slater are applicable is that of the normal state of helium, which we have discussed in Sec. 48. Hartree's and Fock's equations are identical in this case since the electrons have antiparallel spins and the exchange terms are zero. The total energy of the atom, as given¹ by this approximation, is found to be 0.076 Rydberg unit higher than the observed value of 5.810 Rydberg units, a fact indicating that electronic correlations are important to the extent of 0.45 ev per electron. Henceforth, we shall call an energy difference such as this 0.45 ev, which measures the error in the energy derived from a one-electron approximation, a "correlation energy." The connotation of this term is evident from the discussion in preceding sections.

The method employed to determine the one-electron function ψ in the case of helium is characteristic of the self-consistent-field computations of Hartree and his school. This procedure is completely described by Condon and Shortley² and will not be thoroughly discussed here. It need only be mentioned that the procedure consists, essentially, in assuming a starting function for each electron, determining the potential integrals appearing in Hartree's equation from these, solving the equations for the new wave functions, and comparing these functions with the ones originally assumed. If the two sets agree, the system is said to be self-consistent and the equations solved; if not, the procedure is repeated until the initial and final functions do agree. Naturally, there is no fixed plan that ensures that this procedure will converge rapidly, since a great deal depends upon a good choice of starting functions. Other workers, such as Brown³ and Torrance,⁴ have developed variations of the scheme originally used by Hartree. All these methods involve practically the same steps of approximation and will be regarded as the same here.

We are fundamentally more interested in the solutions of Fock's equations than in those of Hartree's equations, since the former should lead to more accurate results. Solutions of one or both of these equations have been obtained for a number of atoms listed at the end of this

¹ HARTREE, *op. cit.*

² CONDON and SHORTLEY, *op. cit.*

³ F. W. BROWN, *Phys. Rev.*, **44**, 214 (1933).

⁴ C. C. TORRANCE, *Phys. Rev.*, **46**, 388 (1934).

section. Among these, those for beryllium and carbon, determined respectively by Hartree and Hartree¹ and by Torrance, are of principal interest since the absolute binding energies of these atoms are known.

We shall use the results of the atomic computations to study two topics, namely, the accuracy of the energy states of a given atom containing n electrons relative to the lowest energy state of that atom with $n - 1$ electrons, and the error in the absolute energy of an entire atom. The first of the two quantities gives a rough estimate of the relative accuracy of Hartree's and Fock's equations, whereas the second indicates the absolute error and will give a correlation energy. To date, beryllium and carbon are the only cases, other than helium, in which the absolute energy has been completely investigated from the theoretical standpoint.

When an atomic configuration involves only closed shells, it may easily be proved² that Fock's equations possess a self-consistent solution for which

$$\psi = R_l(r)\Theta_{lm}(\theta)\Phi_m(\varphi)$$

where R , Θ , Φ are functions of each of the spherical coordinates, respectively, and the spherical harmonics $\Theta_{lm}(\theta)$, $\Phi_m(\varphi)$ are chosen to agree with the conventions of the one-electron approximation for atomic spectra. This theorem is not valid for Hartree's equations, for they do not possess the same symmetry as Fock's equations. The effective potential for an electron in a closed shell arising from the same closed shell is not spherically symmetrical in Hartree's equations, for the summations in these equations are primed. However, when dealing with closed shells, Hartree generally takes only the spherically symmetrical part of the potential in order that the equations may be separated in spherical polar coordinates. For this reason, his results are not exact solutions except in special cases, such as the normal states of helium and beryllium, in which the configurations are closed shells of s functions. We shall now proceed to discuss particular cases.

*a. Beryllium.*³—Both Fock's and Hartree's equations have been solved for the normal $1s^2 2s^2$ state of beryllium to a high degree of accuracy. The resulting $1s$ functions are very nearly alike, but the $2s$ functions show considerable difference. The two types of $2s$ function are difficult to compare with one another because the solution of Fock's equation is orthogonal to the $1s$ function, whereas the solution of Hartree's is not. The latter may be made orthogonal to the $1s$, however, and the results are given in Fig. 3.

The total energies of Be and of Be^{++} , as determined by different methods, appear in Table XLVII.

¹ D. R. HARTREE and W. HARTREE, *Proc. Roy. Soc.*, **150A**, 9 (1935).

² CONDON and SHORTLEY, *op. cit.*

³ HARTREE and HARTREE, *op. cit.*

The difference between the total energy of beryllium as computed from Fock's equations and the observed value is about 0.19 Rydberg unit, or about 2.56 ev, whereas the difference between the two values for Be^{++} is 0.07 Rydberg unit, or 0.9 ev. In order to obtain an estimate of the correlation energy per electron, we shall divide the first of these by 4 and the second by 2, obtaining 0.65 ev and 0.45 ev, respectively. It

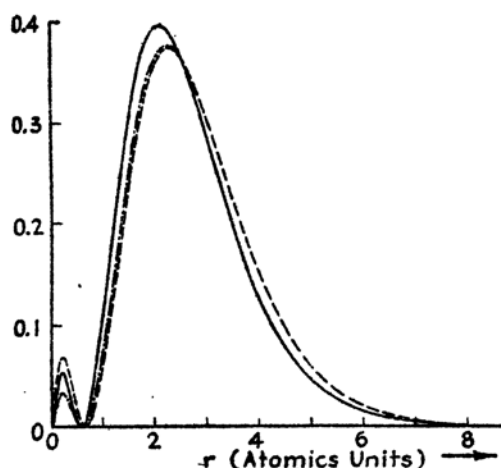


FIG. 3.—A comparison of the square of the 2s functions of beryllium obtained by solving Hartree's and Fock's equations. The full curve represents the solution of Fock's equations; the broken curves represent the orthogonalized and non-orthogonalized solutions of Hartree's equations.

might seem at first sight that it would be more proper to divide the first by 4! and the second by 2!, since there are 4! and 2! interacting pairs in each case. This procedure would not be so reasonable, however, since the correlation effect is larger for electrons in the same shell than for electrons in different shells. According to these results, the mean

TABLE XLVII

	Hartree	Fock	Experimental
$E(\text{Be})$, R.u.	-29.115	-29.140	-29.331 ± 0.008
$E(\text{Be}^{++})$	-27.235	-27.235	-27.307 ± 0.008

correlation energy increases slightly as the number of electrons increases, a fact showing that the one-electron approximation becomes less accurate. The correlation energy of 0.07 Rydberg unit for Be^{++} is almost exactly the same as the value of 0.077 for He.

If we assume that the error in $E(\text{Be}^{++}) - E(\text{Be})$ arises purely from the correlations between the 2s electrons, this correlation energy is found

to be 0.119 Rydberg unit, or about 0.81 ev per electron. Part of this error actually arises from correlations of the 1s and the 2s electrons, but this is probably a small fraction.

b. Carbon.—Calculations of the energy levels of carbon have been carried out by Torrance¹ and Ufford.² Torrance solved Fock's equations for the $1s^2 2s^2 2p^2$ configuration. The energies of the 3P , 1D , and 1S states were evaluated by Ufford using the functions obtained from the solution. Since this is not a closed-shell configuration, Torrance had to replace the asymmetric fields that arise from p electrons by spherically symmetrical ones. In addition, he used the form of Fock's equations discussed in Sec. 51, which is valid for a single determinantal eigenfunction, although the wave functions of the lowest states do not actually have this form.³ For these reasons, Ufford's results should be somewhat higher than those that would be derived from more accurate one-electron functions. The computed cohesive energy of the normal atom is 1,019.66 ev, which may be compared with the observed one of 1,024.84 ev. The difference gives a total correlation energy of 5.18 ev, which is about 0.86 ev per electron.

Torrance also computed one-electron functions for the $1s^2 2s 2p^3$ configuration from which Ufford was able to derive energy values for excited states of the carbon atom. In addition, Ufford computed matrix components of the Hamiltonian between the two configurations and determined new energy levels by taking into account the perturbing effect that the configurations exert on one another. This procedure is equivalent to using new wave functions that are linear combinations of the unperturbed wave functions for the normal and excited states. The relative positions of the levels

in both approximations are shown in Fig. 4 and are compared with the observed values. The computed cohesive energy of the atom is changed from 1,019.66 ev to 1,020.09 ev by the interaction between configurations,

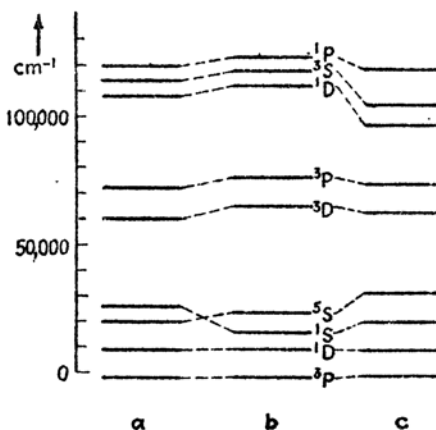


FIG. 4.—The relative positions of the levels of the $2s^2 2p^2$ and $2s 2p^3$ configurations of carbon in different approximations. *a*, strict one-electron approximation; *b*, approximation in which interaction between configurations is taken into account; *c*, observed term values. It should be noted that the positions of the 1S and 1D levels are inverted in going from *a* to *b*.

¹ TORRANCE, *op. cit.*

² C. W. UFFORD, *Phys. Rev.*, **53**, 568 (1938).

³ Modified forms of Fock's equations have been developed by G. H. Shortley, *Phys. Rev.*, **50**, 1072 (1936).

so that this perturbation does not account for a very large part of the error in cohesive energy.

c. Oxygen.—Hartree and Black¹ have solved Hartree's equation for oxygen in various states of ionization, using the approximation described above, in which only the radial part of the field arising from *p* functions is employed. For this reason and because of the fact that Hartree's rather than Fock's equations were solved, the results are not so significant as in the case of beryllium and carbon. One point that may give us some confidence in them, however, is the fact that the energies derived from Hartree's and Fock's equations do not differ appreciably in the case of beryllium. The energies of several states of O^{++} , O^+ , and O , relative to the ground state of the atom having one less electron, are listed in Table XLVIII.

TABLE XLVIII

		Calculated, Rydberg units	Observed, Rydberg units	Difference
O^{++}	1P	3.976	4.050	0.074
	1D	3.778	3.868	0.090
	1S	3.482	3.659	0.176
O^+	4S	2.516	2.602	0.086
	3D	2.258	2.334	0.076
	3P	2.084	2.210	0.126
O	3P	0.832	1.00	0.168
	1D	0.686	0.856	0.170
	1S	0.468	0.694	0.226

If the equations that were solved had been Fock's and if they had been solved exactly, we might regard the energy difference on the right of Table XLVIII as the correlation energy of the electron that is removed to give the ground state. Actually, these values are only approximately equal to the correlation energies. It should be noted that the correlation energy is greater for the excited states, particularly for those having low values of multiplicity and angular momentum. The dependence on multiplicity is probably related to the fact that states with low multiplicity are least affected by the Pauli principle so that the effects of "accidental correlations" are not so prominent as in the other cases; hence, other correlations that are not provided by the Pauli principle are more important.

d. Other Cases.—A large number of other atoms have been investigated by Hartree and additional workers. Although the Hartree

¹ D. R. HARTREE and M. M. BLACK, *Proc. Roy. Soc.*, **139**, 311 (1933).

potentials arising from inner-shell electrons prove to be very valuable for the process of determining wave functions of valence electrons in solids, we shall not discuss these cases here, since they would carry us too far afield. We shall refer to some of the results in later sections, however, and so we shall list those atoms for which Hartree fields have been obtained.

Ag ⁺	M. M. Black, <i>Mem. and Proc. Manchester Lit. Phil. Soc.</i> (1934-1935).
Al ³⁺	D. R. Hartree, <i>Proc. Roy. Soc.</i> , 151 , 96 (1935).
A	D. R. Hartree and W. Hartree, <i>Proc. Roy. Soc.</i> , 166 , 450 (1938).
B	F. W. Brown, J. H. Bartlett, and C. G. Dunn, <i>Phys. Rev.</i> , 44 , 296 (1933).
Be, Be ⁺	D. R. Hartree and W. Hartree, <i>Proc. Roy. Soc.</i> , 150 , 9 (1935); 154 , 588 (1936).
Ca	D. R. Hartree and W. Hartree, <i>Proc. Roy. Soc.</i> , 149 , 210 (1935); 164 , 167 (1938).
Cs ⁺	D. R. Hartree, <i>Proc. Roy. Soc.</i> , 143 , 506 (1934).
Cl ⁻	D. R. Hartree, <i>Proc. Roy. Soc.</i> , 141 , 281 (1933); 156 , 45 (1936).
Cu ⁺	D. R. Hartree, <i>Proc. Roy. Soc.</i> , 141 , 281 (1933); 157 , 490 (1936).
F, F ⁻	D. R. Hartree, <i>Proc. Roy. Soc.</i> , 151 , 96 (1935); F. W. Brown, <i>Phys. Rev.</i> , 44 , 214 (1933).
He	D. R. Hartree, <i>Cambridge Phil. Soc.</i> , 24 , 89 and 111 (1928).
Hg	D. R. Hartree and W. Hartree, <i>Proc. Roy. Soc.</i> , 149 , 210 (1935).
K	D. R. Hartree, <i>Proc. Roy. Soc.</i> , 143 , 506 (1934), 166 , 450 (1938); <i>Proc. Cambridge Phil. Soc.</i> , 34 , 550 (1938).
Li	J. Hargreaves, <i>Proc. Cambridge Phil. Soc.</i> , 25 , 75 (1928).
Na	V. Fock and Mary Petrashen, <i>Physik. Z.</i> , 6 , 368 (1934); D. R. Hartree and W. Hartree, <i>Proc. Cambridge Phil. Soc.</i> , 34 , 550 (1938).
Ne	F. W. Brown, <i>Phys. Rev.</i> , 44 , 214 (1933).
O, O ⁺ , O ⁺⁺ , O ³⁺	D. R. Hartree and M. M. Black, <i>Proc. Roy. Soc.</i> , 139 , 311 (1933).
Rb ⁺	D. R. Hartree, <i>Proc. Roy. Soc.</i> , 151 , 96 (1935).
Si ⁴⁺	J. McDougall, <i>Proc. Roy. Soc.</i> , 138 , 550 (1932).
W	M. F. Manning and J. Millman, <i>Phys. Rev.</i> , 49 , 848 (1936).

53. Types of Solution of Fock's Equations for Multiatomic Systems.—

Two independent types¹ of solution of Fock's equations have been widely used in multiatomic systems, namely: the Heitler-London, or atomic, type; and the Hund-Mulliken-Bloch, or molecular, type. In the Heitler-London scheme, it is assumed that the ψ_i are large only about single atoms or ions. Thus, in H₂, the two one-electron functions have the form shown roughly in Fig. 5a. This type of solution is accurate when the atoms of the multiatomic system are far from one another and the atomic or ionic properties of the constituent atoms are pronounced. In the Hund-Mulliken-Bloch scheme, on the other hand, it is

¹ The workers after which these schemes are named actually did not use them in connection with Fock's equations but simply used one-electron functions of the corresponding form. In the case of solids, we shall call the second scheme simply the "Bloch scheme" or the "band scheme."

assumed that each ψ extends over the entire system of atoms and has equal amplitude at equivalent atoms. In H_2 , for example, both electrons have the type of wave function that is illustrated in Fig. 5b. The use of this type of function is equivalent to assuming that the atoms of the system are affected by combination to such an extent that the

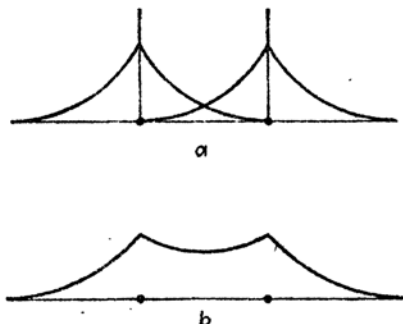


FIG. 5.—In the Heitler-London approximation the electrons of H_2 have separate wave functions of the type shown in *a*. In the Hund-Mulliken approximation both have the function *b*, which is distributed symmetrically between both atoms.

valence electrons belong to the entire molecule rather than to a single atom.

It has not been shown whether or not both types of solution always exist and whether or not they are the only solutions of Fock's equations. We shall show, however, that both types can exist.

Let us consider Fock's equations in the form

$$-\frac{\hbar^2}{2m}\Delta\psi_i(\mathbf{r}_1) + \left[V_i(\mathbf{r}_1) + \sum_j' e^2 \int \frac{|\psi_j|^2}{r_{12}} d\tau_2 \right] \psi_i(\mathbf{r}_1) - \sum_j' \left(e^2 \int \frac{\psi_j^* \psi_i d\tau_2}{r_{12}} \right) \psi_j(\mathbf{r}_1) = \epsilon \psi_i(\mathbf{r}_1). \quad (1)$$

We shall discuss the case of atomic functions first. The coulomb terms

$$\sum_j' e^2 \int \frac{|\psi_j|^2}{r_{12}} d\tau_2$$

obtained from Heitler-London functions screen part of the contribution to the ionic potential V from all atoms except the i th and make the attractive coulomb field largest at the i th atom. Thus, if we neglect the exchange term, we may expect the solutions of Eqs. (1) which are obtained when the potentials have been derived from atomic functions to be localized about individual atoms, that is, to have the form of atomic functions. Hence, we may expect to find a self-consistent solution of atomic functions in Hartree's case. This conclusion remains

valid when the exchange terms are included if the atoms are far apart, for then these terms are small. Since it has not yet proved possible to discuss the influence of exchange terms in any general way when the atoms are close together, it has not been demonstrated that atomic types of solution always occur in this case.

It is easy to see that the molecular types of solution always exist, for if we write Fock's equations in the form of Eq. (11), Sec. 61, it follows that no atom is preferred as long as the ψ_i prefer no atom. Hence, if a starting set of functions of the molecular type is used in applying a self-consistent scheme, all subsequent solutions, including the self-consistent one, will be of this type.

We shall show in later chapters (*cf.* Chaps. VII and VIII) that the two types of solution lead to identical antisymmetric total wave functions in an important case, namely, that of the normal state of a molecule the constituents of which have rare gas structure.

The molecular types of wave function, particularly those of lowest energy which have few nodes, are smoother than the atomic wave functions. Hence, we may expect that the mean kinetic energy usually will be lower when molecular functions are used. This advantage of the Hund-Mulliken-Bloch scheme is balanced by the fact that the scheme relies upon the accidental correlations introduced by the Pauli principle to reduce the energy of electron repulsion. The Heitler-London scheme, on the other hand, reduces this energy by keeping the electrons on separate atoms. The results for the problem of molecular hydrogen, which we shall discuss in the next chapter, show that in this case the advantages and disadvantages of the two schemes are about equal. Incidentally, the cohesive energy obtained by both schemes is in error by about 0.5 ev per electron, which shows that the solutions of Fock's equations are far from exact.

Both these one-electron schemes have been used extensively in solids, since each has its advantage for different types of problem. For example, the Bloch scheme is preferable in discussing metallic conductivity, whereas the Heitler-London scheme is preferable in discussing cohesion in ionic crystals. We shall develop both approximations in the following chapters, letting physical reasonableness be our guide in their application.

CHAPTER VII

MOLECULAR BINDING

54. Introduction.—There is an intimate connection between the binding properties of molecules and of solids as far as quantum mechanical principles are concerned. For this reason, we shall discuss some of the features of molecular binding. Since there are many important topics in the theory of molecules upon which we shall not touch, the following discussion should not be regarded as complete.

A fairly complete investigation along exact lines has been carried out for many of the simpler molecules such as H_2^+ , H_2 , and Li_2 . The Hartree-Fock scheme, discussed in the last chapter, plays a very minor role in this work, for the variational scheme has been employed directly. However, we shall be able to interpret some of the results in terms of the Hartree-Fock scheme. In cases such as H_2^+ and H_2 , in which the final results are almost as accurate as those obtained by Hylleraas for atomic helium, it is possible to gain an abundance of valuable information.

In addition to these quantitative investigations, there have been a number of qualitative discussions of more complicated molecules based on the Heitler-London and the Hund-Mulliken schemes. This work has proved to be extremely useful in the hands of the physical chemist who is willing to introduce an ample amount of empirical knowledge into any scheme he uses.

The Hamiltonian operator used in discussing the electronic structure of simple molecules is generally the same as the operator (1) of Sec. 47, Chap. VI, in which nuclear coordinates appear parametrically in V_i and I and in which spin interactions are neglected. For this reason, we may take over all the general remarks of the last chapter.

55. The Hydrogen-molecule Ion.—The simple molecule H_2^+ has the Hamiltonian operator

$$-\frac{\hbar^2}{2m}\Delta - \left(\frac{e^2}{r_a} + \frac{e^2}{r_b} - \frac{e^2}{r_{ab}} \right) \quad (1)$$

where r_a and r_b are the distances of the electron from the two protons, which are separated by a distance r_{ab} . The corresponding Schrödinger equation

$$-\frac{\hbar^2}{2m}\Delta\psi - \left(\frac{e^2}{r_a} + \frac{e^2}{r_b} - \frac{e^2}{r_{ab}} \right)\psi = E\psi \quad (2)$$

is separable when expressed in terms of the elliptical variables

$$\xi = \frac{r_a + r_b}{r_{ab}}, \quad \eta = \frac{r_a - r_b}{r_{ab}}, \quad \phi, \quad (3)$$

where ϕ is the angle that the plane containing the electron and the two nuclei makes with a fixed plane passing through the nuclei. In fact, if we set

$$\psi = \Xi(\xi)H(\eta)\Phi(\phi), \quad (4)$$

the separated equations are

$$\frac{d}{d\xi} \left((\xi^2 - 1) \frac{d\Xi}{d\xi} \right) + \left(-\lambda\xi^2 + 2D\xi - \frac{\mu^2}{\xi^2 - 1} + \tau \right) \Xi = 0, \quad (5a)$$

$$\frac{d}{d\eta} \left((1 - \eta^2) \frac{dH}{d\eta} \right) + \left(\lambda\eta^2 - \frac{\mu^2}{1 - \eta^2} - \tau \right) H = 0, \quad (5b)$$

$$\frac{d^2\Phi}{d\phi^2} = -\mu^2\Phi, \quad (5c)$$

where

$$\lambda = \frac{-m r_{ab}^2 \left(E - \frac{e^2}{r_{ab}} \right)}{2\hbar^2}, \quad D = \frac{r_{ab} m e^2}{\hbar^2}, \quad (5d)$$

and μ and τ are separational parameters.

From (5c), it is clear that $\Phi = e^{i\mu\phi}$, whence μ may take on only integer values. In addition, we know that the two states for which $\mu = \pm\mu'$ will have identical energies, for only μ^2 appears in the other two equations. In other words, all levels except those for which μ is zero are two-fold degenerate. Since the angular momentum about the axis joining the nuclei is simply $\mu\hbar$, it is conventional to designate the states for which $\mu = 0, 1, 2, \dots$ by σ, π, δ , respectively, in analogy with the atomic designation involving the angular-momentum quantum number.

On general grounds, we should expect the lowest state to be one for which $\mu = 0$ since this type of function does not have an angular nodal plane. This state was investigated first by Burrau,¹ but a more accurate treatment has been given by Teller² who, in addition, carried through an investigation of higher states. We shall not discuss Teller's work in detail, except to say that he solved (5b) by means of power series and found the eigenfunctions of (5a) by use of the variational equation going with this self-adjoint differential equation. If the number of nodes in Ξ and H is designated by n_ξ and n_η , respectively, it should be possible to label all the states by these two quantum numbers and μ . The

¹ O. BURRAU, *Danske Videnskab. Selskab*, **7**, 14 (1927).

² E. TELLER, *Z. Physik*, **61**, 458 (1930).

convention usually adopted, however, is to employ the quantum numbers n and l of the hydrogen-like state into which a given molecular eigenfunction degenerates when r_{ab} approaches zero. A simple investigation reveals that n_ξ goes into the radial quantum number and that n_η goes into the polar quantum number $l - \mu$, as may be seen from the fact that ξ becomes a radial variable and η becomes the variable $\cos \theta$, where θ is the polar angle in spherical polar coordinates. In other words

$$\left. \begin{aligned} n &= n_\xi + l + 1; \\ l &= n_\eta + \mu. \end{aligned} \right\} \quad (6)$$

The energies of several states are shown as functions of r_{ab} in Figs. 1 and 2. In Fig. 1, the abscissa is the internuclear distance expressed in Bohr units and the ordinate is the purely electronic part of the energy

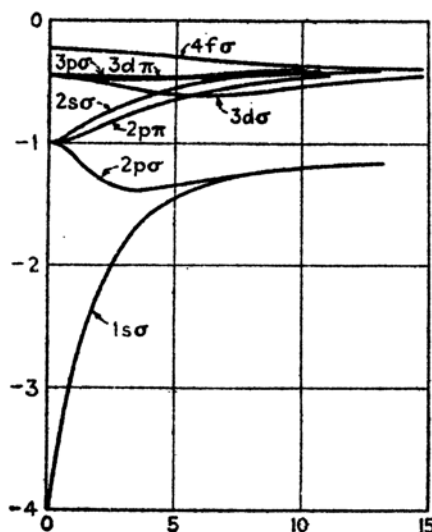


FIG. 1.—The electronic energy of several states of H_2^+ as a function of internuclear distance. (After Teller.) The ordinate is in Rydberg units.

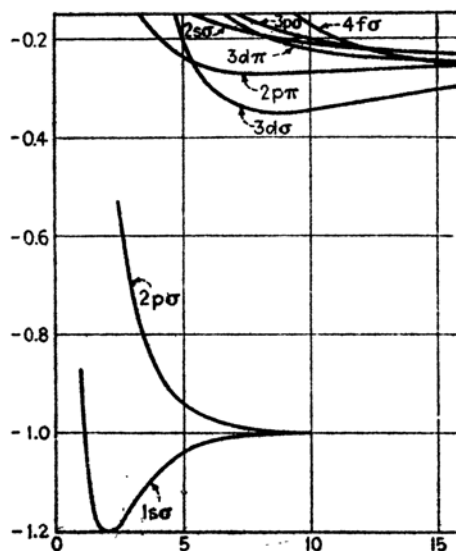


FIG. 2.—The total energy of several states of H_2^+ as a function of internuclear distance. These curves are derived from those of Fig. 3 by adding the energy of nuclear repulsion. (After Teller.)

(that is, the nuclear repulsion is neglected). In Fig. 2, the abscissae are the same as in Fig. 1, whereas the ordinates represent the total energy. The stable states are those of Fig. 2 which possess a minimum for finite values of r_{ab} . It may be seen that the only members of the computed set that possess this property are the $1s\sigma$, $3d\sigma$ and $2p\pi$ states. The energies of the minima relative to a zero in which all particles are separated from each other by an infinite distance are given in Table XLIX.

TABLE XLIX

State	Nuclear separation, a_h	Energy, Rydberg units	Energy of dissociation, ev
1s σ	2.00	-1.20537	2.781
3d σ	11.5	-0.350	1.35
2p π	8	-0.265	0.20

The last column in this table is the energy required to dissociate the molecule into an atom and a proton.

When the parameter r_{ab} becomes very large, the amplitudes of the wave functions become negligibly small at distances midway between the nuclei. The wave function then reduces, for all practical purposes, to two hydrogen wave functions, each centered about one of the protons. Only one of these two states should be regarded as the final state if the separation is sufficiently large, for there is only a negligible probability of the electron jumping from one atom to another. The quantum numbers of this atomic state, which may be determined by a simple analysis involving parabolic coordinates, are connected with those of the molecule in the following way: The total quantum number n' of the final state is related to n_ξ , n_η , and μ by the equations

$$n' = \begin{cases} n_\xi + \frac{n_\eta}{2} + \mu + 1 & (n_\eta \text{ even}) \\ n_\xi + \frac{n_\eta + 1}{2} + \mu + 1 & (n_\eta \text{ odd}). \end{cases} \quad (7)$$

No definite l value may be assigned to the final state, for the hydrogen-atom wave functions obtained by removing a proton from H_2^+ are not eigenfunctions of angular momentum. The values of l entering into these wave functions, when they are expressed as a linear combination of eigenfunctions of angular momentum, range over the allowed values that may be associated with n' , that is, from $n' - 1$ to 0. The value of m for the final state is the same as the value of μ before separation. It may be seen from (6) that the value of n obtained by coalescing the two nuclei is either greater than or equal to the value (7) obtained by separation.

The electronic charge distribution of the lowest state is shown in Fig. 3, for equilibrium separation of protons. It should be noted that the amplitude is large and nearly constant at distances midway between the protons. The disadvantage of having two repelling protons is more

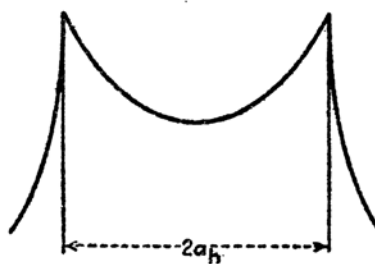


FIG. 3.—The charge distribution associated with the normal wave function of H_2^+ . (After Burrau.)

than compensated by the following two facts: (1) There is a larger region in which the electron may have negative potential energy when there are two protons instead of one. (2) The wave function may be smooth over a larger region of space. The first fact has associated with it a drop in potential energy; the second, a drop in kinetic energy. The $2p\sigma$ state, which combines with the $1s\sigma$ at large internuclear distances, has a nodal plane midway between protons. Thus, neither is it smooth, nor does it allow the electron to have an appreciable probability of being at midway regions where the potential field favors binding. Figure 2 shows that it is entirely repulsive, as we might expect.

56. The Hydrogen Molecule.—The hydrogen molecule has been treated approximately by a large number of workers. We shall consider first the solution obtained by James and Coolidge,¹ since it is considerably more accurate than any other. Their procedure is patterned after the one Hylleraas followed when working with helium (*cf.* Sec. 48, Chap. VI). We shall see that there is a close correspondence between the conclusions that may be drawn from the solutions of both problems.

a. James and Coolidge expressed the Hamiltonian operator

$$-\frac{\hbar^2}{2m}(\Delta_1 + \Delta_2) - \frac{e^2}{r_{1a}} - \frac{e^2}{r_{1b}} - \frac{e^2}{r_{2a}} - \frac{e^2}{r_{2b}} + \frac{e^2}{r_{12}} + \frac{e^2}{r_{ab}} \quad (1)$$

in terms of the four elliptical variables

$$\begin{aligned} \xi_1 &= \frac{r_{1a} + r_{1b}}{r_{ab}}, & \xi_2 &= \frac{r_{2a} + r_{2b}}{r_{ab}}, \\ \eta_1 &= \frac{r_{1a} - r_{1b}}{r_{ab}}, & \eta_2 &= \frac{r_{2a} - r_{2b}}{r_{ab}}, \end{aligned} \quad (2a)$$

which are analogous to the set (3) of the previous section. Instead of using ϕ_1 and ϕ_2 as the remaining pair of electron variables, they chose the set.

$$\theta = \phi_1 + \phi_2 \quad \text{and} \quad \rho = \frac{2r_{12}}{r_{ab}}. \quad (2b)$$

The interelectronic distance was used explicitly in order that electronic correlation, discussed in the last chapter, might play a role appropriate to its importance. The variable θ does not enter into the wave function of the lowest state for the same reason that ϕ does not enter into the wave function of the lowest state of H_2^+ .

The starting wave function for the lowest state was taken in the power-series form

$$e^{-\lambda(\xi_1 + \xi_2)} \sum C_{mnjkp} (\xi_1^m \xi_2^n \eta_1^j \eta_2^k \rho^p + \xi_1^m \xi_2^n \eta_1^k \eta_2^j \rho^p) \quad (3)$$

¹ H. M. JAMES and A. S. COOLIDGE, *Jour. Chem. Phys.*, **1**, 825, (1933).

where δ and the C are parameters that must be determined. Naturally, only the first few terms of the series were used. The various steps that were taken in extending the computation may be summarized in the following way.

1. Only the exponential term was retained in the first approximation. The best value of δ depends upon r_{ab} and is 1.696 for the observed separation of 1.40 Bohr units. In this case, the binding energy of the molecule is 2.56 ev, which should be compared with the experimental value of 4.73 ± 0.04 ev.

The wave function

$$e^{-\delta(\xi_1 + \xi_2)}$$

may be written in the form

$$e^{-\frac{\delta(r_{1a} + r_{1b})}{r_{ab}}} e^{-\frac{\delta(r_{2a} + r_{2b})}{r_{ab}}}, \quad (4)$$

which corresponds to the product of two one-electron wave functions of the Hund-Mulliken type. Hence, the results for this case give us a lower limit to the accuracy that should be expected if the Fock-Hartree procedure based on the Hund-Mulliken scheme were employed. From this result, it is difficult to say to what extent a rigorous solution of Fock's equations would improve the calculated energy. A more accurate solution of the Hund-Mulliken type will be discussed in part b.

2. Neglecting all terms in (3) that depend upon ρ , James and Coolidge found that the best energy they could obtain for $r_{ab} = 1.40a_h$ is about 4.27 ev, which differs from the observed value by approximately $\frac{1}{2}$ ev. Since this approximation is the best possible one in which correlations are not included, it gives an upper limit to the possible accuracy of the Fock-Hartree procedure. It is doubtful whether the energy of the Fock-Hartree approximation would be nearly as good as this, however, since (3) is much more intricate than a product of one-electron functions even when terms in ρ are neglected. Thus, the correlation energy correction would be at least $\frac{1}{2}$ ev per electron, if the Hund-Mulliken scheme were used.

3. The simplest function employed in which the variable ρ occurs is the five-parameter expression

$$e^{-\delta(\xi_1 + \xi_2)}[a_1(\eta_1^2 + \eta_2^2) + a_2\eta_1\eta_2 + a_3(\xi_1 + \xi_2) + a_4\rho]. \quad (5)$$

This leads to an energy of 4.507 ev with $r_{ab} = 1.40a_h$. The parameter values are listed in Table L. It should be noted that δ has a value considerably different from that discussed in case 1. Thus, just as in the case of helium, the inclusion of a linear term in ρ leads to a better energy.

4. As a final step, a thirteen-parameter expression involving quadratic terms in ξ_1 , ξ_2 , and ρ , as well as those appearing in (5), was employed.

The parameters C_{mnkp} are tabulated in Table L, and it may be seen that in some cases they differ greatly from the values discussed in case 3. The final energy is about 4.69 ev, which lies within the experimental error of the observed value. The computed internuclear distance is identical with the observed one.

TABLE L

	Case 1	Case 3	Case 4
δ	1.69609	2.23779	2.2350
C_{00020}	0.80483	1.19279
C_{00110}	-0.27997	-0.45805
C_{10000}	-0.60985	-0.82767
C_{10200}	-0.17134
C_{10020}	-0.12101
C_{10110}	0.12394
C_{20000}	0.08323
C_{00001}	0.19917	0.35076
C_{00021}	0.07090
C_{00111}	-0.01143
C_{10001}	-0.03987
C_{00002}	-0.01197

b. We shall now compare the accuracy of the Heitler-London and Hund-Mulliken schemes in so far as computations on H_2 allow us to make a comparison. It is easy to show that Fock's equations for the two schemes are essentially different in this case, as they are in most multiatomic systems (an exceptional case will be discussed in Sec. 58).

For a two-electron system, the singlet eigenfunction based upon one-electron functions is

$$\frac{1}{\sqrt{2}}[\psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2) + \psi_1(\mathbf{r}_2)\psi_2(\mathbf{r}_1)][\eta_1(1)\eta_2(-1) - \eta_2(1)\eta_1(-1)]. \quad (6)$$

Here, ψ_1 and ψ_2 are the one-electron functions that, in the Heitler-London scheme, are centered about different nuclei. Fock's equation for ψ_1 is

$$-\frac{\hbar^2}{2m}\Delta_1\psi_1(\mathbf{r}_1) + \left(-\frac{e^2}{r_{1a}} - \frac{e^2}{r_{1b}} + e^2 \int \frac{|\psi_2(\mathbf{r}_2)|^2}{r_{12}} d\tau_2\right)\psi_1(\mathbf{r}_1) + \left(e^2 \int \frac{\psi_1(\mathbf{r}_2)\psi_2(\mathbf{r}_2)}{r_{12}} d\tau_2\right)\psi_2(\mathbf{r}_1) = e\psi_1(\mathbf{r}_1) + \lambda\psi_2(\mathbf{r}_1), \quad (7)$$

and there is a similar equation for ψ_2 .

The space part of the total wave function (6) reduces to $\psi(\mathbf{r}_1)\psi(\mathbf{r}_2)$ when ψ_1 and ψ_2 are equal, as in the Hund-Mulliken scheme. In this case, Fock's equation for $\psi(\mathbf{r})$ is

$$-\frac{\hbar^2}{2m}\Delta\psi + \left(-\frac{e^2}{r_{1a}} - \frac{e^2}{r_{2a}} + e^2 \int \frac{|\psi(\mathbf{r}_2)|^2}{r_{12}} d\tau_2\right)\psi = \epsilon\psi. \quad (8)$$

Equation (7) reduces not to (8) when we set $\psi_1 = \psi_2$, but instead to

$$-\frac{\hbar^2}{2m}\Delta\psi + \left(-\frac{e^2}{r_{1a}} - \frac{e^2}{r_{2a}} + 2e^2 \int \frac{|\psi(\mathbf{r}_2)|^2}{r_{12}} d\tau_2\right)\psi = \epsilon\psi, \quad (9)$$

in which the coulomb integral contains a factor 2. Thus the different assumptions of the two cases lead to different systems of equations. The two schemes should be regarded as different approximate solutions of the same problem rather than as different solutions of the same set of one-electron equations.

The most accurate attempt to apply the Heitler-London scheme to the hydrogen molecule was made by Wang,¹ who used hydrogen-like atomic wave functions of the form $e^{-\alpha r}$, where r is the distance of an electron from the nucleus and α is an adjustable screening parameter. The best value of the cohesive energy obtained by this procedure is 3.76 ev corresponding to $r_{ab} = 1.41a_h$. The effective proton charge Ze is related to α by the equation

$$Ze = e\alpha a_h.$$

For the best value of α , one finds that $Z = 1.17$. At first sight, it may seem surprising that this is greater than unity. It should be recalled, however, that Z must increase from unity to the value of $\frac{2}{3}$ for atomic helium (cf. Sec. 48) as r_{ab} decreases from infinity to zero.

On the other hand, the most accurate attempt to apply the Hund-Mulliken scheme to the hydrogen molecule was made by Hylleraas.² He constructed a determinantal eigenfunction from one-electron solutions of a two-center system, similar to H_2^+ . He assumed that the charge on the centers was $e/2$, rather than e as in H_2^+ , in order to compensate for the nuclear screening effect that one electron exerts on the other. The cohesive energy obtained by taking the mean value of the Hamiltonian of H_2 with this determinantal eigenfunction is 3.6 ev for an internuclear distance of $1.40a_h$. This fact shows that Fock's equation would lead to the correct energy to within at least 0.55 ev per electron if the Hund-Mulliken scheme were used. In view of the remarks made under 2, part a, we can say that the correlation energy of H_2 for the Hund-Mulliken scheme lies between 0.55 and 0.25 ev per electron.

We may conclude³ from these two cases that the Heitler-London and Hund-Mulliken schemes are about equally successful as far as the

¹ S. C. WANG, *Phys. Rev.*, **31**, 579 (1928).

² E. HYLLERAAS, *Z. Physik*, **71**, 741 (1931).

³ Other methods of treating H_2 are surveyed by J. H. Van Vleck and A. Sherman, *Rev. Modern Phys.*, **7**, 167 (1935).

problem of minimizing the energy is concerned. The error in each case is of the order of 0.5 ev per electron.

57. Molecular Lithium.—Next to hydrogen, diatomic lithium is the simplest stable molecule that contains only one type of atom. We shall outline briefly the computations that deal with it.

In early work, several attempts were made to compute the binding energy of Li_2 by taking into account only the two 2s valence electrons and treating 1s shells as though rigidly fixed. The interactions between the closed shells on different atoms were neglected. It was first pointed out by James¹ that the relatively simple methods used in treating the two valence electrons give a binding energy that agrees with the experimental value of 1.14 ev only because the closed-shell interactions are neglected. He carried out several types of calculation that show the following facts.

a. If the closed-shell interactions are neglected, a binding energy greater than the observed value may be obtained by using essentially the same procedure that James and Coolidge employed for H_2 . There are two reasons for this fact: (1) The ion cores repel one another more strongly when the shells are taken into account than when they are not. The origin of this additional repulsion will be discussed in the next section. (2) The wave function obtained by a variational procedure usually violates the Pauli principle unless the closed-shell wave functions are explicitly included in the varied wave function.

b. A binding energy of 0.62 ev may be obtained by an involved variational computation in which closed-shell wave functions are included in the varied function. The interelectronic distance variables were not introduced into this wave function since they would have made the computations prohibitively complicated.

James's work on Li_2 is interesting from the standpoint of computations dealing with monatomic solids, for many features of the two cases are identical. Since closed inner shells are present in all the interesting solids, it is important to know the extent to which they can be neglected. The preceding discussion shows that the problem of closed shells must be approached with care if relative binding energies are to have much significance. It will become apparent later that there are several redeeming features in the case of solids. Most important among these is the fact that equilibrium distances in solids are usually much larger than in molecules. For example, the closest distance of approach of lithium atoms in the metal is $5.65a_k$, whereas it is $5.05a_k$ in the molecule.

58. Closed-shell Interaction and van der Waals Forces.²—There is one case in which the Heitler-London and Hund-Mulliken schemes are

¹ M. JAMES, *Jour. Chem. Phys.*, **2**, 794 (1934).

² For a review article by H. Margenau, *Rev. Modern Phys.*, **11**, 1 (1939), for a summary of the development of the theory of van der Waals forces.

identical, namely, the case of electron configurations that correspond to the interaction of closed shells. In this case, the singlet eigenfunction is a simple determinant, for all electronic wave functions appear in pairs with opposite spin. The determinant formed from Heitler-London wave functions may be rearranged so as to satisfy the Hund-Mulliken conditions. We shall demonstrate this theorem by considering the interaction of two normal helium atoms. It will be evident that the principles involved in this particular case are generally applicable.

We shall designate the wave functions centered about one of the atoms by ψ_a and those centered about the other by ψ_b . We may assume that ψ_a and ψ_b are symmetrical in the sense that they become interchanged if the two nuclei are interchanged (cf. Fig. 4). The wave function for the molecule is then

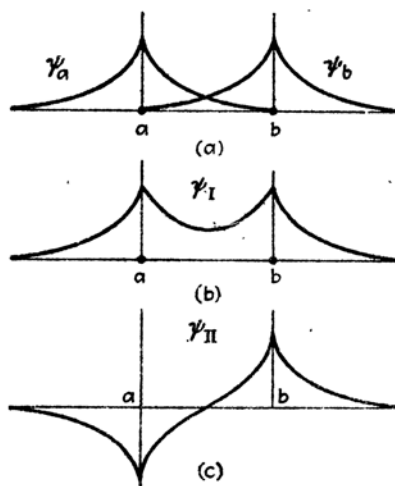


FIG. 4.—*a*, the two Heitler-London functions of H_2 . Their sum *b* and their difference *c* are Hund-Mulliken functions.

$$\begin{vmatrix} \psi_a(r_1)\eta_1(1) & \psi_a(r_2)\eta_2(1) & \psi_a(r_3)\eta_3(1) & \psi_a(r_4)\eta_4(1) \\ \psi_a(r_1)\eta_1(-1) & \psi_a(r_2)\eta_2(-1) & \psi_a(r_3)\eta_3(-1) & \psi_a(r_4)\eta_4(-1) \\ \psi_b(r_1)\eta_1(1) & \psi_b(r_2)\eta_2(1) & \psi_b(r_3)\eta_3(1) & \psi_b(r_4)\eta_4(1) \\ \psi_b(r_1)\eta_1(-1) & \psi_b(r_2)\eta_2(-1) & \psi_b(r_3)\eta_3(-1) & \psi_b(r_4)\eta_4(-1) \end{vmatrix}. \quad (1)$$

It should be noted that corresponding elements in the first and third rows have identical spin functions, as do those in the second and fourth rows. Let us now add the third row to the first, subtract the first row from the second, and repeat this procedure with the second and fourth rows. We then obtain the determinant

$$\begin{vmatrix} \psi_I(r_1)\eta_1(1) & \psi_I(r_2)\eta_2(1) & \psi_I(r_3)\eta_3(1) & \psi_I(r_4)\eta_4(1) \\ \psi_I(r_1)\eta_1(-1) & \psi_I(r_2)\eta_2(-1) & \psi_I(r_3)\eta_3(-1) & \psi_I(r_4)\eta_4(-1) \\ \psi_{II}(r_1)\eta_1(1) & \psi_{II}(r_2)\eta_2(1) & \psi_{II}(r_3)\eta_3(1) & \psi_{II}(r_4)\eta_4(1) \\ \psi_{II}(r_1)\eta_1(-1) & \psi_{II}(r_2)\eta_2(-1) & \psi_{II}(r_3)\eta_3(-1) & \psi_{II}(r_4)\eta_4(-1) \end{vmatrix} \quad (2)$$

where

$$\psi_I = \psi_a + \psi_b, \quad \psi_{II} = \psi_a - \psi_b. \quad (3)$$

Since (2) has been obtained by adding other rows to the rows of (1), the two determinants are equal. ψ_I and ψ_{II} , however, are Hund-Mulliken wave functions (cf. Fig. 4). Hence, in this case, the two schemes are equivalent. In more general cases, the determinant corresponding

to (1) has more rows and columns, but alternate rows have equal spin functions and may again be combined to form Hund-Mulliken wave functions. We shall discuss the crystalline case in Chap. VIII.

The Heitler-London scheme has been used to compute the energy of interaction between closed shells. For convenience, the one-electron functions employed in these computations were based upon approximate atomic functions rather than upon solutions of Fock's equations. We shall discuss several examples in detail.

a. Helium Interaction.—The most accurate treatment of the repulsive interaction of two helium atoms has been given by Slater¹ who computed the mean value of the interaction potential H_1 of the two molecules, using a total wave function that was constructed by taking an appropriate linear combination of atomic eigenfunctions. His work is not a strict application of the Heitler-London method, since the atomic wave functions that he used were not constructed of one-electron functions alone but contained the interelectronic distance variable. In analytical form, his atomic wave function $\Phi(1,2)$ is

$$\Phi(1,2) = \begin{cases} 1.392e^{-2(r_1+r_2)+0.5r_{12}+0.0107(r_1^2+r_2^2)}; & (r_1, r_2 < 3), \\ 1.241e^{-2r_2-1.244r_1r_2-0.255\left(1+\frac{0.0707}{r_1}\right)}; & (r_1 > 3; r_2 < 3), \\ 1.241e^{-2r_1-1.344r_2r_1-0.255\left(1+\frac{0.0707}{r_2}\right)}; & (r_2 > 3; r_1 < 3). \end{cases} \quad (4)$$

Here r_1 , r_2 , and r_{12} , are the radial distances between the electrons and the nucleus and between the electrons, expressed in units of the Bohr radius. If we designate the two nuclei by a and b and let a subscript on Φ indicate the nucleus about which the electrons are centered, the complete antisymmetric wave function for He_2 that is obtainable from (4) is

$$\Psi(1,2,3,4) = \Phi_a(1,2)\Phi_b(3,4) - \Phi_a(1,4)\Phi_b(2,3) - \Phi_a(2,3)\Phi_b(1,4) + \Phi_a(3,4)\Phi_b(1,2), \quad (5)$$

in which it is implied that each Φ has an appropriate spin factor. The function (5) reduces to (3) when $\Phi(1,2)$ is simplified to a product of one-electron functions.

The repulsive interaction $V(R)$ that is derived from (5) is a fairly complicated function of the internuclear distance R , but it may be accurately approximated by means of the analytical expression

$$V(R) = 481e^{-\frac{R}{0.412}}, \quad (6)$$

in units of ev, for R greater than $2a_h$.

¹ J. C. SLATER, *Phys. Rev.*, **32**, 339 (1928).

b. Neon Interaction.—Bleick and Mayer¹ have determined the interaction energy in the case of two neon atoms, using one-electron atomic functions that were obtained by approximating Brown's results² analytically. The radial parts of the 2s and 2p functions are

$$\left. \begin{aligned} R_{2s} &= 13.6e^{-3.22r} - r(14.7e^{-3.69r} + 4.76e^{-2.15r}), \\ R_{2p} &= r(17.9e^{-3.80r} + 2.30e^{-1.69r}). \end{aligned} \right\} \quad (7)$$

The internuclear distances used in their work were so large that the 1s functions did not overlap appreciably. The results of a straightforward, though laborious, computation of the interaction energy are given in Table LI for three values of the internuclear distance.

TABLE LI	
R , Bohr Units	$V(R)$, ev.
3.41	2.15
4.35	0.22
6.05	0.003

These three values may be fitted by the simple function

$$1.18 \cdot 10^4 e^{-\frac{R}{0.395}} \quad (8)$$

with an error of about 10 per cent.

The fact that $V(R)$ is approximately exponential in both the cases *a* and *b* provides a rough justification for the exponential term that Born and Mayer used to express the repulsive interaction of ions in crystals.

These repulsive terms do not describe the interaction energy properly at large distances since the tendency of electrons to avoid one another has not been taken into account. An additional attractive term is found when this is done. Since the additional term was implicitly postulated by van der Waals when he proposed his equation of state for gases, it is called the "van der Waals interaction."

Let us consider the van der Waals energy³ for two atoms *a* and *b* that have *m* and *n* electrons, respectively, and are separated by a distance *R*. We shall assume that the atoms lie along the *x* axis and shall designate the Cartesian coordinates of the electrons relative to the nucleus of the atom on which they reside by (x_{ai}, y_{ai}, z_{ai}) and (x_{bi}, y_{bi}, z_{bi}) , respectively, where *i* ranges from 1 to *m* and *j* ranges from 1 to *n*. The interaction potential H_i of the two atoms may be computed by straightforward principles of electrostatic theory and may be expanded in terms of the Cartesian variables. The result is

¹ W. E. BLEICK and J. MAYER, *Jour. Chem. Phys.*, **2**, 252 (1934).

² W. G. BROWN (*cf.* Sec. 52).

³ F. LONDON, *Z. Physik*, **63**, 245 (1930).

$$H_I = \sum_{i=1}^m \sum_{j=1}^n \left\{ -\frac{e^2}{R^3} (2x_{ai}x_{bj} - y_{ai}y_{bj} - z_{ai}z_{bj}) + \frac{3}{2} \frac{e^2}{R^4} [r_{ai}^2 x_{bj} - x_{ai} r_{bj}^2 + (2y_{ai}y_{bj} + 2z_{ai}z_{bj} - 3x_{ai}x_{bj})(x_{ai} - x_{bj})] + \frac{3}{4} \frac{e^2}{R^5} [r_{ai}^2 r_{bj}^2 - 5r_{ai}^2 x_{bj}^2 - 5r_{bj}^2 x_{ai}^2 - 15x_{ai}^2 x_{bj}^2 + 2(4x_{ai}x_{bj} + y_{ai}y_{bj} + z_{ai}z_{bj})^2] + \dots \right\}. \quad (9)$$

The terms in $1/R^3$, $1/R^4$ and $1/R^5$ are called, respectively, the dipole-dipole, dipole-quadrupole, and quadrupole-quadrupole interaction terms because they are similar to the interaction energy of the corresponding types of multipole.

The expressions (6) and (8) are mean values of H_I for the approximate wave functions constructed from atomic wave functions, and they may be looked upon as the first-order terms in the perturbation formula

$$E_I(R) = \int \Psi_0^* H_I \Psi_0 d\tau - \sum_a \frac{|\int \Psi_0^* H_I \Psi_a d\tau|^2}{E_0 - E_a} + \dots \quad (10)$$

where Ψ_0 is the lowest state and the Ψ_a are higher states. Hence, the second term is the van der Waals energy, if we may assume that higher terms in the expansion (10) are negligible.

London has derived a somewhat rough but general expression for the contribution to the van der Waals energy from the dipole-dipole potential in (9). For simplicity, he assumed that the unperturbed wave functions of the normal and excited states may be represented by products of functions of the separate atoms, thus,

$$\Psi = \Phi_\mu \Phi_\nu \quad (11)$$

in which μ and ν correspond to different atomic states. This approximation is poor when the atoms are very close together, for it violates the Pauli principle. It is accurate, however, for large atomic separations when the wave functions of different atoms do not overlap. On this assumption the integrals in the numerator of the second term in (10) degenerate into sums of products of integrals over separate atoms, for each of the terms in (9) is a product of terms involving variables of electrons on different atoms. These one-atom integrals, which have the form

$$\int \Phi_a^* \left(\sum_i x_{ai} \right) \Phi_a d\tau_a, \quad (12)$$

play an important role in the theory of optical radiation (cf. Chap. V). They vanish unless a dipole transition is allowed between the states Φ_a^0

and Φ_a' . Moreover, there is a close connection between the integrals which involve $\sum_i x_{ai}$ and those which involve $\sum_i y_{ai}$ and $\sum_i z_{ai}$, for the Φ are wave functions for spherically symmetrical atoms. For these reasons, the numerator of the second term in (10) may be simplified considerably, and the sum may be reduced to

$$E_{d-d}(R) = 6 \frac{e^4}{R^6} \sum_{\mu\nu} \frac{\left| \left(\sum_i z_{ai} \right)_{0\mu} \right|^2 \left| \left(\sum_j z_{bj} \right)_{0\nu} \right|^2}{E_a^0 + E_b^0 - E_a^\mu - E_b^\nu} \quad (13)$$

where

$$\left. \begin{aligned} \left(\sum_i z_{ai} \right)_{0\mu} &= \int \Phi_a^{0*} \left(\sum_i z_{ai} \right) \Phi_a^\mu d\tau, \\ \left(\sum_j z_{bj} \right)_{0\nu} &= \int \Phi_b^{0*} \left(\sum_j z_{bj} \right) \Phi_b^\nu d\tau, \end{aligned} \right\} \quad (14)$$

and the E are the energy values of the wave functions that appear in these integrals.

E_a^μ and E_b^ν are sometimes replaced by constant mean values \bar{E}_a and \bar{E}_b in order that the numerator may be summed alone. These mean values, which are not defined independently of Eq. (13), are usually assumed to be approximately equal to the ionization energy of the atoms. By the use of this approximation, Eq. (13) becomes

$$E_{d-d}(R) = 6 \frac{e^4}{R^6} \frac{\left[\left(\sum_i z_{ai} \right)^2 \right]_{00} \left[\left(\sum_j z_{bj} \right)^2 \right]_{00}}{E_a^0 + E_b^0 - \bar{E}_a - \bar{E}_b}, \quad (15)$$

since

$$\sum_\mu \left| \left(\sum_i z_{ai} \right)_{0\mu} \right|^2 = \left[\left(\sum_i z_{ai} \right)^2 \right]_{00}. \quad (16)$$

The polarizability α of an atom is related to its energy $E(E)$, in a field of intensity E , by means of the equation

$$E(E) = E_0 + \frac{1}{2} \alpha E^2. \quad (17)$$

Thus, according to perturbation theory, we have

$$\alpha = -2e^2 \sum_\mu \frac{\left| \left(\sum_i z_{ai} \right)_{0\mu} \right|^2}{E_0 - E_\mu}, \quad (18)$$

which is approximately equal to

$$-2e^2 \frac{\left[\left(\sum_i z_{ai} \right)^2 \right]_{00}}{E_0 - \bar{E}_a}.$$

Hence, (15) is approximately equal to

$$E_{d-d} \cong -\frac{3}{2} \frac{1}{R^6} \frac{(E_0 - \bar{E}_a)(E_0 - \bar{E}_b)}{(E_a^0 + E_b^0 - \bar{E}_a - \bar{E}_b)} \alpha_a \alpha_b \quad (19)$$

where α_a and α_b are, respectively, the polarizabilities of atoms a and b .

Margenau¹ and Mayer² have applied similar methods in order to obtain the contribution to the van der Waals energy from the dipole-quadrupole and quadrupole-quadrupole terms. The end result for the dipole-quadrupole term that corresponds to (19) is

$$E_{d-q} = -\frac{3}{2} \frac{1}{R^8} \frac{[\alpha_a \alpha_b^2 (E_a^0 - \bar{E}_a)(E_b^0 - \bar{E}_b)^2 + \alpha_a^2 \alpha_b (E_a^0 - \bar{E}_a)^2 (E_b^0 - \bar{E}_b)]}{(E_a^0 + E_b^0 - \bar{E}_a - \bar{E}_b)} \quad (20)$$

The quadrupole-quadrupole term may be developed in a similar way, as is described in Margenau's paper.

We shall now discuss the results of computations for hydrogen and helium.

a. Hydrogen.—For two hydrogen atoms, (19) becomes

$$-\frac{6}{R^6} \frac{e^2}{a_h} \quad (21)$$

in which R is expressed in units of a_h , when $\bar{E} - E_0$ is set equal to the ionization energy $e^2/2a_h$. In the same approximation, the complete expression for the van der Waals energy, through quadrupole-quadrupole terms, is

$$-6 \frac{e^2}{a_h} \left(\frac{1}{R^6} + \frac{22.5}{R^8} + \frac{236}{R^{10}} \right). \quad (22)$$

This may be used to estimate the relative magnitude of the different terms.

More accurate expressions for the dipole-dipole term have been derived by Eisenschitz and London³ and by Slater and Kirkwood.⁴ The first workers summed (13) directly and obtained

$$-\frac{6.47}{R^6} \frac{e^2}{a_h} \quad (23)$$

whereas the second employed a variational method and found

$$-\frac{6.49}{R^6} \frac{e^2}{a_h} \quad (24)$$

¹ H. MARGENAU, *Phys. Rev.*, **38**, 747 (1931); **40**, 237 (1932); see also *op. cit.*

² J. E. MAYER, *Jour. Chem. Phys.*, **1**, 270 (1933).

³ R. EISENSCHITZ and F. LONDON, *Z. Physik*, **60**, 491 (1930).

⁴ J. C. SLATER and J. G. KIRKWOOD, *Phys. Rev.*, **37**, 682 (1931).

b. *Helium*.—Margenau¹ has derived an expression equivalent to (22) for helium. The result is

$$-\frac{1.62}{R^6} \frac{e^2}{a_h} \left(\frac{1}{R^6} + \frac{7.9}{R^8} + \frac{30}{R^{10}} \right). \quad (25)$$

A more exact expression for the dipole-dipole term, which was derived by Slater and Kirkwood, is²

$$-\frac{1.59}{R^6} \frac{e^2}{a_h}. \quad (26)$$

The sum of (26) and (6) has a shallow minimum with a depth of 0.75×10^{-3} ev at $5.5a_h$. This minimum accounts for the weak cohesive energy of liquid helium (see Fig. 5).

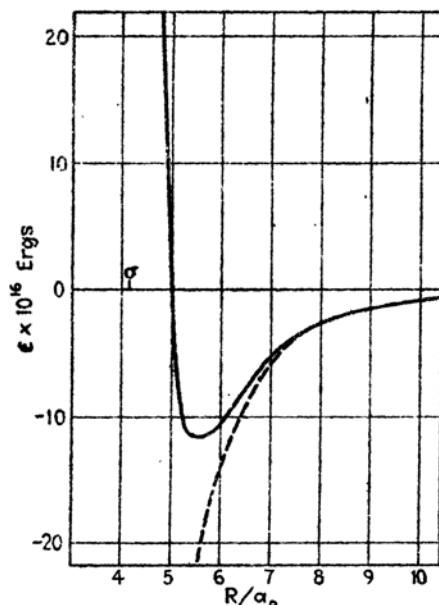


FIG. 5.—The total interaction energy of two helium atoms. The dotted curve is the van der Waals interaction. (After Slater and Kirkwood.)

A similar minimum should be expected for all rare gas atoms. Bleick and Mayer find that the total interaction energy for neon, obtained by adding (8) and (19), has a minimum of $1.3 \cdot 10^{-3}$ ev when R is equal to $6.5a_h$. These values agree roughly with those obtained from empirical data by Lennard-Jones.³

¹ Margenau has recently developed a revised form of Eq. (25) which he regards as a more accurate representation of the true van der Waals energy than either Eq. (25) or (26) (see *Phys. Rev.*, **55**, 1137 (1939)).

² SLATER and KIRKWOOD, *op. cit.*

³ See survey by J. E. Lennard-Jones, *Physica*, **4**, 941 (1937).

59. Molecular Valence.¹—The problem of providing a reason for the inertness of the rare gas atoms in their interaction with one another is solved by the computations discussed in the last section, for they show that closed-shell structures have very weak attractive forces. A similar problem arises in connection with molecules such as H_2 , N_2 , CH_4 , C_2H_6 , since they also form very stable units that do not interact strongly with one another. This problem is of importance to us because these molecules are constituents of an important class of solids.

The origin of the weak intermolecular forces is believed to be largely the same as that in the case of the rare gases, namely, the van der Waals interaction. Computations that support this are discussed in Sec. 88, Chap. X. The problem of understanding the internal stability of these molecules, however, is not so easy to answer in a quantitative way. Since the internal binding energy of the stablest molecules is of the order of magnitude of 1 ev per electron, this problem can be solved quantitatively only by solving the Schrödinger equation for molecular systems to a higher degree of accuracy than is generally feasible at present.

Physical chemists have attempted to avoid some of the difficulties associated with solving the Schrödinger equation accurately by introducing semiempirical schemes. These are usually patterned after one of the one-electron schemes, the matrix components that enter into the theoretical results being judiciously replaced by quantities derived from empirical data. From what is known of the accuracy of one-electron approximations, it is doubtful whether actual computations based on these one-electron schemes would yield results that agree with experimental results as well as those of the semiempirical schemes do. The latter are such a distinct improvement over older valence theories, however, that they have great value in discussing many properties of molecules. We shall present some of the qualitative results of these schemes in the sections of Chap. XIII that deal with valence crystals.

¹ See the review article by J. H. Van Vleck and A. Sherman, *Rev. Modern Phys.*, **7**, 167 (1935).

CHAPTER VIII

THE BAND APPROXIMATION

60. Qualitative Importance of the Band Scheme.—Prior to the introduction of quantum mechanics, it had been believed that insulators have low electronic conductivities because their valence electrons are localized on definite atoms or molecules and cannot jump from one atom to another. The electrons in metals, on the other hand, were considered to be free to roam throughout the lattice, and the high conductivity was believed to arise from this freedom of motion. If we attempt to use these qualitative notions in order to understand both the conductivities and the cohesive energies of solids, we face considerable difficulty unless we are willing to assume that the binding forces arise from essentially different sources in each type of solid.

Suppose, for example, we assume that there are only two kinds of interatomic force, namely: (1) electrostatic forces between bound charge distributions, and (2) undefined forces that are ultimately connected with the presence of free electrons and are of primary importance in metals. It is possible to account for the cohesion and insulating properties of ionic and molecular crystals in terms of (1). We may assume, as is done in the Madelung-Born theory, that the constituents of ionic crystals are ions and that the main part of the cohesion arises from the electrostatic attraction between these. Similarly, we may assume that the molecular constituents of molecular crystals are electrostatically neutral and that the cohesive energy arises from multipole forces of an electrostatic type. Since these forces should be weaker than the forces between ions, we are able to understand the relatively smaller cohesive energies of molecular crystals.

We meet with difficulty in discussing insulating valence crystals such as diamond. In this case, the atoms are electrostatically neutral, as in molecular crystals, and yet the cohesion is as great as in ionic crystals and metals. This difficulty was removed in classical theory by assuming that in addition to (1) and (2) there are valence forces which are responsible for the large cohesion of diamond and quartz.

As in many other cases in which classical views led to complication, the introduction of quantum mechanics produces order in a relatively simple way. In particular, the band concept of solids, which is based

upon the Bloch scheme and which has been developed by many workers,¹ has been very useful in coordinating many of the properties of solids that could not be adequately understood before. It may be recalled that the Bloch scheme is based on a one-electron approximation in which the one-electron functions have the same amplitude at equivalent positions in each unit cell. We shall see in the next section that these functions have the form

$$\psi_{\mathbf{k}} = \chi_{\mathbf{k}}(\mathbf{r})e^{2\pi i \mathbf{k} \cdot \mathbf{r}}, \quad (1)$$

where \mathbf{r} is the position vector, whose components are x, y, z ; $\chi_{\mathbf{k}}(\mathbf{r})$ has the translational periodicity of the lattice; and \mathbf{k} is a wave-number vector. \mathbf{k} may be defined in terms of the reciprocal lattice of the crystal (cf. Sec. 22) if the Born-von Kármán boundary conditions are used in determining the functions (1).

In the simplest case, $\chi_{\mathbf{k}}$ is a constant so that $\psi_{\mathbf{k}}$ is a free-electron wave function for which the dependence of energy on \mathbf{k} is

$$\epsilon(\mathbf{k}) = \frac{h^2}{2m} \mathbf{k}^2 \quad (2)$$

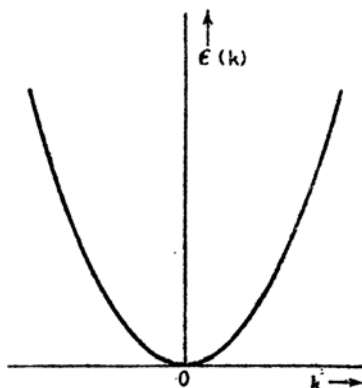


FIG. 1.—The $\epsilon(\mathbf{k})$ curve for perfectly free electrons.

(see Fig. 1). This approximation corresponds to that of the Sommerfeld theory of metals, which was discussed in Chap.

IV. We shall see that its use is equivalent to assuming that the Fock-Hartree field for the electrons is constant—a condition that is nearly satisfied in many simple metals.

In the opposite extreme, corresponding to tightly bound inner-shell electrons, $\chi_{\mathbf{k}}$ is zero everywhere in the unit cell except in the immediate vicinity of the particular atom the inner shells of which are being described. It turns out, in this case, that the portion of $\chi_{\mathbf{k}}$ near the atom is identical with the inner-shell wave function of the free atom. Moreover, the energy $\epsilon(\mathbf{k})$ is practically independent of \mathbf{k} and has a different value for each type of inner-shell electron.

¹The qualitative existence of bands was first pointed out by M. J. O. Strutt, *Ann. Physik*, **84**, 485 (1927); **85**, 129 (1928). The band picture was extended by: L. Brillouin, *Compt. rend.*, **191**, 198, 292 (1930); *Jour. phys.*, (VII), **1**, 377 (1930) [cf. also *Quantenstatistik* (Julius Springer, Berlin, 1931)]; P. M. Morse, *Phys. Rev.*, **35**, 1310 (1930); R. Peierls, *Ann. Physik*, **4**, 121 (1930) [cf. also *Ergebnisse exakt. Natur.*, **11**, 264 (1932)]; R. de L. Kronig and W. G. Penney, *Proc. Roy. Soc.*, **130**, 499 (1931).

In intermediate cases in which $\chi_{\mathbf{k}}$ is neither completely constant nor localized, the energy $\epsilon(\mathbf{k})$, as a function of \mathbf{k} , is not completely quasi-continuous, as in the free-electron case. Instead, $\epsilon(\mathbf{k})$ exhibits the property of *banding*; that is, it is quasi-continuous for large ranges of \mathbf{k} but is discontinuous for certain values of \mathbf{k} . The regions of continuity lie between a set of concentric polyhedra which are centered about the origin of \mathbf{k} space; the points where the discontinuities occur are at the surfaces of the polyhedra. In the following sections, we shall investigate the relationships from which the form of these polyhedra may be deduced. The regions between the polyhedra are called "zones" and the polyhedra are called "zone boundaries." The magnitude of the discontinuities at the zone boundaries depends upon the extent to which $\chi_{\mathbf{k}}$ deviates from a constant value and is zero for perfectly free electrons. Figure 2 illustrates the discontinuities for a typical case in which the $\epsilon(\mathbf{k})$ curve is plotted as a function of the points on a line that passes through the point $\mathbf{k} = 0$. The discontinuities occur at points where this line intersects the different polyhedra.

The transition from the free-electron type of wave function to the rigidly bound electron type may be regarded as taking place when $\chi_{\mathbf{k}}$ changes from a constant to a highly localized function. During this transition $\epsilon(\mathbf{k})$ develops discontinuities which become larger and larger until $\epsilon(\mathbf{k})$ is constant within each zone.

When constructing an antisymmetric wave function from the Bloch functions, we are not allowed to assign a function of given \mathbf{k} to more than two electrons because of the Pauli principle. Hence, we must use a large number of different values of \mathbf{k} in order to assign functions to all electrons in a solid. We shall say that levels are occupied when the wave functions corresponding to them have been assigned to electrons. As in treating atoms and molecules, we shall assume that, in the normal state of the system, the lowest one-electron energy levels are occupied as far as possible. It turns out that the number of states in any zone is an integer multiple of the number of unit cells in the crystal. Hence, it

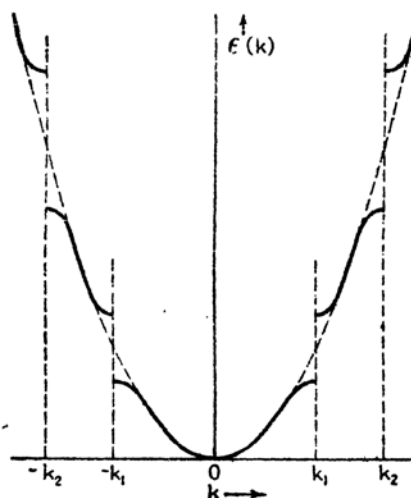


FIG. 2.—Typical $\epsilon(\mathbf{k})$ curve for electrons that are not perfectly free. This corresponds to values of \mathbf{k} that lie on a line passing through the origin of \mathbf{k} space. The discontinuities occur at points k_1 and k_2 at which the line cuts the zone-boundary polyhedra. These points are different for different directions in \mathbf{k} space. The dashed parabola represents the $\epsilon(\mathbf{k})$ curve for perfectly free electrons (cf. Fig. 1).

may happen that a low-lying set of zones is completely occupied and that the levels in zones of higher energy are completely empty. The conditions that must be satisfied if this is to occur in the lowest state are as follows:

a. The number of electrons present must be exactly equal to the number of states in an integer number of zones. It turns out that this condition is satisfied in all insulators and in the alkaline earth metals. It is not satisfied in the alkali metals, for in these there are twice as many states per zone as there are electrons.

b. The highest filled zone must not have any energy levels that lie above the lowest energy levels in the next highest zone. If these energy bands overlap, the energy of the system could be lowered by transferring electrons from the highest levels of the last filled zone to the lower levels of the next zone. Complete filling of zones is prohibited in the alkaline earth metals because condition *b* is not satisfied.

Let us consider the difference between the properties of a substance in which occupied zones are completely filled and those of one in which they are not. In both cases, the electrons normally are paired in such a way that for each electron moving in a given direction there is another moving in the opposite direction with the same speed. Hence, the current carried by each electron cancels that carried by the other, and the resultant current of the entire solid is zero. This statistical balance may be disturbed easily in a substance that does not have completely filled bands; for, by the application of a weak electrostatic field, some of the electrons may be made to jump to the near-lying unoccupied levels, thus changing the average velocity from zero to a finite value. This type of shift in statistical balance was described in the sections of Chap. IV that deal with the Lorentz-Sommerfeld theory of conductivity. On the other hand, the highest occupied levels may be separated from the unoccupied ones by several electron volts if the solid has completely filled bands. In this case, a very strong electrical field would be required to induce the electrons to jump from occupied to unoccupied levels. Hence, the crystal with completely filled bands is an insulator even though its electrons are wandering throughout the lattice.

Thus, we see that the electrical properties of two substances that have similar one-electron functions may be vastly different. In this connection, we may anticipate that the one-electron functions of diamond and of metals are similar so that the cohesive forces have similar origin in both cases.

Since the introduction of the band scheme permits us to modify the classical concept of "bound electrons" in valence crystals, it is natural to ask whether or not we need retain the classical concept when discussing ionic and molecular crystals. This question may be answered fairly

unambiguously in both cases. In ionic crystals, the classical picture is a fair approximation but is not rigorously correct since the valence-electron wave functions are not entirely localized about the cations. This means that the Bloch functions of ionic crystals possess properties similar to those for metals. Conductivity is absent because of the way in which zones are filled rather than because the electrons are not free to move from one atom to another. Since the amplitude of the wave functions unquestionably is small at regions midway between molecules in molecular crystals, we should expect small electronic conductivity even if the zone structure did not exist. Nevertheless, the presence of filled zones plays a predominant role in entirely prohibiting electronic conductivity.

It should be kept in mind that the zone scheme rests upon an approximation. Although it serves a useful purpose in providing a model of a solid that is adequate for describing many important properties in a simple, straightforward way, the picture is not a perfect one, and it may lead to incorrect results if it is not used with sufficient care. In particular, it should not be applied without reserve to problems that involve excited states of insulators, for reasons which are discussed in Chap. XII.

61. The Connection between Zone Structure and Crystal Symmetry. Since the concept of zone structure is based upon a particular type of one-electron approximation, it is natural to ask for those properties of the Fock-Hartree equations that in this case lead to the existence of zones. This question has the relatively simple answer that zone structure is characteristic of any eigenvalue equation in which the operator remains unchanged under the primitive translations of the lattice. Thus the eigenvalues E of any equation that has the form

$$H\psi = E\psi, \quad (1)$$

where H has crystallographic symmetry, are practically continuous except for certain unallowed regions. This topic may be made the basis of a very elegant and practical group-theoretical discussion¹ in which we shall not indulge at this point. We shall consider, instead, several examples of equations of type (1) in which the symmetry conditions that are required for zone structure occur. From these, we may derive general conditions from which the precise form of zone structure may be determined in any case.

Case a. The One-dimensional Oscillating Lattice in Classical Mechanics.—One of the simplest problems in which zone properties occur is that of determining the vibrational modes of a long one-dimensional

¹ The group-theoretical side of the existence of zones is discussed in the following papers: F. Seitz, *Ann. Math.*, **37**, 17 (1936); L. P. Bouckaert, R. Smoluchowski, and E. Wigner, *Phys. Rev.*, **50**, 58 (1936); C. Herring, *Phys. Rev.*, **52**, 361 (1937), **52**, 365 (1937); M. I. Chodorow and M. F. Manning, *Phys. Rev.*, **52**, 731 (1937).

lattice of particles that interact with harmonic forces. Several special cases of this problem were discussed in Sec. 21, Chap. III. We shall review these now.

First, we considered the case in which the particles have equal mass and are separated by a distance a . The equations of motion for this case are

$$m \frac{d^2 x_n}{dt^2} = -\mu[(x_n - x_{n-1}) - (x_{n+1} - x_n)] \quad (2)$$

where x_n is the displacement of the n th atom from its equilibrium position, μ is Hooke's constant for interaction between nearest neighbors, and m is the mass of a particle. Since we are searching for solutions that are periodic in time, we may replace Eq. (2) by

$$-m(2\pi\nu)^2 x_n = -\mu[(x_n - x_{n-1}) - (x_{n+1} - x_n)]. \quad (3)$$

This equation is of the type (1) since we may regard the x_n as components of a vector \mathbf{X} . Thus,

$$\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \\ x_{n+1} \\ \vdots \end{pmatrix}. \quad (4)$$

Using (4) we may write Eq. (3) in the form

$$\sum_m M_{nm} X_m = -\nu^2 X_n, \quad (5)$$

or

$$\mathbf{M} \cdot \mathbf{X} = -\nu^2 \mathbf{X} \quad (5a)$$

where \mathbf{M} is a tensor or matrix the components of which are

$$M_{n,n} = -\frac{2\mu}{4\pi^2 m}, \quad M_{n,m} = 0, \quad m \neq \begin{cases} n \\ \text{or} \\ n \pm 1 \end{cases}$$

$$M_{n,n+1} = M_{n-1,n} = \frac{\mu}{4\pi^2 m},$$

The matrix \mathbf{M} clearly has the symmetry of the lattice since Eqs. (2) are the same for all masses. The solutions of these equations were found to,

have the form

$$x_n = Ae^{2\pi i \sigma n a}. \quad (6)$$

Here σ is the wave number l/Na , where l is an arbitrary integer and N is the number of cells in the lattice. The frequency associated with (6) is

$$\nu_l = \frac{1}{2\pi} \sqrt{\frac{\mu}{m}} |\sin \pi \sigma a|. \quad (7)$$

The independent values of l may be chosen to range from $-N/2$ to $+N/2$. Equation (7), for the corresponding range of σ , is shown in Fig. 7, Sec. 21. This curve does not exhibit the discontinuities characteristic of zone structure, for only one zone occurs in the present problem.

Next, we shall consider the extended problem in which two different masses occur in the unit cell. If the particles are separated uniformly by a distance $a/2$ and are labeled by integers extending from zero to $2N$, the normal modes are

$$\begin{aligned} x_{2n} &= Be^{2\pi i \sigma \left(\frac{2n}{2}a\right)}, \\ x_{2n+1} &= Ae^{2\pi i \sigma \left(\frac{2n+1}{2}a\right)}, \end{aligned} \quad (8)$$

where the wave number σ is again equal to l/Na . The frequencies and the constants A and B may be determined by solving an appropriate second-order secular equation which was discussed in Sec. 21. We there found the result

$$4\pi^2 \nu^2 = \frac{\mu}{mM} (M + m \pm \sqrt{M^2 + m^2 + 2Mm \cos 2\pi \sigma a}),$$

in which we may choose the independent range of σ to extend from $-1/a$ to $1/a$ and obtain the result shown in Fig. 8, Sec. 21. Discontinuities characteristic of zone structure occur at $\sigma = \pm 1/2a$, so that there are two zones in this case.

If another mass is added to the unit cell in such a way that neighboring masses are separated by a distance $a/3$, the new normal coordinates are

$$\begin{aligned} x_{3n} &= Ae^{2\pi i \sigma \frac{3n}{3}a}, \\ x_{3n+1} &= Be^{2\pi i \sigma \frac{3n+1}{3}a}, \\ x_{3n+2} &= Ce^{2\pi i \sigma \frac{3n+2}{3}a}, \end{aligned} \quad (9)$$

and so forth, if the masses are at points labeled by integers extending from zero to $3N$. There are three zones in this case, since discontinuities occur when σ is equal to $\pm 1/2a$ and $\pm 1/a$.

It should be noted that we may write Eqs. (6), (8), and (9) in the form

$$x(\xi) = \chi(\xi)e^{2\pi i\sigma\xi} \quad (10)$$

where ξ is the positional coordinate of a given mass, $x(\xi)$ is the displacement of the mass from its equilibrium position, and $\chi(\xi)$ is a discontinuous function having periodicity a that takes values different from zero only at points where atoms are situated. Thus, in the case corresponding to Eq. (9), χ is equal to A at $\xi = 3na/3$, to B at $\xi = (3n + 1)\frac{a}{3}$, and to C at $\xi = (3n + 2)\frac{a}{3}$ and is zero everywhere else. The form of

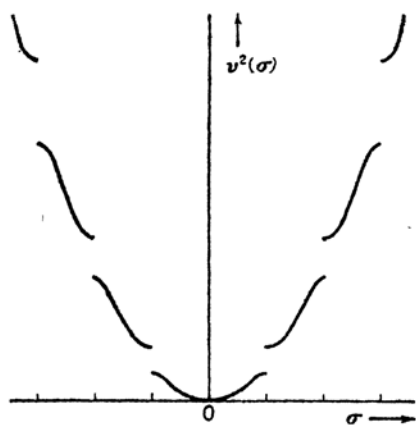


FIG. 3.— $v^2(\sigma)$ curve for the normal modes of a continuous string having periodically varying density and Hooke's constant (extended zone scheme). If the string were perfectly uniform the $v^2(\sigma)$ curve would be a parabola. This curve should be compared with that of Fig. 2.

the function (10) evidently is the same as that of the function (1) in the preceding section.

If we continue to add masses at equivalent points within the unit cell, we eventually find it convenient to use a density function $\rho(\xi)$ and a force function $\mu(\xi)$, both of which are continuous and have periodicity a . The normal coordinate should still be expressible in the form (10) but $\chi(\xi)$ should now be a continuous function of ξ with periodicity a . The independent modes of vibration then correspond to values of σ extending from $-\infty$ to $+\infty$, and v^2 is discontinuous whenever σ is equal to $\pm r/2a$, where r is an arbitrary integer. This case, in which there evidently is an infinite number of zones, is represented schematically in

Fig. 3. The differential equation that is satisfied by $x(\xi)$ is the wave equation

$$-4\pi^2 v^2 x = \frac{\mu}{\rho} \frac{d^2 x}{d\xi^2}$$

which has the eigenvalue form (1).

Case b. The One-dimensional Schrödinger Equation. 1. Kramers' general treatment.—From among the many methods that have been developed to show that the eigenvalues of the Schrödinger equation exhibit band structure when the potential function is periodic, we shall select a particularly good one that is due to Kramers.¹

¹ H. A. KRAMERS, *Physica*, **2**, 483 (1935).

Suppose that we have an equation of the form

$$\frac{d^2\psi}{dx^2} + [E - V(x)]\psi = 0 \quad (11)$$

where

$$V(x + a) = V(x). \quad (12)$$

Then, if $\psi_1(x)$ and $\psi_2(x)$ are solutions of this equation, the functions $\psi_1(x + a)$ and $\psi_2(x + a)$ also are solutions. Since a second-order equation possesses only two independent solutions, we must have the relationships

$$\begin{aligned} \psi_1(x + a) &= a_{11}\psi_1(x) + a_{12}\psi_2(x), \\ \psi_2(x + a) &= a_{21}\psi_1(x) + a_{22}\psi_2(x), \end{aligned} \quad (13)$$

if $\psi_1(x)$ and $\psi_2(x)$ are independent. From these equations, we may derive the relationship

$$\begin{vmatrix} \psi_1(x + a) & \psi_2(x + a) \\ \psi'_1(x + a) & \psi'_2(x + a) \end{vmatrix} = \begin{vmatrix} \psi_1(x) & \psi_2(x) \\ \psi'_1(x) & \psi'_2(x) \end{vmatrix} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \quad (14)$$

where $\psi' = d\psi/dx$. Now, the quantity

$$\begin{vmatrix} \psi_1(x) & \psi_2(x) \\ \psi'_1(x) & \psi'_2(x) \end{vmatrix}, \quad (15)$$

which is called the "*Wronskian*," is a constant in the present case.¹ Hence, we may conclude that

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = 1. \quad (16)$$

We may now choose linear combinations φ_1 and φ_2 of ψ_1 and ψ_2 that have the property

$$\begin{aligned} \varphi_1(x + a) &= \lambda_1\varphi_1(x) \\ \varphi_2(x + a) &= \lambda_2\varphi_2(x). \end{aligned} \quad (17)$$

The coefficients λ_1 and λ_2 may be determined from the a in Eq. (13) by means of the equation

$$\begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = \lambda^2 - (a_{11} + a_{22})\lambda + 1 = 0. \quad (18)$$

The quantity $\mu = (a_{11} + a_{22})$ is real since ψ_1 and ψ_2 may always be chosen to be real functions.

Kramers distinguishes between the three cases $|\mu| > 2$, $|\mu| < 2$, $|\mu| = 2$, which we shall discuss categorically.

¹See, for example, WHITTAKER and WATSON, *Modern Analysis* (Cambridge University Press, 1935).

i. When $|\mu|$ is greater than 2, Eq. (18) has unequal real roots so that φ_1 and φ_2 satisfy the equations

$$\varphi_1(x+a) = \lambda_1 \varphi_1(x), \quad \varphi_2(x+a) = \frac{1}{\lambda_1} \varphi_2(x). \quad (19)$$

ii. When $|\mu|$ is less than 2, the roots are complex conjugates of one another, whence

$$\begin{aligned} \varphi_1(x+a) &= e^{ia} \varphi_1(x), \\ \varphi_2(x+a) &= e^{-ia} \varphi_2(x). \end{aligned} \quad (20)$$

It may be shown very easily that in this case φ_1 and φ_2 are complex conjugates of one another.

iii. When $|\mu|$ is equal to 2, λ is ± 1 , both roots are equal, and Eq. (6) is replaced by

$$\begin{aligned} \varphi_1(x+a) &= \pm \varphi_1(x), \\ \varphi_2(x+a) &= \pm \varphi_2(x) + b \varphi_2(x). \end{aligned} \quad (21)$$

Thus, both functions φ_1 and φ_2 satisfy the equation

$$\varphi(x+a) = \pm \varphi(x) \quad (21a)$$

only when b vanishes.

In case *i*, the ratio

$$\frac{\varphi(x+na)}{\varphi(x)} \quad (22)$$

certainly becomes infinite when n approaches either $+\infty$ or $-\infty$. Hence, this type of eigenfunction must be excluded if the periodic field extends over the entire range of x between $+\infty$ and $-\infty$.

In case *ii*, both eigenfunctions are periodic and satisfy the relation

$$|\varphi(x+a)| = |\varphi(x)|.$$

These solutions are allowable as long as they remain finite in any unit cell.

In case *iii*, there is at least one solution of type (21a) and possibly two, depending upon whether b is zero or not.

The solutions for case *i* correspond to the unallowed regions of energy, whereas those for cases *ii* and *iii* correspond to the allowed regions. We shall see that the functions corresponding to case *iii* are solutions associated either with points in wave-number space at the zone boundaries or at points such as $\delta = 0$ for which the solutions have periodicity a .

It should be mentioned that the solutions for case *i* need not always be excluded if the periodic field does not extend to infinity; for in the finite case, which corresponds to an actual crystal, the wave functions

associated with some values of E in the allowed region may not diverge. In these cases, which were first pointed out by Tamm, the functions have their maximum absolute values near the boundary of the lattice and decrease rapidly on both sides of this point. We shall discuss these solutions further in Sec. 70.

Kramers was able to express the quantities μ , $d\mu/dE$, and $d^2\mu/dE^2$ in terms of integrals that involve the functions φ_1 and φ_2 . From these expressions, he deduced that $\mu(E)$ has the form illustrated in Fig. 4,

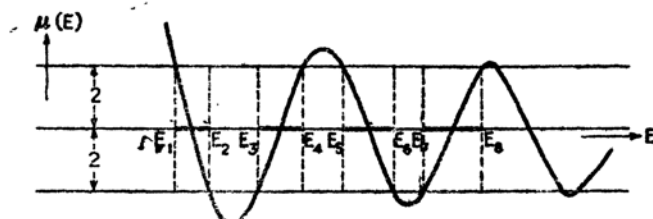


FIG. 4.—The $\mu(E)$ curve. The allowed values of E correspond to the ranges in which μ lies between 1 and -1 . (After Kramers.)

which approaches $2 \cos a\sqrt{E}$ when E approaches ∞ and approaches $e^{a\sqrt{-E}}$ when E approaches $-\infty$. In intermediate regions, it oscillates between values greater than 2 and less than -2 , in the manner shown. Thus, there are alternate continuous bands of allowed and unallowed levels. The unallowed bands become vanishingly small when E is large and positive. All values of E that are sufficiently negative are unallowed because of the monotonic increase of μ .

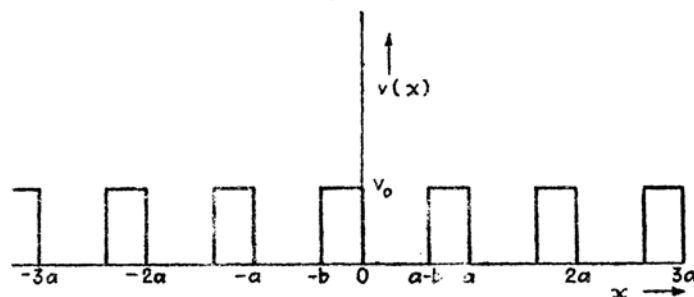


FIG. 5.—The one-dimensional periodic potential of Kronig and Penney.

2. *The case of Kronig and Penney.*—One of the simplest examples of a one-dimensional periodic field has been treated by Kronig and Penney.¹ This example merits attention because of the direct way in which it yields the general features of zone structure. Let us consider the periodic potential illustrated in Fig. 5, for which

¹ KRONIG and PENNEY, *op. cit.*, 499 (1931). See also V. ROJANSKY, *Introductory Quantum Mechanics*, Sec. 49 (Prentice-Hall, Inc., New York, 1938).

$$\begin{aligned} V &= V_0, & -b \leq x \leq 0, \\ V &= 0, & 0 \leq x \leq a - b, \\ V(x + a) &= V(x). \end{aligned} \quad (23)$$

Thus, V has the constant value V_0 for a range of length b in each unit cell. In the regions where V is zero, the general solution of the Schrödinger equation is

$$\psi_1 = Ae^{i\alpha x} + Be^{-i\alpha x} \quad (24)$$

where

$$\alpha = \frac{\sqrt{2mE}}{\hbar}.$$

In the other region, we have

$$\psi_2 = Ce^{\beta x} + De^{-\beta x} \quad (25)$$

where

$$\beta = \frac{\sqrt{2m(V_0 - E)}}{\hbar}.$$

Since we are searching for solutions of Kramers' class *ii*, we must have

$$\begin{aligned} \psi_2(-b) &= e^{-i\alpha a} \psi_1(a - b), \\ \psi_2'(-b) &= e^{-i\alpha a} \psi_1'(a - b), \end{aligned} \quad (26)$$

where $\lambda/2\pi$ is the wave number. In addition, we must have

$$\begin{aligned} \psi_2(0) &= \psi_1(0), \\ \psi_2'(0) &= \psi_1'(0), \end{aligned} \quad (27)$$

because of the continuity requirement at $x = 0$. We find, after substituting (24) and (25) in (26) and (27) and solving the determinantal compatibility equation, the condition

$$\cos \lambda a = \frac{(\beta^2 - \alpha^2)}{2\alpha\beta} \sinh \beta b \sin \alpha(a - b) + \cosh \beta b \cos \alpha(a - b), \quad (28)$$

which may be used to determine the allowed values of E .

Following Kronig and Penney, we may, at this point, introduce the simplifying conditions

$$\begin{aligned} b &\rightarrow 0, \\ V_0 &\rightarrow \infty, \end{aligned}$$

and we may stipulate that these limiting values are approached in such a way that the quantity

$$\frac{mV_0}{\hbar^2} b(a - b) \quad (29)$$

remains constant. This restriction assures us that the "potential area" $V_0 b$ is finite. With these conditions, the quantity βb in Eq. (28) is equal to

$$\sqrt{\frac{2cb}{a-b}}$$

where c is the limiting value of (29). Equation (28) approaches the value

$$\cos \lambda a = \frac{c}{\alpha a} \sin \alpha a + \cos \alpha a \quad (30)$$

in the limit as $b \rightarrow 0$. This equation may be satisfied whenever the quantity on the right-hand side lies in the interval from -1 to $+1$ since λ may then take real values. As is shown in Fig. 6, we obtain allowed

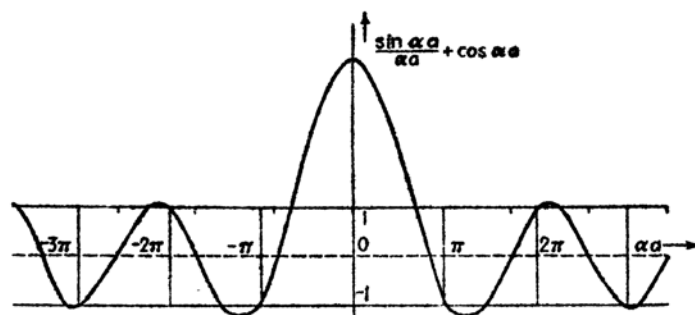


FIG. 6.—The function $\{c \sin \alpha a / \alpha a + \cos \alpha a\}$. The allowed values of E are given by those ranges of $\alpha = \sqrt{2mE}/\hbar$ in which this function lies between -1 and $+1$. It may be seen that the unallowed ranges become smaller as E increases. This curve is analogous to that of Fig. 4. (After Kronig and Penney.)

bands of energy that become closer and closer as E approaches infinity. It may be verified that the coefficients A and B in (24) have the ratio

$$-\frac{A}{B} = \frac{1 - e^{-i(\lambda+\alpha)a}}{1 - e^{-i(\lambda-\alpha)a}} \quad (31)$$

in the range of x that extends from 0 to a . The values of A and B in any other range, such as that extending from na to $(n+1)a$, may be obtained from the values in the range from 0 to a by multiplying by the factor $e^{i\alpha(n+1)a}$.

Case c. The Two-dimensional Schrödinger Equation.—By discussing three types of two-dimensional cases, we may obtain some of the most important principles of zone-structure theory. The first case, which was investigated by Brillouin,¹ deals with electrons that are practically free. The periodic potential then may be treated as a small perturbation. In

¹ Cf. *Quantenstatistik*, *op. cit.*

the second case, the periodic potential is larger than in the first but may be expressed in the form

$$V(x, y) = V_x(x) + V_y(y) + V_p(x, y) \quad (32)$$

where $V_x(x)$ and $V_y(y)$ are large compared with $V_p(x, y)$. The Schrödinger equation is separable in this case if $V_p(x, y)$ is neglected. The zones that are obtained when V_p is included by perturbation theory usually do not have the same form as in Brillouin's case. We shall see, however, that the two sets may be made identical by suitable rearrangement. The third case to be discussed is the more general one in which perturbation methods cannot be employed.

1. *Brillouin's case.*—We shall assume that the Schrödinger equation has the form

$$-\frac{\hbar^2}{2m} \left(\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} \right) + [V(x, y) - E]\psi = 0.$$

For simplicity, we shall discuss the case in which $V(x, y)$ has the periodicity a in both the x and y directions. The normalized unperturbed eigenfunctions then have the form

$$\psi_{\mathbf{k}}^0 = \frac{1}{\sqrt{S}} e^{2\pi i \mathbf{k} \cdot \mathbf{r}} \quad (33)$$

where S is the area of the lattice, which we shall assume is a square having the edge length Na . If we adopt the Born-von Kármán boundary conditions, the permissible components of \mathbf{k} are

$$k_x = \frac{n_x}{Na}, \quad k_y = \frac{n_y}{Na} \quad (34)$$

where n_x and n_y are arbitrary integers. The unperturbed energy of the wave function (33) is

$$E_{\mathbf{k}}^0 = \frac{\hbar^2}{2m} \mathbf{k}^2, \quad (35)$$

and the entire energy spectrum above zero is quasi-continuous.

To the first order in perturbed quantities, the eigenfunctions of Eq. (32) are

$$\psi_{\mathbf{k}}(x, y) = \frac{1}{\sqrt{S}} e^{2\pi i \mathbf{k} \cdot \mathbf{r}} + \sum_{\mathbf{k}'} \frac{\left[\frac{1}{S} \int e^{-2\pi i \mathbf{k}' \cdot \mathbf{r}'} V(x', y') e^{2\pi i \mathbf{k} \cdot \mathbf{r}'} dx' dy' \right] e^{2\pi i \mathbf{k}' \cdot \mathbf{r}}}{E_{\mathbf{k}}^0 - E_{\mathbf{k}'}^0} \quad (36)$$

where the summation extends over all values of k'_x and k'_y and the integration extends over the entire lattice. The integrals in Eq. (36) may

be simplified by use of the translational symmetry of $V(x, y)$. We have in fact

$$\int_S e^{2\pi i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} V(x, y) dx dy = \left[\int_A e^{2\pi i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} V(x, y) dx dy \right] \sum_{\mathbf{d}} e^{2\pi i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{d}} \quad (37)$$

where \mathbf{d} is the vector

$$\mathbf{d} = \begin{pmatrix} na \\ ma \end{pmatrix}, \quad (m, n = 0, 1, 2, \dots, N)$$

and the integral in the right-hand term extends over the unit cell, that is, over the ranges of x and y lying between zero and a . The summation in (37) vanishes identically unless the vector $(\mathbf{k} - \mathbf{k}')$ satisfies the relation

$$\begin{cases} (k_x - k'_x)a = p_x, \\ (k_y - k'_y)a = p_y, \end{cases} \quad (38)$$

where p_x and p_y are integers. Hereafter, we shall reserve the letter \mathbf{K} for wave-number vectors of which the scalar products with the primitive translation vectors of the lattice are integers. Thus, we may write Eq. (38) in the form

$$\mathbf{k}' - \mathbf{k} = \mathbf{K}. \quad (39)$$

Vectors of type \mathbf{K} possess the important property that the functions $e^{2\pi i\mathbf{K}\cdot\mathbf{r}}$ have the same periodicity as the lattice. Using (38), we may now write Eq. (36) as

$$\psi_{\mathbf{k}} = \chi_{\mathbf{k}}(x, y) e^{2\pi i\mathbf{k}\cdot\mathbf{r}} \quad (40)$$

where the function

$$\chi_{\mathbf{k}} = \frac{1}{\sqrt{S}} \left(1 + \sum_{\mathbf{K}} \frac{\frac{e^{2\pi i\mathbf{K}\cdot\mathbf{r}}}{S} \int_S e^{-2\pi i\mathbf{K}\cdot\mathbf{r}'} V dx' dy'}{E_{\mathbf{k}}^0 - E_{\mathbf{k}+\mathbf{K}}^0} \right)$$

has the periodicity of the lattice. As we have remarked previously, the form of (40) may be deduced rigorously on the basis of symmetry.

Turning now to the energy, we have in the second approximation

$$E_{\mathbf{k}} = \frac{\hbar^2}{2m} \mathbf{k}^2 + \frac{1}{S} \int V dx dy + \frac{1}{S^2} \sum_{\mathbf{K}'} \frac{|\int e^{2\pi i(\mathbf{k}-\mathbf{K}')\cdot\mathbf{r}} V dx dy|^2}{E_{\mathbf{k}}^0 - E_{\mathbf{K}'}^0}. \quad (41)$$

The first two terms on the right do not affect the continuity of E as a function of \mathbf{k} . It may be seen, however, that the third term becomes infinite when \mathbf{k} and \mathbf{k}' satisfy Eq. (39) and the relation

$$E_{\mathbf{k}}^0 = E_{\mathbf{K}'}^0 \quad (42)$$

is fulfilled. It is precisely in this case that the perturbation method we have used needs revision. Instead of using the unperturbed wave functions (33), we should select those "proper linear combinations" of the degenerate functions that make the integral

$$V_{\mathbf{k}'\mathbf{k}} = \int \psi_{\mathbf{k}'}^0 V(x, y) \psi_{\mathbf{k}}^0 dx dy \quad (43)$$

vanish. Then the numerators of the troublesome terms vanish, and the degeneracy (42) is removed, so that there is a discontinuity in E at the points where Eqs. (39) and (42) are satisfied.

The origin of this discontinuity may be illustrated by approximate considerations of the following type. We shall assume that the values of \mathbf{k} and \mathbf{k}' for which degeneracy nearly occurs and for which Eq. (39) is satisfied occur in pairs. If $\psi_{\mathbf{k}}^0$ and $\psi_{\mathbf{k}+\mathbf{K}}^0$ are the functions associated with these pairs, the proper linear combinations $\psi_{\mathbf{k}}^{0'}$ and $\psi_{\mathbf{k}'}^{0'}$ have the form

$$\begin{cases} \psi_{\mathbf{k}}^{0'} = a\psi_{\mathbf{k}}^0 + b\psi_{\mathbf{k}+\mathbf{K}}^0 \\ \psi_{\mathbf{k}'}^{0'} = c\psi_{\mathbf{k}}^0 + d\psi_{\mathbf{k}+\mathbf{K}}^0 \end{cases}$$

in which a , b , c , and d must be chosen so that the off-diagonal components of the energy matrix

$$\begin{pmatrix} E_{\mathbf{k}}^0 + V_{\mathbf{k}\mathbf{k}} & V_{\mathbf{k},\mathbf{k}+\mathbf{K}} \\ V_{\mathbf{k}+\mathbf{K},\mathbf{k}} & E_{\mathbf{k}+\mathbf{K}}^0 + V_{\mathbf{k}+\mathbf{K},\mathbf{k}+\mathbf{K}} \end{pmatrix}$$

vanish. We shall assume that the V^0 are continuous functions of \mathbf{k} . The equations corresponding to this condition are

$$\begin{cases} \epsilon_{\mathbf{k}} a + V_{\mathbf{k},\mathbf{k}'}^0 b = \lambda a, \\ V_{\mathbf{k}',\mathbf{k}}^0 a + \epsilon_{\mathbf{k}'} b = \lambda b, \end{cases} \quad (44)$$

in which we have set $\mathbf{k}' = \mathbf{k} + \mathbf{K}$ and

$$\epsilon_{\mathbf{k}} = E_{\mathbf{k}}^0 + V_{\mathbf{k}\mathbf{k}}^0.$$

The λ for which Eqs. (44) have solutions are the new unperturbed energies and satisfy the equation

$$\lambda = \frac{(\epsilon_{\mathbf{k}} + \epsilon_{\mathbf{k}'}) \pm \sqrt{(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'})^2 + 4|V_{\mathbf{k}\mathbf{k}'}^0|^2}}{2} \quad (45)$$

When $|V_{\mathbf{k}\mathbf{k}'}^0|$ is negligible in comparison with $\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'}$, the roots of this equation are $\epsilon_{\mathbf{k}}$ and $\epsilon_{\mathbf{k}'}$. When $\epsilon_{\mathbf{k}}$ and $\epsilon_{\mathbf{k}'}$ become nearly alike, however, the roots do not become closer than $2|V_{\mathbf{k}\mathbf{k}'}^0|$, which implies a discontinuity in the energy versus \mathbf{k} curve. This is illustrated schematically in Fig. 7.

Now, Eq. (42) is satisfied whenever

$$|\mathbf{k}| = |\mathbf{k}'|, \quad (46)$$

because of Eq. (35). Hence, (39) and (46) are the only equations that must be satisfied if there is to be a discontinuity. It may be seen that these conditions determine a set of lines which satisfy the equations

$$\mathbf{k} \cdot \frac{\mathbf{K}}{2} = \frac{\mathbf{K}^2}{4} \quad (47)$$

where \mathbf{K} may be any one of the \mathbf{K} -type vectors. These lines are illustrated in Fig. 8 for the lattice under discussion.

Brillouin has pointed out that all similarly shaded regions in Fig. 8 may be pieced together in a unique way to form a square that is identical with the central zone. When this is done, points in any two squares that overlap when the squares are placed on top of one another differ by a vector of type \mathbf{K} . For this reason, similarly shaded regions are said to belong to the same zone. It may be observed that this piecing process requires only that the sections be *translated* to the central zone by a vector of type \mathbf{K} ; that is, the sections do not need to be folded or rotated. If it is recalled that $e^{2\pi i \mathbf{K} \cdot \mathbf{r}}$ has the translational periodicity of the lattice, it may be seen that the functions that go with a given zone may be written in the form

$$\chi_{\mathbf{k}} e^{2\pi i \mathbf{k} \cdot \mathbf{r}},$$

where \mathbf{k} ranges over all points in the inner zone and $\chi_{\mathbf{k}}$ has translational symmetry. In addition, it may be seen from Eq. (45) that $E_{\mathbf{k}}$ is continuous when regarded only as a function of the values of \mathbf{k} in the first zone. In other words, we may represent $E_{\mathbf{k}}$ as a multiple-valued function of the points in the first zone instead of as a single-valued, discontinuous function of the points in the entire \mathbf{k} plane. The correspondence between the two types of representation may be made clear by cutting up each of the surfaces in the multiple-valued representation and by placing them over the outer zones. The reduced-zone scheme, which uses only the inner zone, has the advantage that it requires knowing the geometrical form of only one zone.

It should be emphasized that the symmetry of the zone structure shown in Fig. 8 arises from the high degree of symmetry of $E_{\mathbf{k}}^0$, the

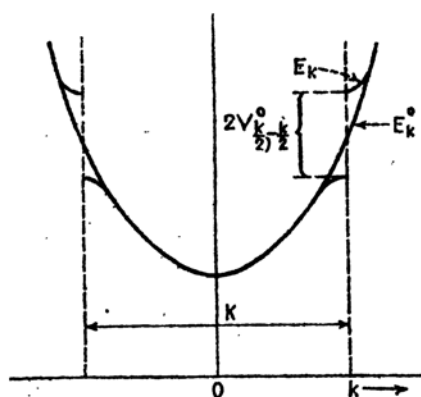


FIG. 7.—Schematic diagram showing the effect of the perturbing potential V on the energy levels. The unperturbed energy curve $E_{\mathbf{k}}^0$ is continuous and the energy associated with the state \mathbf{k} is equal to that of the state $-\mathbf{k}$. The matrix components of V vanish unless the two states satisfy the relation $\mathbf{k}' = \mathbf{k} + \mathbf{K}$, where \mathbf{K} is a principal lattice vector. If the matrix of V connecting these two states is diagonalized, the new $E_{\mathbf{k}}$ curve possesses a discontinuity at $\mathbf{k} = \pm \mathbf{K}/2$.

unperturbed energy surface. The zone structure does not have the same form when this symmetry is absent. We shall illustrate other possibilities in case II.

The important zone relations

$$\mathbf{k}' - \mathbf{k} = \mathbf{K}, \quad (48a)$$

$$k^2 = k'^2, \quad (48b)$$

occur in the theory¹ of diffraction of X rays by crystals in which they are called Laue's equations. An X ray having wave-number vector \mathbf{k} that impinges on a crystal can have its wave-number vector changed to \mathbf{k}' by diffraction only if these equations are satisfied. Thus, we may

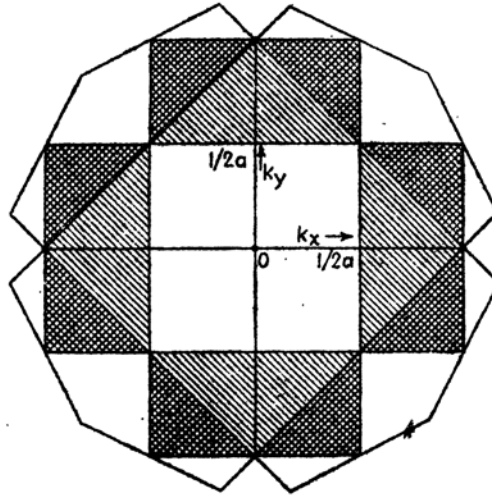


FIG. 8.—The first four Brillouin zones for a square, two-dimensional lattice. The similarly shaded areas may be translated into the first zone by vectors of type \mathbf{K} and will then exactly cover this zone. The heavy dots indicate principal vectors in \mathbf{k} space.

determine the X-ray diffraction pattern of a crystal from the Brillouin zone pattern, and vice versa. This identification of Eqs. (48a) and (48b) with Laue's equations shows that the occurrence of zones in the electronic problem is intimately connected with the wave properties of electrons.

2. *Case II.*—We shall now consider the Schrödinger equation when the potential has the form (32). If V_p is neglected, we know that the unperturbed equation may be separated into the ordinary equations

$$\left. \begin{aligned} -\frac{\hbar^2}{2m} \frac{d^2 \xi(x)}{dx^2} + (V_x - E_x) \xi(x) &= 0, \\ -\frac{\hbar^2}{2m} \frac{d^2 \eta(y)}{dy^2} + (V_y - E_y) \eta(y) &= 0. \end{aligned} \right\} \quad (49)$$

¹ Cf. A. H. COMPTON and S. K. ALLISON, *X-Rays* (D. Van Nostrand Company, Inc., New York, 1934).

The complete unperturbed wave function is

$$\psi^0(x, y) = \xi(x)\eta(y), \quad (50)$$

and the unperturbed energy is

$$E^0 = E_x + E_y. \quad (51)$$

From parts *a* and *b* of this section, it follows that

$$\left. \begin{aligned} \xi &= \chi_{k_x}(x)e^{2\pi i k_x x}, \\ \eta &= \chi_{k_y}(y)e^{2\pi i k_y y}, \end{aligned} \right\} \quad (52)$$

where k_x and k_y satisfy Eqs. (34). Hence,

$$\psi_{\mathbf{k}}^0 = \chi_{\mathbf{k}}^0 e^{2\pi i \mathbf{k} \cdot \mathbf{r}}, \quad (53)$$

just as in Eq. (40), where

$$\chi_{\mathbf{k}}^0 = \chi_{k_x} \chi_{k_y}.$$

Moreover, both E_x and E_y possess discontinuities at points that satisfy the relations

$$\begin{aligned} k_x &= \frac{n_x}{2a} \\ k_y &= \frac{n_y}{2a} \end{aligned} \quad (n_x, n_y = 0, \pm 1, \pm 2, \dots).$$

Since E^0 has the same discontinuities as E_x and E_y , we know that discontinuities appear in the unperturbed problem along the lines shown in Fig. 9, which illustrates the energy contours of the unperturbed energy function in a typical case.

When applying the perturbation theory, we may use the same simplifications that were used in simplifying Eq. (36). This procedure is permissible because χ in Eq. (53) has the same translational periodicity as V_p . The perturbed functions now are

$$\psi_{\mathbf{k}} = e^{2\pi i \mathbf{k} \cdot \mathbf{r}} \left(\chi_{\mathbf{k}}^0 + \sum_{\mathbf{K}} \chi_{\mathbf{k}+\mathbf{K}}^0 e^{2\pi i \mathbf{K} \cdot \mathbf{r}} \frac{\int_S \psi_{\mathbf{k}}^0 V \psi_{\mathbf{k}+\mathbf{K}}^{*0} dx' dy'}{E_{\mathbf{k}}^0 - E_{\mathbf{k}+\mathbf{K}}^0} \right). \quad (54)$$

Hence, the form of Eq. (40) is maintained, and many of the remarks made in connection with $E_{\mathbf{k}}^0$ in the preceding case are valid again. Thus, $E_{\mathbf{k}}$, the perturbed energy, is discontinuous whenever

$$E_{\mathbf{k}}^0 = E_{\mathbf{k}+\mathbf{K}}^0. \quad (55)$$

This fact may be derived from an equation analogous to (41). One difference between this case and the last is that Eq. (55) is not necessarily

satisfied when $|\mathbf{k}| = |\mathbf{k} + \mathbf{K}|$, as it was in Brillouin's problem, for the energy contours usually are not circles. Instead, they have lower symmetry which depends upon the symmetry of $V_x + V_y$. Let us assume for simplicity that the unperturbed potential has the symmetry of a square, which is the highest possible. We may then use one of the elementary results of the theory of symmetry,¹ which states that Eq. (55) is satisfied whenever $\psi_{\mathbf{k}}^0$ is sent into $\psi_{\mathbf{k}+\mathbf{K}}^0$ by the symmetry operators under which the Schrödinger equation is invariant. From the form of (54), it follows that this happens in the present highly symmetric case

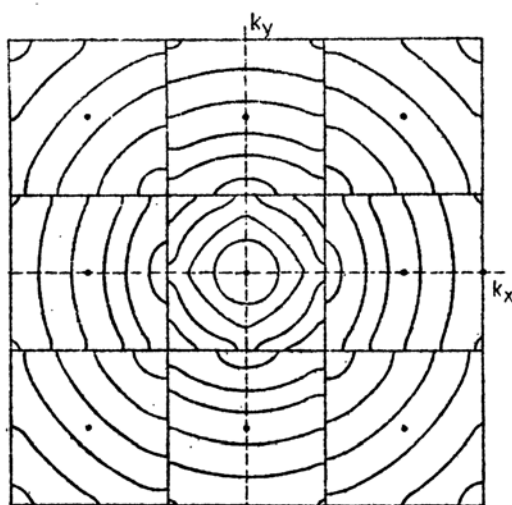


FIG. 9.—Energy contours of the zero-order eigenvalues when the potential has the form (32). Discontinuities occur only at the lines corresponding to the zones of one-dimensional lattices.,,

whenever \mathbf{k} is sent into $\mathbf{k} + \mathbf{K}$ by one of the eight symmetry operations of a square, that is, whenever

$$\begin{cases} k_x = \pm(k_x + K_x) \\ k_y = \pm(k_y + K_y) \end{cases} \quad \text{or} \quad \begin{cases} k_x = \pm(k_y + K_y) \\ k_y = \pm(k_x + K_x) \end{cases} \quad (56)$$

The \pm signs in the two sets of cases may be taken in arbitrary combinations. The conditions (56) include all the lines shown in Fig. 9 and all the additional straight lines shown in Fig. 10. Many of the lines that appear in Brillouin's pattern are absent, however, because (56) is not so stringent as the condition $|\mathbf{k}| = |\mathbf{k}'|$.

There is a large amount of accidental degeneracy in addition to the symmetrical degeneracy that occurs at the points satisfying (56). It is not difficult to see that this accidental degeneracy adds just enough zones to make up for the difference between the pattern of Fig. 10 and Bril-

¹ Cf. E. WIGNER, *Gruppentheorie* (Vieweg, Braunschweig, Germany, 1931).

loun's pattern. The additional zone boundaries, however, are usually curved rather than straight lines. The amount by which these curves deviate from straight lines depends upon the amount by which the energy surface $E_{\mathbf{k}}^0 = E_x + E_y$ deviates from a parabola of revolution. An additional curved zone is illustrated in Fig. 10 for the case corresponding to Fig. 9. The resulting pattern possesses square symmetry.

The sections having similar shading in Fig. 10 have the same total area as the central zone and fit exactly into the first zone if cut along the lines of the Brillouin pattern. Moreover, overlapping points of the two squares still correspond to values of \mathbf{k} that differ by a vector of type \mathbf{K} .

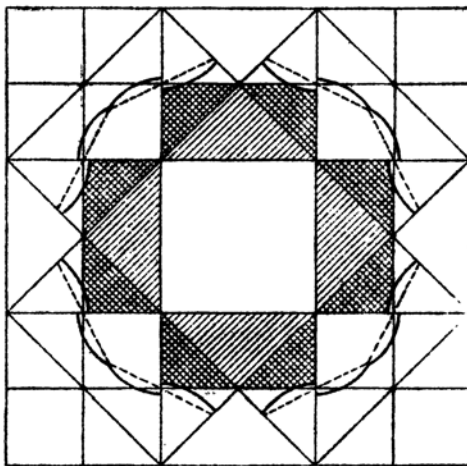


FIG. 10.—The lines of discontinuity in the case in which the non-diagonal matrix components of the perturbing term in (11) are taken into account. Only the straight lines are the same as for Brillouin's case, in which the zero-order eigenfunctions are free-electron functions (cf. Fig. 8). The additional straight lines of Brillouin's pattern are replaced by curved lines, of which one set is shown. The corresponding Brillouin zone is represented by dotted lines.

We shall present these statements without proof since they are easy to demonstrate. It follows that the differences between Brillouin's pattern and Fig. 10 have only superficial importance, for by properly cutting and translating the zones of Fig. 10 it is possible to piece them together to form the Brillouin pattern. As a result of this process, some of the starting eigenfunctions are relabeled, since a ψ previously associated with a point \mathbf{k} may be associated with a point $\mathbf{k} + \mathbf{K}$ after the redistribution. These facts show that knowledge of the central zone is sufficient to provide a complete description of zone structure, for we may always regard the energy surface as a multiple-valued function in this domain.

If the potential function (32) has no symmetry, aside from translational periodicity, we may anticipate that the symmetry of $E_{\mathbf{k}}^0$ is much lower than it was previously. It is true that the separation into Eqs. (49)

proceeds as before and that the unperturbed energy (51) has the discontinuity pattern of Fig. 9, but $E_{\mathbf{k}}$ does not have equal values at points that are connected by Eqs. (56). The remaining symmetry, which is called the "natural symmetry of the Schrödinger equation," is expressed by the theorem that ψ and ψ^* are independent solutions of the Schrödinger equation having the same energy, if ψ is a complex solution and the potential function is real. Now, the conjugates of the solutions ξ_{k_x} and η_{k_y} of Eq. (49) are ξ_{-k_x} and η_{-k_y} , respectively. Thus, the symmetry of the one-dimensional energy curves relative to the origin of \mathbf{k} space is due to the natural symmetry¹ of the Schrödinger equation. This symmetry

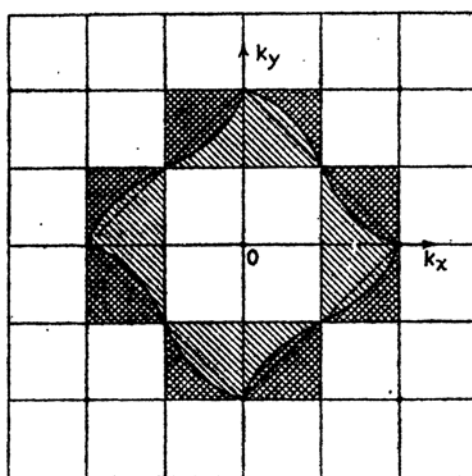


FIG. 11.—First, second and third zones in a case in which the potential function (32) has no symmetry other than translational symmetry. The only zones that are straight lines, as in Brillouin's scheme, are the vertical and horizontal lines of Fig. 9. All other zones are curved lines. Only one curved zone, namely the second, is shown in this figure. The fourth zone is curved as in Fig. 10.

carries over to the two-dimensional energy-level diagram and accounts for the straight lines of Fig. 9. Since all other degeneracy that occurs is accidental, the other zone boundaries usually are irregular curves, as in Fig. 11, which have the residual symmetry expressed by the equations

$$E_{k_x, k_y} = E_{-k_x, -k_y}, \dots$$

As before, the similarly shaded regions may be cut and pieced together to form the Brillouin pattern.

3. *The general case.*—It follows from the discussion of the two preceding cases that Brillouin's zone pattern usually does not occur unless a perturbation method in which the unperturbed wave functions are free-electron functions is used. The only unambiguous way in which to

¹ Cf. E. WIGNER, *Gött. Nachr.*, 546 (1932).

arrive at Brillouin's pattern in a general case is to use the reduced-zone scheme, in which $E_{\mathbf{k}}$ is determined as a many-valued function in the first Brillouin zone, and to cut and distribute the energy surfaces among the outer zones. The distribution must take place in such a way that an energy value associated with the point \mathbf{k} in the inner zone goes to a point $\mathbf{k} + \mathbf{K}$ in an outer one. If the single-valued representation is obtained by another method, the zone structure usually depends upon the method employed. Certain zone boundaries, which may be established by an investigation of symmetry degeneracy, always appear in cases of high symmetry. Other boundaries are not unique and may be altered by choice. There will be certain fixed points through which the zones must pass, however, even in the case of lowest symmetry. These points are determined by the relation

$$E_{\mathbf{k}} = E_{-\mathbf{k}}, \quad (57)$$

which expresses the natural symmetry of the Schrödinger equation. Thus, the zone lines must pass through the points

$$\mathbf{k} = \frac{\mathbf{K}}{2}. \quad (58)$$

The natural advantages of the Brillouin zone scheme are (1) that it determines zones by the Laue conditions, which have significance in the theory of X-ray diffraction, and (2) that it preserves a correspondence between the \mathbf{k} vectors for perfectly free electrons and those in the lattice. If the wave functions for electrons in the crystal are labeled according to Brillouin's scheme and if the crystalline potential field is adiabatically decreased to zero, the wave-number values will be preserved and will agree with those for the resulting free electrons.

In spite of these advantages, the Brillouin scheme is not always the simplest one to use, particularly when one is dealing with cases in which primitive translations are not equal or are not orthogonal to one another. Let us consider a case in which the primitive translations are orthogonal but not equal. The lines defined by Eq. (58), which are shown in Fig. 12, do not form zones that are as simple as those of Fig. 8. It is easy, however, to determine a set that is almost as simple. Such a set, for example, is shown in Fig. 13. These lines are parallel to the fundamental \mathbf{K} vectors instead of orthogonal to them, as Brillouin's zone boundaries are. The relative simplicity of this scheme is illustrated in an even more striking manner in an oblique case. It is still true, of course, that both methods of construction may be brought into coincidence by appropriately cutting and rearranging zones.

Case d. The Three-dimensional Schrödinger Equation.—The three-dimensional case is a straightforward generalization of the two-dimen-

sional one. All the remarks in case *c* may be extended at once to cover the case in which \mathbf{k} is a three-dimensional vector. If we introduce the Born-von Kármán boundary conditions, \mathbf{k} takes on the discrete, though dense, values defined by the equations

$$N_i \mathbf{k} \cdot \boldsymbol{\tau}_i = n_i \quad (n_i = 0, \pm 1, \pm 2, \pm \dots), \quad (59)$$

which are identical with Eqs. (8), Sec. 22, Chap. III.¹ Here the n_i are arbitrary integers, the $\boldsymbol{\tau}_i$ ($i = 1, 2, 3$) are the three primitive translations

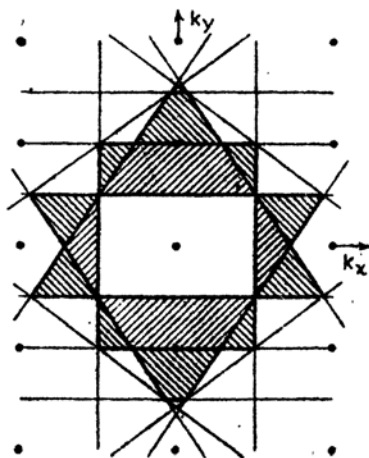


FIG. 12.—Brillouin zone scheme for a two-dimensional rectangular lattice. The zone boundaries, which are defined by Eq. (27), are more complex than those of Fig. 8 because they are orthogonal to the \mathbf{K} vectors.

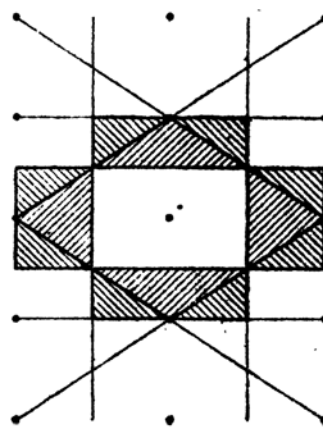


FIG. 13.—Alternative to Fig. 12 in which the zone boundaries are drawn parallel to the \mathbf{K} vectors. This simplifies the pattern.

of the lattice, and the N_i are the number of cells that extend along an axis of the crystal parallel to $\boldsymbol{\tau}_i$.

The three-dimensional Brillouin zones are determined by planes that satisfy the equations

$$\mathbf{k} \cdot \frac{\mathbf{K}}{2} = \frac{\mathbf{K}^2}{4}, \quad (60)$$

¹ It may be noted at this point for future reference that the density of points in \mathbf{k} space, as determined from the solutions

$$\mathbf{k} = \frac{n_1}{N_1} \frac{\boldsymbol{\tau}_2 \times \boldsymbol{\tau}_3}{|\boldsymbol{\tau}_1 \boldsymbol{\tau}_2 \boldsymbol{\tau}_3|} + \frac{n_2}{N_2} \frac{\boldsymbol{\tau}_3 \times \boldsymbol{\tau}_1}{|\boldsymbol{\tau}_1 \boldsymbol{\tau}_2 \boldsymbol{\tau}_3|} + \frac{n_3}{N_3} \frac{\boldsymbol{\tau}_1 \times \boldsymbol{\tau}_2}{|\boldsymbol{\tau}_1 \boldsymbol{\tau}_2 \boldsymbol{\tau}_3|}$$

of Eq. (34), is V , where V is the volume of crystal. This may be shown by computing the reciprocal of the volume of the unit cell of the reciprocal lattice. If spin degeneracy is included, the density of states is $2V$.

which are analogous to (48). The vectors \mathbf{K} are defined by the relations

$$\mathbf{K} \cdot \mathbf{r}_i = l_i \quad (61)$$

where the l are arbitrary integers. These equations, which generalize (38), are special cases of (59) and their solutions [cf. Eq. (8a), Sec. 22, Chap. III] are:

$$\mathbf{K} = l_1 \frac{\mathbf{r}_2 \times \mathbf{r}_3}{|\mathbf{r}_1 \mathbf{r}_2 \mathbf{r}_3|} + l_2 \frac{\mathbf{r}_3 \times \mathbf{r}_1}{|\mathbf{r}_1 \mathbf{r}_2 \mathbf{r}_3|} + l_3 \frac{\mathbf{r}_1 \times \mathbf{r}_2}{|\mathbf{r}_1 \mathbf{r}_2 \mathbf{r}_3|}. \quad (62)$$

The three-dimensional eigenfunctions always have the form

$$\psi_{\mathbf{k}} = \chi_{\mathbf{k}} e^{2\pi i \mathbf{k} \cdot \mathbf{r}},$$

where $\chi_{\mathbf{k}}$ has the translational periodicity of the lattice. This fact may be proved either by perturbation theory, as was done in case *c*, or by use of group theory.

We may obtain zone patterns that are different from Brillouin's by basing the determination of zones on some scheme that does not start from free-electron waves. Some of the simpler zones determined by other methods are the same as those obtained by use of (61) in cases of high symmetry. As before, the differences in zone pattern are superficial, since any pattern may be made to coincide with the Brillouin pattern by appropriately rearranging points in \mathbf{k} space. We shall discuss further details of the three-dimensional case in the next section.

Although the discussion of *c* and *d* has been restricted to the Schrödinger equation, the conclusions are valid for the solutions of any eigenvalue problem in which there is translational symmetry. The problem of determining the normal modes of vibration of a crystal is a case of this type. We have seen in Chap. III that zone theory plays an important role in the classification of solutions of this problem.

62. Survey of Rules and Principles Concerning Three-dimensional Zones.—In this section, we shall amplify the remarks of part *c* of the previous section by tabulating rules for constructing three-dimensional zones. Some of these rules were discussed in Sec. 22; others may be derived by applying the principles introduced there. The rules are as follows:

a. All zones have equal volume in wave-number space. This rule follows from the fact that all zones may be mapped in one another.

b. It is possible to neglect all zones except the first if $\epsilon(\mathbf{k})$ is regarded as a multiple-valued function in this zone. This reduced-zone scheme is most useful when one is determining surfaces by a direct solution of

¹ We shall usually designate the eigenvalues of the three-dimensional periodic one-electron functions by $\epsilon(\mathbf{k})$.

Fock's equations, for then one may derive the energy bands without becoming involved in many of the geometrical complexities of zone structure.

c. Each zone contains $2N$ states, where N is the number of cells in the lattice. The factor 2 arises from the two possible orientations of spin, and the factor N from the fact that N values of \mathbf{k} are associated with each zone. It follows that the maximum number of electrons that may occupy any zone is $2N$.

d. There is a close correspondence between the laws for determining Brillouin zones and those for determining X-ray diffraction. This correspondence is indicated in part by the fact that the equations of Brillouin zone boundaries are the same as Laue's equations for X-ray diffraction; however, the correlation may be much closer. Mott and Jones¹ have pointed out, for example, that in any monatomic solid in which the electronic potential may be approximated by a sum $\sum_s V_s$

of contributions V_s from each atom, the energy discontinuities are small for values of wave numbers for which the X-ray structure factor is small and are large when the structure factor is large. This type of correlation usually does not occur in polyatomic solids because the scattering powers of atoms for low-energy electrons and for X rays are not necessarily proportional for different atoms.

e. Lattices that have the same type of translational symmetry have equivalent zone patterns since zone structure is determined by the \mathbf{K} vectors and these are determined by the primitive translation vectors. The magnitude of the gaps, which is determined by the distribution of potential in the unit cell, however, may be entirely different for crystals that have the same zone structure. Hence, the zone boundaries at which the largest gaps occur may be completely different in translationally similar crystals. This fact is analogous to the fact that the intensities of X-ray diffraction spots may be very different in crystals that have the same translational symmetry.

f. If Brillouin's zones are used to describe the states of all electrons in the solid, it may be convenient to regard the K -shell electrons as filling the first set, the L -shell electrons the next, etc. The valence electrons then occupy zones at the outer fringe of the filled region. It is usually more convenient, however, to disregard inner-shell electrons when one is discussing properties of the solid that do not involve them explicitly. The valence electrons then may be regarded as occupying the central zones.

¹ N. F. MOTT and H. JONES. *The Theory of Metals and Alloys*, Chap. V (Oxford University Press, New York, 1936).

g. Two neighboring zones may have overlapping energy levels even though the energy associated with the outer zone is higher than that associated with the inner at any particular point of the boundary. An example of a case in which overlapping occurs is shown in Fig. 14.

h. All occupied zones cannot be completely filled in a substance that has an odd number of electrons per unit cell. Hence, occupied and unoccupied levels are immediately adjacent in a substance of this type. Since this is the condition for metallic conductivity, these substances are

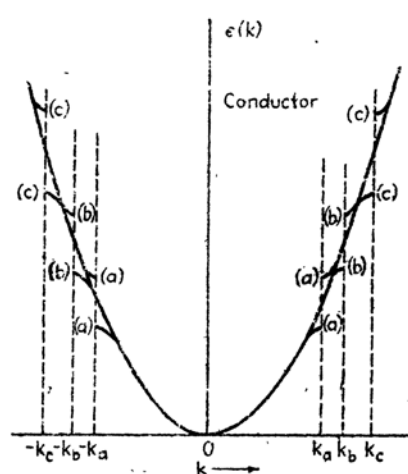


FIG. 14.

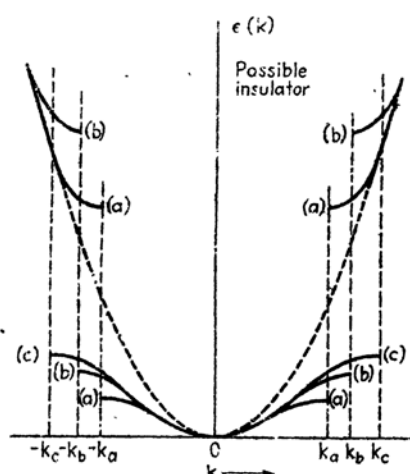


FIG. 15.

FIG. 14.— $\epsilon(k)$ curves for the case in which two zones overlap. This figure gives superposition of energy curves for values of k on lines that pass through the origin of k space and extend in three prominent directions. The intercepts of these lines with the zone boundary occur, respectively, at k_a , k_b , and k_c . The highest point at k_a , which belongs to the second zone, lies below the lowest point of k_b , which belongs to the first zone, etc. (compare with Fig. 15).

FIG. 15.— $\epsilon(k)$ curves for the case in which two zones do not overlap. This figure is the same as Fig. 14 except that the gaps are so large that the uppermost levels of the lower band are always below the lowest levels of the upper band. A substance having $\epsilon(k)$ curves of this type is an insulator if there is an even number of electrons.

metals (*cf.* Sec. 60). If the substance has an even number of electrons per unit cell, the type of conductivity depends upon the nature of the gaps at the boundary of the filled region. Suppose that there are $2m$ electrons per unit cell. The m th zone is not completely filled if the m th and the $m + 1$ st zones overlap, for then some electrons prefer the lowest levels of the $m + 1$ st zone to the highest levels of the m th zone. Thus, the substance should be a metal in this case. On the other hand, it should be an insulator if they do not overlap. Cases corresponding to both these types are shown in Figs. 14 and 15.

We shall not discuss those properties of zones and energy surfaces which may be derived by application of group theory, for to do so would

carry us too far afield. The development of this topic may be found in the references of footnote 1, page 275.

63. Examples of Zone Structure.—For illustrative purposes, we shall describe in this section the simplest zones of several lattices. The precise description of zones is important, at present, only for a few simple crystals, such as the monovalent and divalent metals and the alkali halides. More complex solids can be handled only in rough approximation. Thus, for γ brass all that is important is to know the shape of one or two zones near the limit of the filled region and whether the gaps at the boundaries of these zones are large or small. We shall discuss these cases as we need them in later sections.

a. Face-centered Cubic Lattice.—The primitive translations of the face-centered cubic lattice may be chosen as

$$\tau_1 = \begin{pmatrix} a \\ a \\ 0 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} a \\ 0 \\ a \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} 0 \\ a \\ a \end{pmatrix}. \quad (1)$$

The components of these vectors are expressed in Cartesian coordinates. Solving the equations

$$\mathbf{k} \cdot N_i \tau_i = n_i \quad (i = 1, 2, 3), \quad (2)$$

which define the values of \mathbf{k} , we find

$$\mathbf{k} = \frac{n_1}{2N_1a} \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} + \frac{n_2}{2N_2a} \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} + \frac{n_3}{2N_3a} \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}. \quad (3)$$

Hence, the reciprocal lattice is body-centered.

The \mathbf{K} vectors are given by those values of (3) for which n_1 , n_2 , and n_3 are integer multiples of N_1 , N_2 , and N_3 , respectively. The first three sets of these vectors are:

$$\frac{1}{2a} \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}, \quad \frac{1}{2a} \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}, \quad \frac{1}{2a} \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \quad \frac{1}{2a} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \quad (I)$$

$$\frac{1}{2a} \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \quad \frac{1}{2a} \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}, \quad \frac{1}{2a} \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}. \quad (II)$$

$$\left. \begin{aligned} &\frac{1}{2a} \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}, \quad \frac{1}{2a} \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}, \quad \frac{1}{2a} \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix} \\ &\frac{1}{2a} \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}, \quad \frac{1}{2a} \begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix}, \quad \frac{1}{2a} \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} \end{aligned} \right\} \quad (III)$$

The first two zones are shown in Figs. 16a and b and are bounded by planes orthogonal to (I) and (II). The higher zones are more complicated and involve other \mathbf{K} vectors. It may be seen that the illustrated zones are uniquely specified by symmetry conditions.

b. *Body-centered Lattice*.—The primitive translations in this case are

$$\tau_1 = \begin{pmatrix} a \\ a \\ a \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} -a \\ a \\ a \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} -a \\ -a \\ a \end{pmatrix}, \quad (4)$$

whence

$$\mathbf{k}_1 = \frac{n_1}{2N_1a} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \frac{n_2}{2N_2a} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \frac{n_3}{2N_3a} \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}. \quad (5)$$

In other words, the inverse lattice in this case is face-centered.

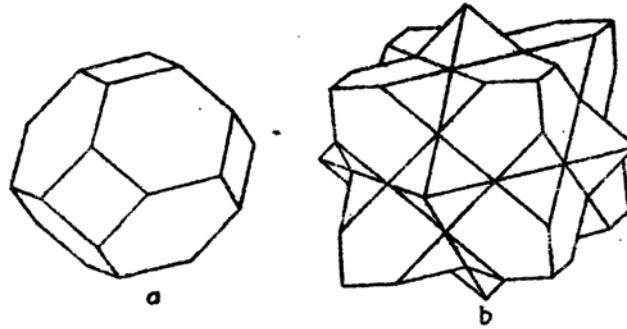


FIG. 16.—The first and second zones for a face-centered cubic lattice. The first has half the volume of the cube that is determined by extending the six square faces. The second has the same volume as this cube.

The first two sets of \mathbf{K} vectors are:

$$\frac{1}{2a} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \frac{1}{2a} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \frac{1}{2a} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \frac{1}{2a} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \frac{1}{2a} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \frac{1}{2a} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \quad (I)$$

$$\frac{1}{2a} \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \quad \frac{1}{2a} \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}, \quad \frac{1}{2a} \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}. \quad (II)$$

The first two zones appear in Fig. 17.

c. *Close-packed Hexagonal Lattice*.—The primitive translations of the close-packed hexagonal lattice are

$$\tau_1 = \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 0 \\ b/2 \\ \sqrt{3}b/2 \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} 0 \\ -b/2 \\ \sqrt{3}b/2 \end{pmatrix}, \quad (6)$$

for which

$$\mathbf{k}_1 = \frac{n_1}{N_1 a} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \frac{2n_2}{N_2 b \sqrt{3}} \begin{pmatrix} 0 \\ \sqrt{3}/2 \\ 1/2 \end{pmatrix} + \frac{2n_3}{N_3 b \sqrt{3}} \begin{pmatrix} 0 \\ -\sqrt{3}/2 \\ 1/2 \end{pmatrix}. \quad (7)$$

Thus, this lattice is its own reciprocal. The first zone is the hexagonal

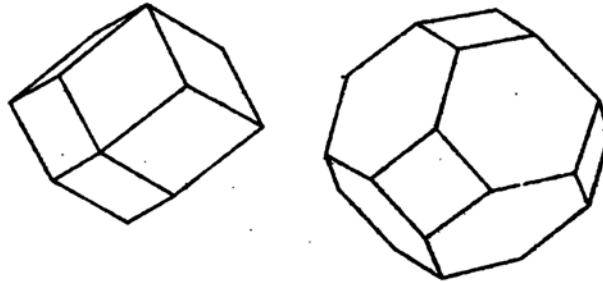


FIG. 17.—The first two zones for the body-centered cubic lattice. The surfaces of the first are normal to \mathbf{K} vectors (I) of the text, whereas the surfaces of the second are normal to \mathbf{K} vectors (II) and \mathbf{K} vectors that lie in the (111) direction.

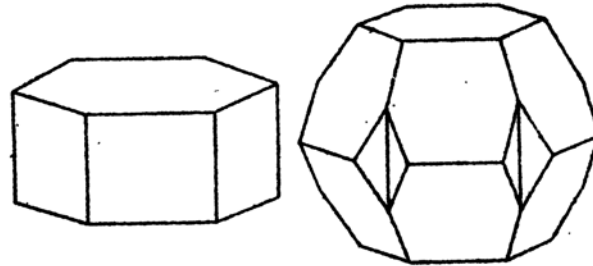



FIG. 18.—The first two Brillouin-type zones for a close-packed hexagonal crystal. The second zone is not uniquely defined by symmetry and may be drawn in many different ways.

prism shown in Fig. 18, which is determined by the following two sets of \mathbf{K} vectors:

$$\frac{1}{a} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \quad (I)$$

$$\frac{2}{\sqrt{3}b} \begin{pmatrix} 0 \\ \sqrt{3}/2 \\ 1/2 \end{pmatrix}, \quad \frac{2}{\sqrt{3}b} \begin{pmatrix} 0 \\ -\sqrt{3}/2 \\ 1/2 \end{pmatrix}, \quad \frac{2}{\sqrt{3}b} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (II)$$

The form of the second zone, which is determined by the sets (I) and (II) and the set



$$\begin{array}{ccc} \frac{2}{\sqrt{3b}} \begin{pmatrix} \sqrt{3b/2a} \\ \sqrt{3/2} \\ 1/2 \end{pmatrix}, & \frac{2}{\sqrt{3b}} \begin{pmatrix} \sqrt{3b/2a} \\ -\sqrt{3/2} \\ -1/2 \end{pmatrix}, & \frac{2}{\sqrt{3b}} \begin{pmatrix} \sqrt{3b/2a} \\ -\sqrt{3/2} \\ 1/2 \end{pmatrix}, \\ \frac{2}{\sqrt{3b}} \begin{pmatrix} \sqrt{3b/2a} \\ \sqrt{3/2} \\ -1/2 \end{pmatrix}, & \frac{2}{\sqrt{3b}} \begin{pmatrix} \sqrt{3b/2a} \\ 0 \\ 1 \end{pmatrix}, & \frac{2}{\sqrt{3b}} \begin{pmatrix} \sqrt{3b/2a} \\ 0 \\ -1 \end{pmatrix}, \end{array}$$

depends upon the axial ratio $\sqrt{3b/2a}$. The form for an ideally close-packed lattice is illustrated in Fig. 18.

64. Cases of Coalescence of the Heitler-London and Bloch Schemes*

We saw in Sec. 58, Chap. VII, that the Heitler-London and Hund-Mulliken schemes are equivalent when applied to the state of a molecule whose constituent atoms or ions have only closed-shell configurations. In this case, the total wave function is a single determinant in both the one-electron schemes, and it is possible to transform one determinant into the other by adding to each row an appropriate linear combination of other rows. A similar theorem is valid for solids. The Heitler-London and Bloch schemes are identical whenever the Heitler-London scheme leads to completely filled shells. To prove this, all that is necessary is to show that a complete zone of N Bloch functions may be constructed by taking linear combinations of the N Heitler-London functions associated with translationally equivalent atoms in the N unit cells of the lattice, for each of these Heitler-London functions appears once with each spin in the determinantal eigenfunction.

Let $\psi(\mathbf{r} - \mathbf{r}(n))$ be the closed-shell function that is centered about the point $\mathbf{r}(n)$ in the n th unit cell. The functions formed from these by taking the linear combinations

$$\psi_{\mathbf{k}} = \sum_n e^{-2\pi i \mathbf{k} \cdot \mathbf{r}(n)} \psi(\mathbf{r} - \mathbf{r}(n)), \quad (1)$$

where $\mathbf{r}(n)$ is summed over all unit cells in the lattice and \mathbf{k} is a vector in the reciprocal lattice of the crystal, are Bloch functions. This may be seen¹ by changing \mathbf{r} to $\mathbf{r} - \boldsymbol{\tau}$ where $\boldsymbol{\tau}$ is any primitive translation of the lattice, for then (1) becomes

$$\sum_{\mathbf{r}(n) + \boldsymbol{\tau}} e^{2\pi i \mathbf{k} \cdot \boldsymbol{\tau}} e^{-2\pi i \mathbf{k} \cdot [\mathbf{r}(n) + \boldsymbol{\tau}]} \psi(\mathbf{r} - (\mathbf{r}(n) + \boldsymbol{\tau})) = e^{2\pi i \mathbf{k} \cdot \boldsymbol{\tau}} \psi_{\mathbf{k}},$$

since in an ideal unbounded lattice the summation over $\mathbf{r}(n) + \boldsymbol{\tau}$ is equivalent to a summation over $\mathbf{r}(n)$. Thus, the function in (1) is multiplied by a factor $e^{2\pi i \mathbf{k} \cdot \boldsymbol{\tau}}$ when the crystal is translated by an amount

¹ F. BLOCH, *Z. Physik*, **52**, 555 (1928), pointed out first that the sum (1) is the combination of atomic functions which satisfies the periodic boundary conditions.

N independent functions of type (1) may be constructed from the $\psi(r - r(n))$, and this number is just sufficient to fill a zone.

The converse would be a theorem stating that the Heitler-London and Bloch schemes are equivalent whenever the Bloch scheme leads to a set of completely filled zones. It does not seem to be possible to prove this theorem generally. It is probably valid, however, in many special cases. Consider the N functions

$$\psi(r - r(n)) = \sum_{\mathbf{k}} \chi_{\mathbf{k}} e^{-2\pi i \mathbf{k} \cdot (r - r(n))} \quad (2)$$

where $\chi_{\mathbf{k}} e^{2\pi i \mathbf{k} \cdot r}$ is a Bloch function, \mathbf{k} is summed over all the N values in a single zone, and $r(n)$ ranges over N translationally equivalent positions in the N cells of the lattice. The function (2) is localized in the n th unit cell if $\chi_{\mathbf{k}}$ varies sufficiently slowly with \mathbf{k} ; hence, in this case, it is a function of the Heitler-London type. Suppose, for simplicity, that $\chi_{\mathbf{k}}$ is independent of \mathbf{k} . Then the function (2) degenerates to

$$\chi(r) \sum_{\mathbf{k}} e^{-2\pi i \mathbf{k} \cdot (r - r(n))} \quad (3)$$

The phases of the exponent in the summands of this function range roughly between the values $\pm 2\pi |r - r(n)| k(\text{max.})$, where $k(\text{max.})$ is the distance from $k = 0$ to the farthest point in the zone. Hence, the sum in (3) has its maximum value when r is in the n th unit cell and is very small, when $|r - r(n)|$ is large. Although this property is affected if χ depends upon \mathbf{k} , we may expect (2) to be localized as long as the dependence is not strong.

Thus, in some cases, we may transform a determinantal wave function that is based upon a filled zone of Bloch functions into a determinant of Heitler-London functions that are localized about any set of points in the unit cell we choose. It may happen, however, that the Heitler-London wave functions which are obtained in this way do not have desirable symmetry characteristics and that a better set can be found.

Before leaving this topic, we shall discuss the connection between the energy parameters in Fock's equations for the Bloch scheme and the Heitler-London scheme when the two are equivalent. For the Bloch scheme, the equations may be chosen in the form

$$H\psi_{\mathbf{k}} = \epsilon(\mathbf{k})\psi_{\mathbf{k}} \quad (4)$$

since the $\psi_{\mathbf{k}}$ are automatically orthogonal (cf. Sec. 51). The corresponding equations for the Heitler-London functions are

$$H\psi_n = \epsilon_0 \psi_n + \sum_m' \alpha_{nm} \psi_m \quad (5)$$

in which

$$\psi_n = \psi(\mathbf{r} - \mathbf{r}(n))$$

is localized about the atom at the position $\mathbf{r}(n)$ and the α_{nm} are normalization parameters that must be introduced because the ψ_n form a highly degenerate set of functions not automatically orthogonal.

The ψ_k and ψ_n are connected by the equations

$$\psi_k = \frac{1}{\sqrt{N}} \sum_m e^{2\pi i \mathbf{k} \cdot \mathbf{r}(m)} \psi_m. \quad (6)$$

It may be verified that the operators H in Eqs. (4) and (5) are the same because of Eqs. (6). The relation between $\epsilon(\mathbf{k})$ and ϵ_0 may be found by multiplying Eq. (5) by $e^{2\pi i \mathbf{k} \cdot \mathbf{r}(n)}$, summing over n , and subtracting Eq. (4). The result is

$$(\epsilon_0 - \epsilon(\mathbf{k}))\psi_k + \frac{1}{\sqrt{N}} \sum_{n,m}' e^{2\pi i \mathbf{k} \cdot \mathbf{r}(n)} \alpha_{nm} \psi_m = 0. \quad (7)$$

If we multiply this by ψ_k^* and integrate, we find

$$\epsilon(\mathbf{k}) = \epsilon_0 + \frac{1}{N} \sum_{n,m}' \alpha_{nm} e^{2\pi i \mathbf{k} \cdot [\mathbf{r}(n) - \mathbf{r}(m)]}. \quad (8)$$

We shall discuss relationships of this type in more detail in the next section.

65. Approximate Bloch Functions for the Case of Narrow Bands*.

It is possible to use Eq. (6) of the preceding section to obtain Bloch type functions from atomic functions even when the Heitler-London and Bloch schemes are not equivalent; indeed, these functions are sometimes very useful for semiquantitative discussions. The relationship between the energies in this case may be derived in the following way: We shall assume, for simplicity, that we have a monatomic lattice of atoms that possess one valence electron each and that the Schrödinger equation for the wave function $\psi(\mathbf{r} - \mathbf{r}(n))$ of the electron in the n th atom is

$$\left\{ -\frac{\hbar^2}{2m} \Delta + V(\mathbf{r} - \mathbf{r}(n)) \right\} \psi(\mathbf{r} - \mathbf{r}(n)) = \epsilon \psi(\mathbf{r} - \mathbf{r}(n)) \quad (1)$$

when the atom is free. The function $V(\mathbf{r} - \mathbf{r}(n))$ is the electronic potential, which we may expect to be negative. When the atoms are brought together, the potential at the n th atom is changed by the addition of the term

$$\sum_m' V(\mathbf{r} - \mathbf{r}(m))$$

in which m is summed over all atoms except the n th. In order to determine¹ the energy of the Bloch wave function

$$\psi_{\mathbf{k}} = \frac{1}{\sqrt{N}} \sum_n e^{2\pi i \mathbf{k} \cdot \mathbf{r}(n)} \psi(\mathbf{r} - \mathbf{r}(n)), \quad (2)$$

it is necessary to evaluate the integral

$$\begin{aligned} \epsilon(\mathbf{k}) &= \int \psi_{\mathbf{k}}^* \left\{ -\frac{\hbar^2}{2m} \Delta + \sum_m V(\mathbf{r} - \mathbf{r}(m)) \right\} \psi_{\mathbf{k}} d\tau \\ &= \frac{1}{N} \left(\sum_{n,m} e^{2\pi i \mathbf{k} \cdot [\mathbf{r}(n) - \mathbf{r}(m)]} \int \psi^*(\mathbf{r} - \mathbf{r}(m)) \left\{ -\frac{\hbar^2}{2m} \Delta + \right. \right. \\ &\quad \left. \left. V(\mathbf{r} - \mathbf{r}(n)) \right\} \psi(\mathbf{r} - \mathbf{r}(n)) d\tau + \right. \\ &\quad \left. \sum_{n,m'} e^{2\pi i \mathbf{k} \cdot [\mathbf{r}(n) - \mathbf{r}(m)]} \int \psi^*(\mathbf{r} - \mathbf{r}(m)) \sum_{l \neq n} V(\mathbf{r} - \mathbf{r}(l)) \psi(\mathbf{r} - \mathbf{r}(n)) d\tau \right) \quad (3) \end{aligned}$$

We shall assume that the atomic functions satisfy the relations

$$\int \psi^*(\mathbf{r} - \mathbf{r}(m)) \psi(\mathbf{r} - \mathbf{r}(n)) d\tau = \delta_{m,n}$$

and that

$$\int \psi^*(\mathbf{r} - \mathbf{r}(m)) V(\mathbf{r} - \mathbf{r}(l)) \psi(\mathbf{r} - \mathbf{r}(n)) d\tau$$

is zero unless one of the conditions

$$n = l, \quad n = m, \quad l = m$$

is satisfied. Equation (3) may then be simplified to

$$\epsilon(\mathbf{k}) = \epsilon' + \frac{1}{N} \sum_{n,m} \alpha'_{nm} e^{2\pi i \mathbf{k} \cdot [\mathbf{r}(n) - \mathbf{r}(m)]} \quad (4)$$

where

$$\epsilon' = \epsilon + \sum_l \int |\psi(\mathbf{r})|^2 V(\mathbf{r} - \mathbf{r}(l)) d\tau \quad (5)$$

and

$$\alpha'_{nm} = \int \psi^*(\mathbf{r} - \mathbf{r}(m)) V(\mathbf{r} - \mathbf{r}(n)) \psi(\mathbf{r} - \mathbf{r}(n)) d\tau. \quad (6)$$

Both α'_{nm} in this equation and α_{nm} in Eq. (8) of the preceding section are larger when the atoms overlap a great deal than when they are far apart; moreover, in the limit as the atoms become free, (4) approaches the energy ϵ' of the electron in the free atom for all values of \mathbf{k} . We may

¹ *Ibid.*

conclude that in the solid there is a band of Bloch levels corresponding to one zone of \mathbf{k} space for each electronic state of the free atom (cf. Fig. 19).

Since the lattice has translational symmetry, it follows that α'_{n+l} is independent of n . Thus, Eq. (4) may be written in the form

$$\epsilon(\mathbf{k}) = \epsilon' + \sum_l \alpha'_l e^{2\pi i \mathbf{k} \cdot \mathbf{r}(l)} \quad (7)$$

in which

$$\alpha'_l = \alpha'_{n+l,n} \quad (8)$$

and the summation extends over all values of l . Equation (8) of the preceding section evidently can be placed in the same form. Now, atomic functions may always be chosen as real if there is no magnetic field present. Thus, the α'_l may be regarded as real, and Eq. (7) may be placed in the form

$$\begin{aligned} \epsilon(\mathbf{k}) &= \epsilon' + \sum_l \alpha'_l \cos 2\pi \mathbf{k} \cdot \mathbf{r}(l) \\ &= \epsilon' + \sum_l \alpha'_l - 2 \sum_l \alpha'_l \sin^2 \pi \mathbf{k} \cdot \mathbf{r}(l). \end{aligned} \quad (9)$$

Only the α_l corresponding to very near neighbors are important when the bands are comparatively narrow. Hence, only the first few terms in the series (8) need be retained in this case. The α corresponding to immediate neighbors are equal in the three cubic lattices and in the ideally close-packed hexagonal crystal when the atomic functions are s type. Hence, in this case the nearest-neighbor terms in (9) may be written as

$$\epsilon(\mathbf{k}) = \epsilon'' - 2\alpha \sum_{\mathbf{r}_n} \sin^2 \pi \mathbf{k} \cdot \mathbf{r}_n \quad (10)$$

where \mathbf{r}_n ranges over the vectors connecting an atom with its nearest neighbors and

$$\epsilon'' = \epsilon' + z\alpha$$

in which z is the number of nearest neighbors. Now, if $V(\mathbf{r} - \mathbf{r}(n))$ in the integral (6) is negative, as the electronic potential of an atom

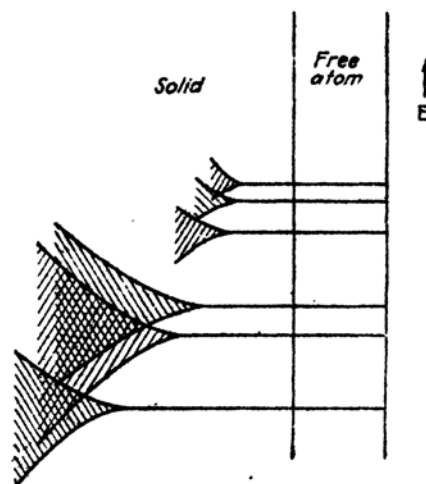


FIG. 19.—The one-electron levels of the free atom split into bands when the atom enters the solid.

should be, and if the portions of the ψ that overlap have the same sign, as we may expect for well-separated s functions, the α in (9) are negative. Thus, we may conclude that before appreciable overlapping occurs the $\epsilon(\mathbf{k})$ curves for bands that arise from s electrons increase with increasing values of $|\mathbf{k}|$ in any direction starting from the origin of \mathbf{k} space (cf.

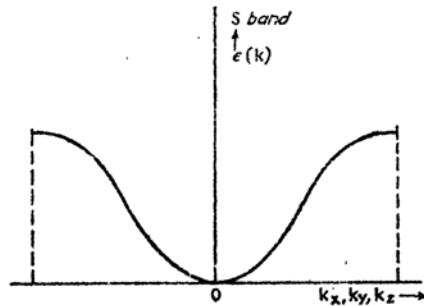


FIG. 20.—The $\epsilon(\mathbf{k})$ curve for an s electron band is concave upwards for all directions in \mathbf{k} space.

Fig. 20).

The same conclusion cannot be drawn in cases in which the atomic functions are not s type. Suppose, for example, that the atomic ψ are p functions of the form

$$\frac{x}{r}f(r). \quad (11)$$

Then the overlapping portions of the ψ of atoms that have different x coordinates have opposite signs, and their α are positive (cf. Fig. 21).

Moreover, although the overlapping portions for atoms that lie in the same x plane have the same sign, the product is small because (11) has a node in this plane. Thus, in this case the function $\epsilon(\mathbf{k})$ should decrease with increasing values of k in the x direction of \mathbf{k} space and should increase in the y and z directions (cf. Fig. 22). Cases in which the functions are higher than s or p type must be considered individually.



FIG. 21.—The overlapping parts of p functions of type (11) that have different x coordinates have opposite sign.

It is readily found that, in monatomic crystals having the three simplest cubic structures, Eq. (10) reduces to

$$\epsilon(\mathbf{k}) \cong \epsilon'' - 4\pi^2 a a^2 k^2 \quad (12)$$

in the neighborhood of the origin of \mathbf{k} space, where a is the edge length of the fundamental cube. This may be made identical with the expression for effectively free electrons, namely,

$$\epsilon(\mathbf{k}) = \frac{\hbar^2}{2m^*} \mathbf{k}^2,$$

by setting

$$\frac{1}{m^*} = -\frac{8\pi^2}{\hbar^2} \alpha a^2.$$

Since α is negative for s electrons, they have a positive effective mass. On the other hand, it is clear from what has been said in the preceding paragraph that the effective mass of electrons near the origin of a p band is negative for motion in the x direction. The significance of negative masses will be discussed in Sec. 68.

Mott and Jones¹ have used $\epsilon(\mathbf{k})$ functions of type (10) to compute the function $dN/d\epsilon$, which gives the number of states per unit energy range.² Since the density of points in \mathbf{k} space is uniform and is equal³ to $2V$, it follows that

$$\delta N = 2V d\tau_{\mathbf{k}},$$

where $d\tau_{\mathbf{k}}$ is the differential of volume in \mathbf{k} space and δN is the number of states in this volume. Now, if we choose coordinates so that the differential of volume is bounded on two faces by surfaces of constant energy,

$$d\tau_{\mathbf{k}} = dS \left(\frac{dk}{d\epsilon} \right) d\epsilon,$$

where dS is the differential area of the surface of constant energy, $d\epsilon$ is the energy difference of the two bounding surfaces, and $dk/d\epsilon$ is the change in k in going from one energy surface to the other. Evidently,

$$\left(\frac{dk}{d\epsilon} \right) = \frac{1}{|\text{grad}_{\mathbf{k}} \epsilon(\mathbf{k})|}, \quad (13)$$

whence

$$n(\epsilon') = \left(\frac{dN}{d\epsilon} \right)_{\epsilon=\epsilon'} = \iint \frac{2V dS}{|\text{grad}_{\mathbf{k}} \epsilon(\mathbf{k})|} \quad (14)$$

in which the integral extends over the surface of energy ϵ' .

¹ Cf. Mott and Jones, *op. cit.*

² In the following sections of this book, we shall usually designate $dN/d\epsilon$ by $n(\epsilon)$.

³ See footnote 1, p. 294.

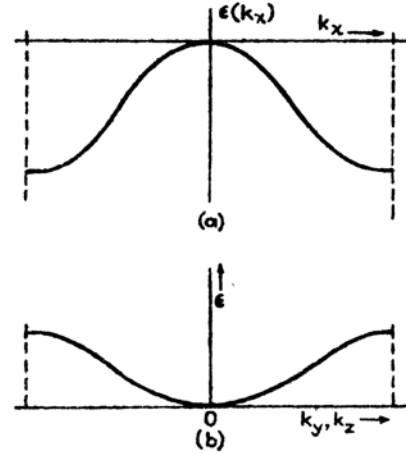


FIG. 22.— $\epsilon(\mathbf{k})$ curves for a p band. The $\epsilon(k_x)$ curve is concave downward, whereas the $\epsilon(k_y)$ and $\epsilon(k_z)$ curves are concave upward. The curvature is less in the second case because the overlapping is smaller.

Figure 23 gives a comparison of the $n(\epsilon)$ curves for the first zone of a body-centered lattice that are obtained with the assumptions that $\epsilon(\mathbf{k})$ is given by Eq. (10) and by the free-electron equation

$$\epsilon(\mathbf{k}) = \frac{h^2}{2m^*} k^2,$$

respectively. In the second case, the function varies as $\sqrt{\epsilon}$ until the spherical contours intersect the zone boundary, whereupon the density of levels decreases.

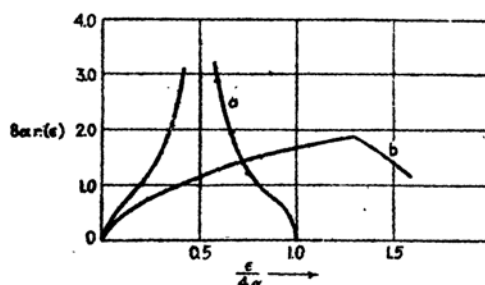


FIG. 23.— $dN/d\epsilon$ curves for a body-centered cubic lattice. Curve b is derived by assuming the electrons are perfectly free; curve a is obtained from an $\epsilon(\mathbf{k})$ relation of the type (10).

66. The Total Electronic Wave Function of the Solid.—Let us digress momentarily from the one-particle approximation and discuss the wave functions of the solid as a whole. To begin with, we shall assume that the constituents are far apart. If the lowest states of these atoms, ions, or molecules are nondegenerate as in rare gas solids, ordinary molecular crystals, or ionic crystals that are referred to a normal state of ions with rare gas configurations, the lowest state of the entire assembly also is nondegenerate. On the other hand, if the constituents are twofold degenerate, as in the alkali metals, the total degeneracy of the system of N separated atoms is 2^N . More generally, the degeneracy of the lowest state of the entire assembly is g^N if the lowest state of the constituent is g -fold degenerate. The first excited state of the nondegenerate assembly is Nf -fold degenerate if the first excited state of the constituents is f -fold degenerate. The factor N appears because any one of the N constituents may be excited.

Let us now bring the atoms or molecules closer together in order to form a solid. We may expect some of the levels of the entire solid that are highly degenerate at infinite separation to split into levels of lower degeneracy as the interatomic distance decreases to the point where the charge distributions of different atoms overlap. This splitting may occur in any one of several ways. (1) The components of the splitting level

may be separated from one another by finite distances (*cf.* Fig. 24*a*). (2) The level may split into a quasi-continuous band (Fig. 24*b*). (3) It may break into separate continuous bands (Fig. 24*c*). (4) It may break into bands and discrete levels (Fig. 24*d*). A nondegenerate level remains nondegenerate, of course, although it may cross or mingle with the progeny of degenerate levels (Figs. 25*a* and 25*b*).

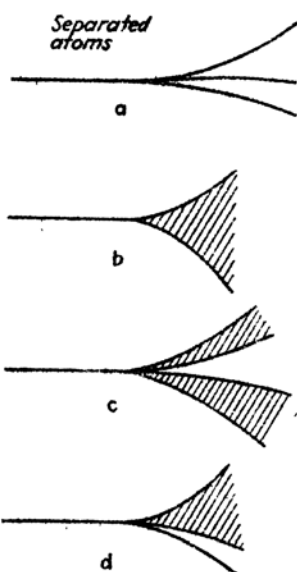


FIG. 24.

FIG. 24.—Possible behavior of a degenerate level of a system of atoms as the atoms are combined to form a solid. In *a* the level splits but remains discrete; in *b* it splits into a quasi-continuous band; in *c* it splits into two bands; in *d* it splits into a band and a discrete level. The discrete levels in *a* and *d* may be highly degenerate.

FIG. 25.—The upper level is highly degenerate whereas the lower is nondegenerate. In case *a* the lower level does not merge with the continuum; in case *b* it does.

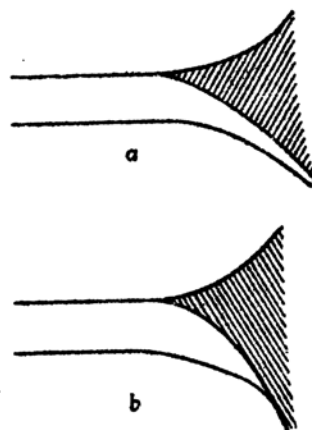


FIG. 25.

symmetry theory that the wave functions of the entire crystal have the form

$$\chi_{\mathbf{k}_1, \dots, \mathbf{k}_n}(\mathbf{r}_1, \dots, \mathbf{r}_n) e^{2\pi i(\sum_i \mathbf{k}_i \cdot \mathbf{r}_i)} \quad (1)$$

Here, \mathbf{r}_i is the position vector of the i th electron; $\chi_{\mathbf{k}_1, \dots, \mathbf{k}_n}$ is a periodic function in the sense that

$$\chi_{\mathbf{k}_1, \dots, \mathbf{k}_n}(\mathbf{r}_1 + \boldsymbol{\tau}, \mathbf{r}_2 + \boldsymbol{\tau}, \dots, \mathbf{r}_n + \boldsymbol{\tau}) = \chi_{\mathbf{k}_1, \dots, \mathbf{k}_n}(\mathbf{r}_1, \dots, \mathbf{r}_n) \quad (2)$$

where $\boldsymbol{\tau}$ is any translational vector of the lattice. If periodic boundary conditions are used, the \mathbf{k}_i satisfy the relations

$$\left(\sum_i \mathbf{k}_i\right) \cdot \mathbf{T}_\alpha = l_\alpha \quad (\alpha = 1, 2, 3),$$

where \mathbf{T}_α is one of the three vectors corresponding to the edge lengths of the crystal and l_α is an integer. The function (1) evidently is a

$3n$ -dimensional generalization of a Bloch function, and $\sum_i \mathbf{k}_i$ is a generalized wave-number vector. For each set of integers l_1, l_2 , and l_3 , there is an independent value of $\sum_i \mathbf{k}_i$. When the atoms are widely separated, we may expect $\chi_{\mathbf{k}_1}, \dots, \chi_{\mathbf{k}_n}$, as in the one-dimensional case, to be a function that has a relatively large amplitude only when the electrons are near the atoms. The modulation factor $e^{2\pi i(\sum_i \mathbf{k}_i \cdot \mathbf{r}_i)}$ then has relatively minor importance; and, as in the case of bound electrons in the Bloch scheme, the degeneracies may be high since many wave functions that have different values of $\sum_i \mathbf{k}_i$ may have the same energy. On the other hand, we may expect χ to have an appreciable amplitude for a much wider range of the \mathbf{r}_i when the atoms are close together. The value of $\sum_i \mathbf{k}_i$ then affects the energy, and the degeneracy is at least partly removed.

If the lowest level is not degenerate and if, as in Fig. 25a, it does not mingle with a continuum, the solid in its normal state is an insulator. In this case, we may regard the effect of an electrostatic field as a perturbation and may express the perturbed state as a linear combination of the unperturbed functions. The amplitude with which the excited states appear in this function is proportional to the field intensity and is small as long as the field is not strong. Hence, we should expect a finite electronic polarizability for small electrostatic fields, just as for ordinary atoms and molecules in which the lowest level is discrete.

Consider next the cases illustrated in Fig. 26 in which there are low-lying continuous bands. The lowest state now has a large number of other states arbitrarily near to it. When this solid is placed in an electrostatic field, it may be possible to construct from the lowest levels of the continuum a stable perturbed wave function which takes maximum advantage of the applied field by neutralizing this field within the solid. This property corresponds to infinite polarizability and is characteristic of metals. A continuum of levels is a necessary but not a sufficient condition for metallic polarizability, as we shall see presently. We may conclude that Fig. 26 is representative of the energy-level diagram of a metal. Figure 26b then corresponds to the case of an alkaline earth metal. The normal state at infinite separation is not degenerate in this case because the atoms have closed-shell (s^2) configurations.

It is not possible to describe a continuum of the type required for metallic behavior by means of the Heitler-London approximation.

Since the Hartree-Fock field for the Heitler-London functions must have a trough in the region where the electrons are localized, just as in the case of atoms, we may expect the lowest energy levels that are associated with this trough to be discrete. Hence, the lowest and first excited levels of the system should be separated by a finite energy.

The Bloch scheme, on the other hand, leads to a continuum of the required type, as we have seen in the preceding sections of this chapter. Thus the Bloch scheme is preferable to the Heitler-London for a purely qualitative description of the properties of metals.

Although the lowest state of insulators may be described equally well by either approximation, the Bloch scheme may be inferior for the lowest excited states. Consider, for example, the following system. We shall start with a set of N infinitely separated atoms having nondegenerate normal states. One excited state of this system is that in which an electron is removed from one atom and is placed on another, a positive and a negative ion thus being formed. We shall assume for simplicity that this is the first excited state. Its degeneracy is $g_p g_n N(N-1)$, where g_p and g_n are the degeneracies of the positive and negative ions, respectively. The factor $N(N-1)$ occurs because an electron may be taken from any one of the N ions and placed on any one of the $N-1$ remaining atoms. Some of this degeneracy disappears as the atoms are brought within a finite distance of one another, because of the interaction between positive and negative ions. The energy of a pair of spherically symmetrical nonoverlapping ions decreases by an amount $-e^2/r$, as the ions approach one another, where r is the distance between their centers. Thus, the degeneracy of the new levels is $n_r g_p g_n N$, where n_r is the number of neighbors at distance r , from a given atom. This remaining degeneracy is partly removed as the atoms overlap, and each level may be expected to split into a band. Suppose that the splitting has reached the point corresponding to the line A, Fig. 27, when the system attains equilibrium relative to interatomic displacements. The first band of excited states then differ from the normal state in that an electron has

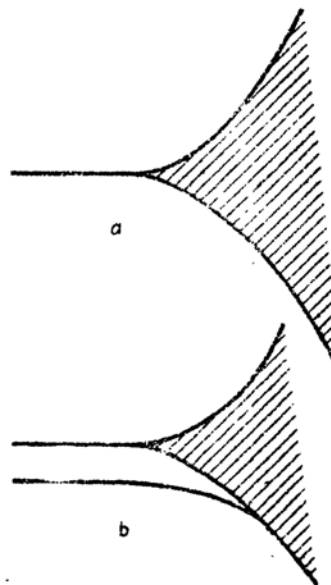


FIG. 26.—Cases in which the lowest levels form a quasi-continuous band. In case *a* there is no nondegenerate level and the low-lying degenerate level breaks into a band. In case *b* the nondegenerate level is absorbed into the continuum of the band arising from an excited level. Case *a* is characteristic of the alkali metals, case *b* of the alkaline earth metals.

been transferred from one atom to an immediate neighbor. Since the negative and positive ions that are produced remain at a fixed distance relative to one another, the crystal does not possess electronic conductivity when in any one of the levels of this band of excited states. It is not possible to describe nonconducting excited states of this type by means of the Bloch approximation. A crystal in which all occupied zones are completely filled is an insulator, but it becomes an electronic conductor as soon as one electron has been removed to an unoccupied zone. Thus, the use of the Bloch approximation for an insulator is equivalent to assuming that a continuum which is similar to the continua of metals lies above the lowest nondegenerate level of the crystal.

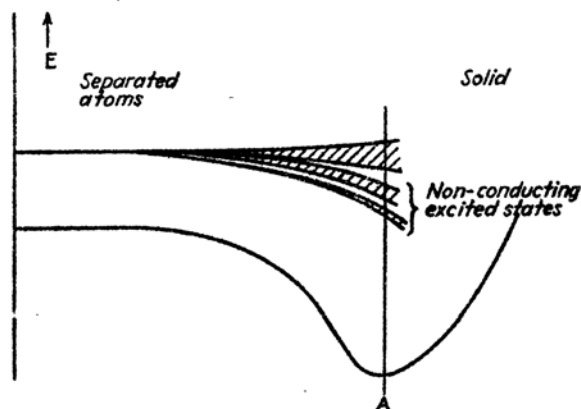


FIG. 27.—Case in which the lowest level is nondegenerate and the excited degenerate level splits into several bands, the lowest of which is nonconducting. The nonconducting states cannot be described by the Bloch approximation.

The Heitler-London approximation is somewhat more suitable than the Bloch scheme for describing the nonconducting excited states. However, it is open to the objection that in using it the assumption must be made that a particular atom is ionized and that the electron is localized on the neighbors of this atom. Since all atoms in the model discussed above are equivalent, the excitation would be described more appropriately if it extended throughout the lattice. In Sec. 96, we shall discuss an approximational scheme that satisfies this condition.

The Bloch scheme is not always invalid when the lowest state of the solid is not degenerate for large separations. Free alkaline-earth metal atoms are normally in a 1S_0 state, and yet they combine to form metals. In this case, the degenerate level that corresponds to all atoms being in the first excited state apparently drops below the nondegenerate state and broadens into a continuous band as the atoms are brought together (cf. Fig. 26b). This band presumably is similar to the low-lying band of the alkali metals and gives the alkaline earths their metallic charac-

teristics. It follows, from this example, that the first excited state of an insulator may be conducting for actual interatomic separations even though this state is nonconducting for large distances. A continuum that arises from degenerate states in which many atoms are excited or ionized can cross or become very close to the nonconducting levels, leaving the solid in the condition corresponding to that shown in Fig. 28.

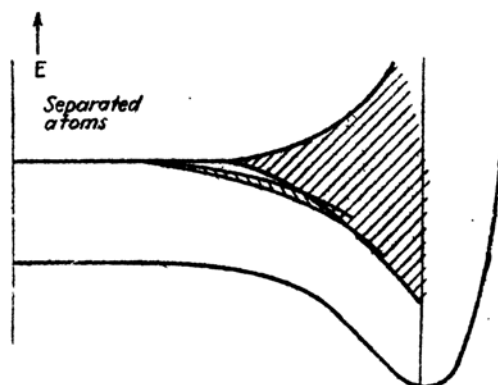


FIG. 28.—Case in which the conducting levels arising from the first excited state of an insulator overlap or come very close to the nonconducting levels at the actual interatomic distances. In this case, unlike that of Fig. 27, the first excited state is conducting.

67. Koopmans's Theorem*.—We shall now prove a theorem, due to Koopmans,¹ which states that the energy parameter ϵ_i in Fock's equations for a solid, namely,

$$H^r \varphi_i = \epsilon_i \varphi_i, \quad (1)$$

is the negative of the energy required to remove the electron in the state φ_i from the solid when the space part of the φ are Bloch type functions.

Suppose that the Hamiltonian for the entire solid is designated by H . Then when the j th electron is removed, the new Hamiltonian is

$$H' = H - \left(-\frac{\hbar^2}{2m} \Delta_j + \sum_i \frac{e^2}{r_{ij}} + V_j \right) \quad (2)$$

where V_j is the ion-core potential. The wave function of the initial state is

$$\Psi = \frac{1}{\sqrt{n!}} \begin{vmatrix} \varphi_1(\mathbf{r}_1) & \dots & \varphi_{j-1}(\mathbf{r}_1) & \varphi_j(\mathbf{r}_1) & \varphi_{j+1}(\mathbf{r}_1) & \dots & \varphi_n(\mathbf{r}_1) \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \varphi_1(\mathbf{r}_n) & \dots & \varphi_{j-1}(\mathbf{r}_n) & \varphi_j(\mathbf{r}_n) & \varphi_{j+1}(\mathbf{r}_n) & \dots & \varphi_n(\mathbf{r}_n) \end{vmatrix} \quad (3)$$

¹ T. KOOPMANS, *Physica*, 1, 104 (1933).

whereas the wave function for the state in which the electron having the function φ_j has been removed is

$$\Psi_j = \frac{1}{\sqrt{(n-1)!}} \begin{vmatrix} \varphi_1(\mathbf{r}_1) & \dots & \varphi_{j-1}(\mathbf{r}_1) & \varphi_{j+1}(\mathbf{r}_1) & \dots & \varphi_n(\mathbf{r}_1) \\ \vdots & & \vdots & \vdots & & \vdots \\ \varphi_1(\mathbf{r}_n) & \dots & \varphi_{j-1}(\mathbf{r}_n) & \varphi_{j+1}(\mathbf{r}_n) & \dots & \varphi_n(\mathbf{r}_n) \end{vmatrix} \quad (4)$$

Since the Fock Hamiltonian is practically unchanged when a Bloch type function is removed, because the electron charge is spread throughout the entire crystal, the φ in (3) and (4) are practically identical. This conclusion clearly is valid only when the electronic system is very large and when the one-electron functions are of the extended type.

The work done in removing the electron is simply

$$\int \Psi_j^* H' \Psi_j d\tau' - \int \Psi^* H \Psi d\tau \quad (5)$$

in which $d\tau'$ excludes the variables of the j th electron. Now,

$$\begin{aligned} \int \Psi_j^* H' \Psi_j d\tau' &= \sum_{i \neq j} \int \varphi_i^*(\mathbf{r}_1) \left[-\frac{\hbar^2}{2m} \Delta_1 + V(\mathbf{r}_1) \right] \varphi_i(\mathbf{r}_1) d\tau_1 + \\ &\quad \frac{e^2}{2} \sum_{i, k \neq j} \left[\int \frac{|\varphi_i(\mathbf{r}_1)|^2 |\varphi_k(\mathbf{r}_2)|^2}{r_{12}} d\tau_{12} - \int \frac{\varphi_i^*(\mathbf{r}_1) \varphi_k^*(\mathbf{r}_2) \varphi_i(\mathbf{r}_2) \varphi_k(\mathbf{r}_1)}{r_{12}} d\tau_{12} \right]. \end{aligned} \quad (6)$$

$\int \Psi^* H \Psi d\tau$ is equal to the same expression with the difference that the summations include the index j . Thus, (5) is

$$\begin{aligned} - \left\{ \int \varphi_j^*(\mathbf{r}_1) \left[-\frac{\hbar^2}{2m} \Delta_1 + V(\mathbf{r}_1) + \sum_i e^2 \int \frac{|\varphi_i(\mathbf{r}_2)|^2}{r_{12}} d\tau_2 \right] \varphi_j(\mathbf{r}_1) d\tau_1 - \right. \\ \left. \sum_i e^2 \int \frac{\varphi_i^*(\mathbf{r}_1) \varphi_i^*(\mathbf{r}_2) \varphi_i(\mathbf{r}_2) \varphi_j(\mathbf{r}_1)}{r_{12}} d\tau_{12} \right\} = \\ - \int \varphi_j^*(\mathbf{r}_1) H^P \varphi_j(\mathbf{r}_1) d\tau = -\epsilon_j \end{aligned}$$

(cf. Sec. 51), which proves Koopmans's theorem.

It should be noted that Koopmans's theorem also tells us that, in the Fock approximation, the energy required to take an electron from the state φ_k to the state $\varphi_{k'}$ is $\epsilon(k') - \epsilon(k)$ when the one-electron functions are of the Bloch type.

68. Velocity and Acceleration in the Bloch Scheme.—We shall now develop several theorems¹ concerning the behavior of electrons in the

¹ These theorems have been proved independently by several workers. See, for example, A. Sommerfeld and H. Bethe, *Handbuch der Physik*, Vol. XXIV/2; H. Jones

Bloch approximation. In particular, we shall be interested in the relation between electron velocity and energy and that between acceleration and external fields. In quantum theory, the quantities corresponding to velocity, acceleration, and force are represented by operators. The ordinary measured values of these quantities for a given state of a system are the mean values of the corresponding operators. Hence, in order to proceed in a rigorous fashion, we should compute these mean values and should derive relationships among them from the quantum laws. Actually, the correct final results may be obtained much more simply by using wave-packet methods. We shall employ this procedure.

a. The Relation between Velocity and Energy.—In order to determine the velocity of an electron that has a given energy $\epsilon(\mathbf{k}')$ and has the wave-number vector \mathbf{k}' , we shall construct a packet by adding together wave functions associated with values of \mathbf{k} in the vicinity of \mathbf{k}' and by determining the group velocity of this packet. This group velocity corresponds to the measured or mean value of the velocity of an electron having wave number \mathbf{k}' . When the proper time dependence is included, the Bloch wave functions are

$$\psi_{\mathbf{k}} = \chi_{\mathbf{k}} e^{2\pi i \mathbf{k} \cdot \mathbf{r}} e^{-2\pi i \frac{\epsilon(\mathbf{k})}{h} t} \quad (1)$$

where $\chi_{\mathbf{k}} e^{2\pi i \mathbf{k} \cdot \mathbf{r}}$ is the space-dependent part and $e^{-2\pi i \frac{\epsilon(\mathbf{k})}{h} t}$ is the time-dependent part. The packet, whose constituent functions are centered about \mathbf{k}' , is

$$f(\mathbf{r}, t) = \int a(\mathbf{k}) \chi_{\mathbf{k}} e^{2\pi i \left(\mathbf{k} \cdot \mathbf{r} - \frac{\epsilon(\mathbf{k})}{h} t \right)} d\tau(\mathbf{k}) \quad (2)$$

where $|a(\mathbf{k})|$ has a maximum value at $\mathbf{k} = \mathbf{k}'$ and decreases rapidly on either side. The integration extends over the range of values of \mathbf{k} about \mathbf{k}' in which $|a(\mathbf{k})|^2$ is appreciably different from zero and which may be made as small as we please. If we set $\mathbf{k} = \mathbf{k}' + \Delta\mathbf{k}$ and expand $\epsilon(\mathbf{k})$ in a Taylor series about \mathbf{k}' , we obtain

$$\epsilon(\mathbf{k}) = \epsilon(\mathbf{k}') + \text{grad}_{\mathbf{k}'} \epsilon(\mathbf{k}') \cdot \Delta\mathbf{k} + \dots \quad (3)$$

We shall assume that the range in which $\Delta\mathbf{k}$ is important is so small that we need retain only those terms indicated in (3). Equation (2) then becomes

$$f(\mathbf{r}, t) \cong e^{2\pi i \left[\mathbf{r} \cdot \mathbf{k}' - \frac{\epsilon(\mathbf{k}')}{h} t \right]} \int a(\Delta\mathbf{k}) \chi_{\mathbf{k}'} e^{2\pi i \left(\mathbf{r} - \frac{\text{grad}_{\mathbf{k}'} \epsilon}{h} t \right) \cdot \Delta\mathbf{k}} d\tau(\Delta\mathbf{k})$$

and C. Zener, *Proc. Roy. Soc.*, **144**, 101 (1934); H. Fröhlich, *Theorie der Metalle* (Julius Springer, Berlin, 1937).

where the integration now extends over a range of $\Delta\mathbf{k}$ about the value $\Delta\mathbf{k} = 0$. $\chi_{\mathbf{k}}$ varies so slowly with \mathbf{k} that we may take it outside the integral and write

$$f(\mathbf{r}, t) \cong \chi_{\mathbf{k}'}(\mathbf{r}) e^{2\pi i \left[\mathbf{r} \cdot \mathbf{k}' - \frac{\epsilon(\mathbf{k}')}{h} t \right]} \int a(\Delta\mathbf{k}) e^{2\pi i \left(\mathbf{r} - \frac{\text{grad}_{\mathbf{k}'} \epsilon}{h} t \right) \cdot \Delta\mathbf{k}} d\tau(\Delta\mathbf{k}). \quad (4)$$

Equation (4) describes a wave packet whose phase is determined by the function

$$\chi_{\mathbf{k}'} e^{2\pi i \left[\mathbf{k}' \cdot \mathbf{r} - \frac{\epsilon(\mathbf{k}')}{h} t \right]}, \quad (5)$$

which is the Bloch function corresponding to $\mathbf{k} = \mathbf{k}'$. Its group behavior is determined by the function

$$\int a(\Delta\mathbf{k}) e^{2\pi i \left(\mathbf{r} - \frac{\text{grad}_{\mathbf{k}'} \epsilon}{h} t \right) \cdot \Delta\mathbf{k}} d\tau(\Delta\mathbf{k}), \quad (6)$$

of which the argument is $\mathbf{r} - \frac{\text{grad}_{\mathbf{k}'} \epsilon}{h} t$. Thus, the function (6) has constant amplitude at those points for which $\mathbf{r} - t \text{grad}_{\mathbf{k}'} \epsilon$ is constant, that is, at points that move with the constant velocity

$$\mathbf{v}(\mathbf{k}') = \frac{1}{h} \text{grad}_{\mathbf{k}'} \epsilon(\mathbf{k}'), \quad (7)$$

which is the group velocity of the packet. The classical expression $|\mathbf{v}| = \sqrt{2\epsilon/m}$ for the speed is valid only in the special case in which

$$\epsilon = \frac{h^2}{2m} \mathbf{k}^2,$$

that is, for perfectly free electrons. The relation between speed and ϵ is different in all other cases. For example, near the top of a band where the slope of ϵ decreases, the speed decreases with increasing energy and actually may become zero if the slope of ϵ happens to vanish at the boundary of the zone.

The current i carried by an electron having velocity \mathbf{v} is $-e\mathbf{v}$. Hence, we obtain from (7)

$$\mathbf{i} = -e \frac{\text{grad}_{\mathbf{k}} \epsilon(\mathbf{k})}{h}. \quad (8)$$

b. The Effective Electron Mass.—We shall now derive the relation between acceleration and applied force. It follows from Eq. (7) that

$$\frac{d\mathbf{v}}{dt} = \frac{d}{dt} \left[\frac{\text{grad}_{\mathbf{k}} \epsilon(\mathbf{k})}{h} \right] = \frac{\text{grad}_{\mathbf{k}} [d\epsilon(\mathbf{k})/dt]}{h}. \quad (9)$$

The classical relation between force and energy is

$$\frac{d\epsilon}{dt} = \mathbf{F} \cdot \mathbf{v}, \quad (10)$$

which should remain valid for the mean values of quantum theory. Substituting (10) in (9), we find

$$\frac{d\mathbf{v}}{dt} = \frac{\text{grad}_{\mathbf{k}} \mathbf{F} \cdot \mathbf{v}}{\hbar} = \frac{\mathbf{F} \cdot [\text{grad}_{\mathbf{k}} \text{grad}_{\mathbf{k}} \epsilon(\mathbf{k})]}{\hbar^2} \quad (11)$$

where $\text{grad}_{\mathbf{k}} \text{grad}_{\mathbf{k}} \epsilon(\mathbf{k})$ is a tensor the nine components of which are

$$\begin{pmatrix} \frac{\partial^2 \epsilon}{\partial k_x^2} & \frac{\partial^2 \epsilon}{\partial k_y \partial k_x} & \frac{\partial^2 \epsilon}{\partial k_z \partial k_x} \\ \frac{\partial^2 \epsilon}{\partial k_x \partial k_y} & \frac{\partial^2 \epsilon}{\partial k_y^2} & \frac{\partial^2 \epsilon}{\partial k_z \partial k_y} \\ \frac{\partial^2 \epsilon}{\partial k_x \partial k_z} & \frac{\partial^2 \epsilon}{\partial k_y \partial k_z} & \frac{\partial^2 \epsilon}{\partial k_z^2} \end{pmatrix}. \quad (12)$$

Equation (11) is analogous to the classical equation

$$\frac{d\mathbf{v}}{dt} = \frac{1}{m} \mathbf{F}$$

and shows that in the Bloch approximation an electron in an energy band behaves as though it had an effective mass m^* represented by the tensor

$$\frac{1}{m^*} = \text{grad}_{\mathbf{k}} \text{grad}_{\mathbf{k}} \epsilon(\mathbf{k}). \quad (13)$$

Thus, the force and acceleration usually are not in the same direction.

Suppose that the electron is moving in the direction of a principal axis of the tensor (12) and that this direction is chosen to lie along the x axis. Then, the effective mass for acceleration in the x direction is

$$m^* = \frac{1}{d^2 \epsilon(\mathbf{k}) / dk_x^2}. \quad (14)$$

As we may see from Fig. 2, this is usually negative when the electron is near the top of the band. Hence, an electron in this position behaves as though it had a negative mass. The type of force that we shall usually consider is the combination of electrostatic and Lorentz force:

$$\mathbf{F} = -e\mathbf{E} - \frac{e}{c} \mathbf{v} \times \mathbf{H}.$$

Since e appears linearly in this, an electron with a negative effective mass is equivalent to a particle with a positive charge.

It was found while developing the free-electron theory of metals that the electrons that reside in the energy band of width kT at the top of the filled region are principally responsible for the electrical properties of metals. If this band lies above the inflection point of $\epsilon(\mathbf{k})$, all these electrons behave like positive charges. This fact evidently offers an explanation of the anomalous Hall effect: metals and semi-conductors that have a positive Hall constant contain nearly filled bands.

It sometimes is convenient to ascribe all the "anomalous" properties of metals and semi-conductors that are associated with the electrons of negative mass to the vacant levels or holes at the top of the nearly filled band. Thus, these holes may be treated as though they were positively charged particles having effective mass

$$\frac{1}{m^*} = - \text{grad}_{\mathbf{k}} \text{grad}_{\mathbf{k}} \epsilon(\mathbf{k}).$$

This convention is not simply an analogy, for Heisenberg¹ has shown that it is possible to replace the wave equation for the electrons in a nearly filled band by an approximate wave equation for the holes in the band. In this equation, the holes play the same role as positive charges. Heisenberg's theorem may also be applied to nearly filled atomic shells.

In addition to the preceding relations, we shall find the following one useful. By combining Eqs. (10) and (7), we find

$$\mathbf{F} \cdot \mathbf{v} = \frac{d\epsilon}{dt} = \frac{d\mathbf{k}}{dt} \cdot \text{grad } \epsilon = \hbar \frac{d\mathbf{k}}{dt} \cdot \mathbf{v},$$

whence

$$\frac{d\mathbf{k}}{dt} = \frac{\mathbf{F}}{\hbar}. \quad (15)$$

In this connection, Houston² has made an interesting investigation of the motion of an electron in the presence of an electrostatic field for the band approximation. He has shown that when the field is small, or when the electronic wave-number vector is not near the zone boundary, a good approximation for the solution of the time-dependent Schrödinger equation is

$$\chi_{\mu} e^{2\pi i \mu \cdot \mathbf{r}} e^{-\frac{2\pi i}{\hbar} \int \epsilon(\mu) dt}$$

where $\mu = \mathbf{k} - \frac{e\mathbf{E}t}{\hbar}$, \mathbf{E} being the field intensity. This function degenerates to $\psi_{\mathbf{k}} e^{-\frac{2\pi i \epsilon(\mathbf{k})t}{\hbar}}$ when \mathbf{E} is zero. In this approximation the electronic

¹ W. HEISENBERG, *Ann. Physik*, **10**, 888 (1931).

² W. V. HOUSTON, *Phys. Rev.*, **57**, 184 (1940).

wave-number vector moves through \mathbf{k} space at a uniform rate $e\mathbf{E}/\hbar$, or, in the reduced-zone scheme, the wave-number vector moves uniformly to the boundary of the zone in the direction of \mathbf{E} , suffers a Bragg reflection, that is, jumps to the opposite face of the zone, and starts back along the same path, repeating the process indefinitely. During these cycles, the system remains in the same energy surface $\epsilon(\mathbf{k})$, so that there is no transition between zones. In higher approximation, the electron may jump between zones for sufficiently strong fields. Houston has shown that these transitions occur when the wave-number vector is near the zone boundary and that in a one-dimensional case the probability of a transition in a single cycle is

$$\frac{4\pi^2 e^{-\alpha}}{(1 - e^{-\alpha})^2}$$

where $\alpha = mV^2d/eE\hbar$, in which m is the electronic mass, $2V$ is the energy gap, and d is the lattice distance. A similar expression had previously been derived by Zener¹ and applied to a discussion of the problem of dielectric breakdown in insulating crystals. It is believed at present that dielectric breakdown in simple crystals, such as the alkali halides, occurs for fields weaker than those required to cause transitions from the filled to the unoccupied levels by this process (see Sec. 133).

69. Modification of Boltzmann's Equation for the Bloch Scheme.

We are now in a position to modify Boltzmann's equation to suit the quantum mechanical one-electron mode of description instead of the classical one (cf. Sec. 31, Chap. IV). In place of the distribution function $f(x, y, z, v_x, v_y, v_z)$, which gives the number of particles at x, y , and z with velocity components v_x, v_y , and v_z , we shall introduce the function

$$f(x, y, z, k_x, k_y, k_z) \equiv f(\mathbf{r}, \mathbf{k}), \quad (1)$$

which gives the number of particles at x, y, z with wave-number components k_x, k_y, k_z .

The drift term, analogous to (2), Sec. 31, Chap. IV, is

$$\left(\frac{df}{dt}\right)_{\text{drift}} = -\frac{\partial f}{\partial x}v_x - \frac{\partial f}{\partial y}v_y - \frac{\partial f}{\partial z}v_z - \frac{\partial f}{\partial k_x}\dot{k}_x - \frac{\partial f}{\partial k_y}\dot{k}_y - \frac{\partial f}{\partial k_z}\dot{k}_z, \quad (2)$$

which may be expressed in terms of $\text{grad}_{\mathbf{k}} \epsilon(\mathbf{k})$ and \mathbf{F} by means of Eqs. (7) and (15) of the preceding section. The result is

$$\left(\frac{df}{dt}\right)_{\text{drift}} = -\frac{\text{grad}_{\mathbf{k}} \epsilon}{\hbar} \cdot \text{grad}_{\mathbf{k}} f - \frac{\mathbf{F}}{\hbar} \cdot \text{grad}_{\mathbf{k}} f. \quad (3)$$

¹C. ZENER, *Proc. Roy. Soc.*, **145**, 523 (1934).

The collision terms may now be expressed in terms of the function $\Theta(k_x, k_y, k_z; k'_x, k'_y, k'_z) dk'_x dk'_y dk'_z$, which gives the probability that a particle changes its wave-number components from k_x, k_y, k_z to values in the range from k'_x to $k'_x + dk'_x$, etc. The quantities that correspond to a and b , Sec. 31, Chap. IV, are then

$$a = f(x, y, z, k_x, k_y, k_z) \int \Theta(k_x, k_y, k_z; k'_x, k'_y, k'_z) dk'_x dk'_y dk'_z, \quad (4)$$

$$b = \int f(x, y, z, k'_x, k'_y, k'_z) \Theta(k'_x, k'_y, k'_z; k_x, k_y, k_z) dk'_x dk'_y dk'_z. \quad (5)$$

The equation of equilibrium analogous to Boltzmann's equation, is

$$\frac{\text{grad}_k \epsilon}{h} \cdot \text{grad}_r f + \frac{\mathbf{F}}{h} \cdot \text{grad}_k f = b - a, \quad (6)$$

which may be solved when $b - a$ is known.

In general, a and b may be determined from (4) and (5) when Θ has been determined by means of a quantum mechanical analysis. We shall discuss particular cases of this procedure in Chap. XV.

It may be concluded from the results of the present discussion that the developments of Chap. IV can be modified in two important respects, namely:

a. Although the electrons may behave as though effectively free, so that the relation

$$\epsilon = \frac{h^2}{2m^*} k^2 = \frac{p^2}{2m^*} \quad (7)$$

is valid, as assumed in Chap. IV, it should be recognized that the effective mass m^* may be negative in cases in which the bands are nearly filled. As we have seen in the preceding section, this situation may be treated by assuming that both m^* and the electronic charge are positive. Thus, the results of Chap. IV can be employed in these cases by reversing the sign of ϵ .

b. It may be necessary to use the tensor character of the effective mass in order to take into account the anisotropy even of cubic metals. This procedure probably is necessary in discussing the dependence of resistivity on magnetic field strength; for as we have seen in Sec. 34, the results derived on the basis of (7) are badly in error but would be improved if the velocity at the top of the filled region were an anisotropic function.

70. Additional Energy States.—It was mentioned in part *b*, Sec. 61, that some of the nonperiodic states that are labeled under Kramers' class i may be permitted if the crystal is not infinitely large. If the crystal does extend to infinity in all directions, the amplitude of these states diverges in at least one direction. The possibilities for the finite lattice were first pointed out by Tamm¹ who showed that the divergences

¹ I. TAMM, *Physik. Z. Sowj.*, **1**, 733 (1932).

can be avoided in special cases. We shall discuss his computations for a simple model in the following paragraphs. Before doing so, however, we should add that a computation that is based upon the use of a static one-electron field, as Tamm's is, is not completely adequate for establishing the general conditions under which nonperiodic states are permitted, since the actual field in which an electron moves is determined in part by the electron itself.

Tamm considers the simple potential field that is shown in Fig. 29, which, for positive values of x , is the same as that used by Kronig and Penney and has the constant value W for negative values. As before, we shall allow V_0 to approach infinity and shall assume that $mV_0b(a-b)/\hbar^2$

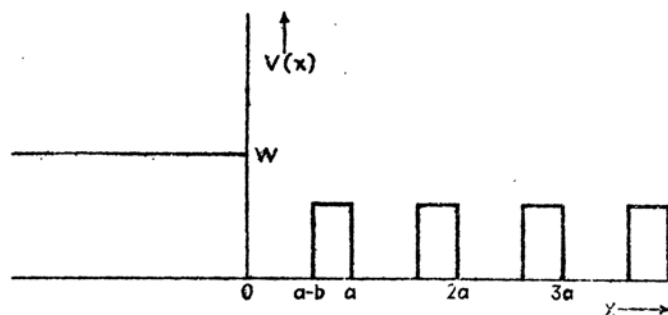


FIG. 29.—The potential function used by Tamm.

approaches the constant value c in the limit. One restriction on the solutions then is

$$\cos \lambda a = \frac{c}{\alpha a} \sin \alpha a + \cos \alpha a \quad (1)$$

[cf. Eq. (30), Sec. 61], where

$$\alpha = \frac{\sqrt{2mE}}{\hbar}$$

and $e^{\alpha a}$ is the factor by which the wave function is multiplied when x changes to $x + a$. We found previously that λ is necessarily real, for otherwise the wave functions would diverge at large values of x . Hence, the allowed values of E were determined by the condition that the right-hand side of (1) should lie between the values -1 and $+1$. These conclusions are no longer valid since the wave functions are different from Kronig and Penney's in the negative region of x . Instead, they are exponential functions of the type

$$\psi = c'e^{\gamma x} \quad (2)$$

where

$$\eta = \frac{\sqrt{2m(W - E)}}{\hbar}.$$

These functions should be joined to Kronig and Penney's at the origin of x .

It turns out that all of Kronig and Penney's solutions going with real values of λ may be joined to functions of type (2), if the phases are chosen properly. Hence, the continuous spectrum of periodic states is not altered. We must consider the case in which λ is imaginary.

The solution that decreases by a factor $ee^{-\mu a}$ in successive cells of the lattice along the positive x axis is

$$\psi_\mu = \left(e^{-i\alpha x} - \frac{1 - \epsilon e^{(\mu - i\alpha)a}}{1 - \epsilon e^{(\mu + i\alpha)a}} e^{i\alpha x} \right), \quad (3)$$

where ϵ is $+1$ when the right-hand side of (1) is greater than $+1$ and is -1 when the right-hand side is less than -1 . The condition (1) is replaced by

$$\epsilon \cosh \mu a = \frac{c}{\alpha a} \sin \alpha a + \cos \alpha a. \quad (4)$$

The additional condition that must be satisfied if (2) and (3) are to join smoothly at the origin is

$$\eta \frac{\sin \alpha a}{\alpha} + \cos \alpha a = \epsilon e^{-\mu a} \quad (5)$$

for the case in which $E < W$. Eliminating μ from (4) and (5), we find that E must satisfy the equation

$$\alpha a \cotan \alpha a = \frac{a^2 \gamma^2}{c} - \sqrt{a^2 \gamma^2 - a^2 \alpha^2} \quad (6)$$

where

$$\gamma^2 = \frac{2mW}{\hbar^2}.$$

This equation possesses one root in each interval of αa extending from $n\pi$ to $(n + 1)\pi$. Since there is one continuous band in each of these intervals, we may conclude that there is one "surface" state in the energy region between each pair of neighboring bands, when E is less than W .

When E is greater than W , the functions (2) are periodic and may be joined to (3) for arbitrary values of E . The resulting solutions are periodic outside the lattice and exponentially damped inside.

We may conclude that the type of surface barrier which is considered here permits states which lie in the "unallowed" energy region.

These states are exponentially damped on both sides of the surface when E is less than W and are periodic outside the lattice and damped inside when E is greater than W . The first type represents electrons that are localized at the surface, whereas the second represents electrons that impinge on the surface from outside and are reflected back.

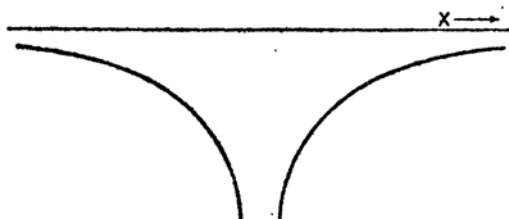


FIG. 30.—A simple potential trough.

Fowler¹ has pointed out that in a finite one-dimensional crystal the surface states occur in pairs, one state being associated with each end of the crystal.

In addition, Shockley² examined the origin of the surface levels more thoroughly on the basis of a more general one-dimensional model and considered the dependence of the levels on lattice parameter. He found that if the potential of the normal lattice may be expressed as a periodic sum of simple troughs of the type shown in Fig. 30 the surface levels can

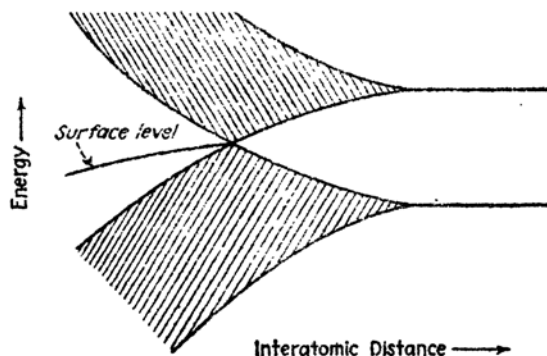


FIG. 31.—Schematic diagram showing the manner in which the surface levels occur in a case in which the potential is a periodic sum of troughs of the type shown in Fig. 30.

occur only if there is a separate potential trough at the surface or if the energy bands arising from separate atomic levels overlap (see Fig. 31). The second case evidently cannot occur at large interatomic separations. It turns out that the bands do not overlap in the Kronig-Penney model and that Tamm obtained surface levels only because he implicitly intro-

¹ R. H. FOWLER, *Proc. Roy. Soc.*, **141**, 56 (1933).

² W. SHOCKLEY, *Phys. Rev.*, **56**, 317 (1939). See also the similar discussion by W. G. POLLARD, *Phys. Rev.*, **56**, 324 (1939).

duced an additional potential trough at the surface in cutting off the potential function in the manner shown in Fig. 29.

Shockley also showed that one level is removed from the lower band and one is removed from the upper band for each pair of surface levels that occurs. Thus, if the lower band were completely filled and the upper band were completely empty at large interatomic separations, two electrons of opposite spin would be forced from the lower band into the surface levels when overlapping occurs.

The generalization for three dimensions is apparently straightforward. Suppose that we have a crystal that is bounded by the plane $x = 0$, the crystal being on the positive side of the x axis. We may expect some wave functions that have energies lying in the unallowed regions and that have the form

$$\psi = \chi e^{-\mu x} e^{2\pi i(k_y y + k_z z)} \quad (7)$$

inside the lattice and the form

$$\psi = e^{-\eta x} e^{2\pi i(k_y y + k_z z)} \quad (8)$$

outside. In (7), χ is a function possessing the periodicity of the lattice; η is real when E is less than W , the potential outside of the lattice, and is imaginary when E is greater than W . As before, the second type of solution is permitted for all values of E in the unallowed regions. We may expect more than one solution of the first type in the unallowed regions in which they may occur, however, for k_x and k_y may take all values associated with a two-dimensional zone system that is determined by the translational symmetry of the surface. Roughly speaking, we may expect as many states in these unallowed regions as there are atoms on the surface. Since this number ordinarily is about one million times smaller than the total number of atoms in a crystal, we should not expect these surface levels to affect the bulk properties of a substance.

There does not seem to be any direct experimental evidence for the existence of surface states although Tamm suggests that charges on the surfaces of some charged insulators may be bound in states of this type.

Shockley has made a qualitative generalization of the results of his investigation of the one-dimensional model. His conclusions are as follows:

Surface levels will not occur between the ordinary X-ray levels or in the forbidden region of most simple insulating salts such as sodium chloride, for neither of these cases corresponds to overlapping-band systems. The basis for these conclusions is discussed in Chap. XIII.

b. Surface states should occur in the forbidden region between the highest filled band and lowest vacant band of diamond, for this gap occurs as a result of the overlapping of an s band and a p band (see Sec. 109).

Since the number of surface states is approximately twice the number of electrons that are forced from the filled band, the surface band should not be completely filled. In an ideal case, this statement would imply that the surface should be conducting; however, various effects, such as surface cracks or adsorbed atoms, could easily impair this type of conductivity, which does not seem to be observed.

c. Surface states should occur near the overlapping bands of all metals. In the monovalent metals, such as the alkali metals, in which the occupied band is usually not filled to the zone boundary, these levels would be completely empty. They should be partly filled in the divalent metals, however.

More important than the existence of surface states is the fact, suggested by these computations, that bound electronic states may be

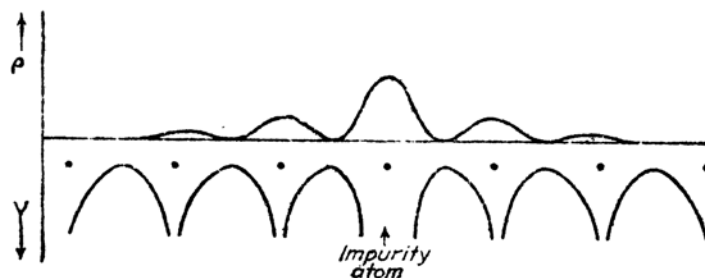


FIG. 32.—Schematic representation of a localized impurity level. The lower curve is the lattice potential which is distorted by an impurity atom. The upper curve represents the localized charge distribution associated with an impurity level that lies in the forbidden region.

associated with any flaw or discontinuity in an otherwise perfectly periodic lattice in the manner illustrated in Fig. 32. Suppose, for example, that we have an infinite one-dimensional lattice, such as that of Kronig and Penney, and that we alter the potential in a single cell by lowering it from zero to $-W'$. It is easy to show by the method used above that in the forbidden energy regions there are then states corresponding to electrons which are localized in the vicinity of the singular cell. If the cell extends from $-a$ to 0 , the allowed forms of the wave function within this range are

$$\sin \gamma' \left(x + \frac{a}{2} \right) \quad \text{or} \quad \cos \gamma' \left(x + \frac{a}{2} \right) \quad (9)$$

where

$$\gamma' = \frac{\sqrt{2m(E + W')}}{\hbar} \quad (10)$$

The localized functions correspond to cases in which the functions (9) are joined smoothly to functions of type (3) at $x = 0$ and $x = a$. The

conditions for this in the two cases (9) are, respectively,

$$\frac{\gamma'}{\alpha} \cotan \frac{\gamma'a}{2} \sin \alpha a + \cos \alpha a = e e^{-\mu a} \quad (11)$$

and

$$-\frac{\gamma'}{\alpha} \tan \frac{\gamma'a}{2} \sin \alpha a + \cos \alpha a = e e^{-\mu a}. \quad (11a)$$

Either one of these conditions and the condition (4) must be satisfied simultaneously. The resulting equations are more complicated than (6) and will not be discussed in detail here. They show that when E is negative we may expect approximately as many trapped electron states as there are discrete states for a simple barrier of width a and depth $-V'$, whereas, for values of E greater than zero, there usually is one state in each forbidden region. However, there may be more or less than one in particular regions. These cases depend in a complex way upon the relative values of αa and $\gamma'a$.

Thus, we may conclude that impurity atoms or lattice imperfections induce additional energy states which correspond to electrons localized in the vicinity of the impurity or defect. These states lie in the regions between continuous bands. We shall find that they can play a very important role, particularly in semi-conductors.

71. Optical Transitions in the Zone Approximation.—Before leaving the zone approximation, we shall find it convenient to discuss optical transition probabilities. We found in Sec. 43 that the probability of an optical transition between two states Ψ_α and Ψ_β contains the following integral

$$\int \Psi_\alpha^* \sum_i \mathbf{p}_i e^{2\pi i \boldsymbol{\eta} \cdot \mathbf{r}_i} \Psi_\beta d\tau(x_1, \dots, x_n, \zeta_1, \dots, \zeta_n) \quad (1)$$

in which \mathbf{p}_i is the momentum operator for the i th electron, \mathbf{n} is the wave-number vector of the light quantum that is absorbed or emitted, and the integration extends over all of the electronic coordinates. In the band approximation, Ψ_α and Ψ_β are determinantal wave functions that are constructed of Bloch one-electron functions. Since the operator

$$\sum_i \mathbf{p}_i e^{2\pi i \boldsymbol{\eta} \cdot \mathbf{r}_i} \quad (2)$$

is a sum of one-electron operators, the integral in (1) vanishes if Ψ_α and Ψ_β differ with respect to more than one Bloch function; moreover, (1) vanishes even in this case, unless the electron spins associated with the two different Bloch functions are the same. Hence, only one electron can change its state during the absorption or emission of a single light quantum.

If $\psi_{\mathbf{k}}$ and $\psi_{\mathbf{k}'}$ are, respectively, the different Bloch functions in the functions Ψ_{α} and Ψ_{β} , (1) may be reduced to the form

$$\int \psi_{\mathbf{k}}^* p e^{2\pi i \mathbf{q} \cdot \mathbf{r}} \psi_{\mathbf{k}'} d\tau(x, y, z) \quad (3)$$

where the integral now extends over the coordinates of one electron. Let us write $\psi_{\mathbf{k}}$ and $\psi_{\mathbf{k}'}$ in the typical Bloch function forms

$$\begin{aligned} \psi_{\mathbf{k}} &= \chi_{\mathbf{k}} e^{2\pi i \mathbf{k} \cdot \mathbf{r}}, \\ \psi_{\mathbf{k}'} &= \chi_{\mathbf{k}'} e^{2\pi i \mathbf{k}' \cdot \mathbf{r}}. \end{aligned} \quad (4)$$

The integral (3) may then be written as

$$\int e^{2\pi i (\mathbf{k}' + \mathbf{q} - \mathbf{k}) \cdot \mathbf{r}} \chi_{\mathbf{k}}^* \left[\frac{\hbar}{i} \text{grad } \chi_{\mathbf{k}'} - \hbar 2\pi (\mathbf{n} + \mathbf{k}') \chi_{\mathbf{k}'} \right] d\tau. \quad (5)$$

The quantity

$$f = \chi_{\mathbf{k}}^* \left[\frac{\hbar}{i} \text{grad } \chi_{\mathbf{k}'} - \hbar 2\pi (\mathbf{n} + \mathbf{k}') \chi_{\mathbf{k}'} \right] \quad (6)$$

has the periodicity of the unit cell, so that (5) may be written in the form

$$\sum_{\mathbf{r}_i} e^{2\pi i (\mathbf{k}' + \mathbf{q} - \mathbf{k}) \cdot \mathbf{r}_i} \int_i e^{2\pi i (\mathbf{k}' + \mathbf{q} - \mathbf{k}) \cdot (\mathbf{r} - \mathbf{r}_i)} f d\tau_i \quad (7)$$

where the integral \int_i extends over the i th unit cell of the lattice and \mathbf{r}_i is a vector extending from a corner of this cell to the origin of coordinates. Since each integral in the sum (7) is the same, this series is

$$A \sum_{\mathbf{r}_i} e^{2\pi i (\mathbf{k}' + \mathbf{q} - \mathbf{k}) \cdot \mathbf{r}_i} \quad (8)$$

where

$$A = \int_0 e^{2\pi i (\mathbf{k}' + \mathbf{q} - \mathbf{k}) \cdot \mathbf{r}} f d\tau_0$$

in which the integral extends over the unit cell at the origin. Now, (8) vanishes unless

$$\mathbf{k}' + \mathbf{n} - \mathbf{k} = \mathbf{K} \quad (9)$$

where \mathbf{K} is a principal vector in the inverse lattice. Hence, transitions are allowed only between one-electron states the wave-number vectors of which satisfy the relation (9). The wave length of a light quantum

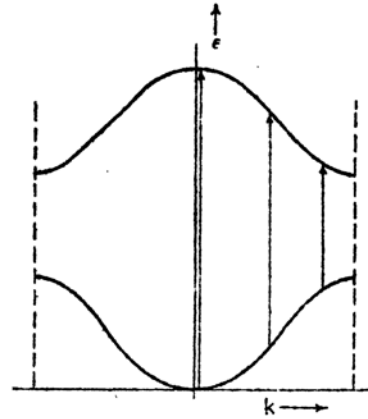


FIG. 33.—Diagram showing the allowed transitions in the reduced-zone scheme.

ordinarily is large compared with the wave lengths of an electron of comparable energy. Hence, \mathbf{n} is usually much smaller than \mathbf{k} or \mathbf{k}' and may be neglected. Equation (9) then simplifies to

$$\mathbf{k}' = \mathbf{k} + \mathbf{K},$$

which states that electrons may make only vertical transitions in the reduced-zone scheme (*cf.* Fig. 33).

CHAPTER IX

APPROXIMATIONAL METHODS

72. Introduction.—One of the most useful methods for obtaining approximate solutions of the Schrödinger equation for solids will be discussed briefly in this chapter. This discussion is supplementary to that of Chap. VI, for the one-electron schemes described there form the basis for the method described here. This method begins by replacing Fock's equations, which usually cannot be separated into one-variable equations, by central field equations that are separable. When accurate one-electron functions have been derived in this way, they are used to compute coulomb and exchange energies. Following this, an attempt is made to estimate the correlation effect and correlation energy. It is difficult to treat these quantities either accurately or concisely; however, they have been handled in a few special cases that will be discussed near the end of the chapter.

73. The Cellular Method.—The primary requirement of a practical plan for solving Fock's equations is that it should replace them by accurate separable equations. Hartree's procedure in the case of free atoms (*cf.* Chap. VI) is a good illustrative example of such a plan: Hartree's equations are not separable when they are applied to an electron configuration that involves an incompletely filled shell of p or d electrons. If the nonspherical part of the coulomb potential of p or d electrons is dropped, however, the equations become separable and may be solved by the methods used for ordinary differential equations. The error made in dropping the nonspherical terms lies within the limits of natural error of the Hartree field method and may be conveniently corrected by perturbation methods.

A similar procedure is possible in solids.¹ Let us restrict the discussion, for the present, to the case of Hartree's equations and overlook the exchange terms. These equations are

$$-\frac{\hbar^2}{2m}\Delta\psi_k(\mathbf{r}_1) + \left[V(\mathbf{r}_1) + \sum_{\mathbf{k}'} c^2 \int \frac{|\psi_{\mathbf{k}'}(\mathbf{r}_2)|^2}{r_{12}} d\mathbf{r}_2 \right] \psi_k(\mathbf{r}_1) = \epsilon \psi_k(\mathbf{r}_1) \quad (1)$$

where $V(\mathbf{r}_1)$ is the total ion-core potential and the sum in the second term extends over all electrons except the k th. The wave function near the

¹ E. WIGNER and F. SEITZ, *Phys. Rev.*, **43**, 804 (1933); **46**, 509 (1934).

nucleus of any atom is determined by the ion-core field of this atom, for this field becomes very large compared with the other potential terms in (1). In simple solids in which the ion cores are closed shells, this field is spherically symmetrical. The potential arising from other parts of the lattice is comparable with that of the ion core of this atom at distances from the nucleus that are of the same order as interatomic dimensions. If the crystal has a high degree of rotational symmetry relative to the nucleus, the potential of the rest of the lattice is nearly spherically symmetrical. Thus, it may be expected that there is a large domain about each nucleus in which the field may be replaced by a spherically symmetrical one. This principle may be used in a wide range of cases, although it is less accurate in crystals having low symmetry than in cubic lattices, such as the alkali metals and alkali halides.

The preceding observation on the symmetry of the field in the neighborhood of each nucleus suggests that the lattice should be partitioned into a set of space-filling polyhedra, which are centered about each of the nuclei, and that the field may be chosen to be centrally symmetrical within each of these polyhedra. Within each polyhedron, Hartree's equations may then be replaced by the equation

$$-\frac{\hbar^2}{2m}\Delta\psi + V(r)\psi = \epsilon\psi \quad (2)$$

where $V(r)$ is the approximate spherically symmetrical field. The solutions of (2) in spherical polar coordinates have the form

$$\psi = f_l(r, \epsilon)\Theta_m^l(\theta, \varphi) \quad (3)$$

where $f_l(r, \epsilon)$ is a radial function that satisfies the radial equation

$$-\frac{\hbar^2}{2m}\frac{d^2f}{dr^2} + \left[V(r) + \frac{\hbar^2}{2m}\frac{l(l+1)}{r^2} \right]f = \epsilon f \quad (4)$$

and $\Theta_m^l(\theta, \varphi)$ is a surface harmonic

$$\Theta_m^l = \sqrt{\frac{(l-m)!}{(l+m)!}} \frac{2l+1}{4\pi} P_l^m(\cos \theta) e^{im\varphi}. \quad (5)$$

Bloch functions may be constructed from functions of the type (3) by forming series of those which are associated with the same value of ϵ . The coefficients in these series may be determined from appropriate boundary conditions which we shall discuss below. This procedure forms the basis of the cellular method.

The manner in which cells are chosen depends upon the lattice for which computations are being made. For monatomic crystals in which all atoms are translationally equivalent, the most convenient cell is

chosen by taking the polyhedron whose plane faces bisect orthogonally the lines joining an atom with its nearest neighbors. Figures 1 and 2 show cells of this type for monatomic face-centered and body-centered cubic lattices. The cell may be chosen in the same way when more than one atom is present in the unit cell if they are equivalent, as in diamond or in closed-packed hexagonal crystals. Figures 3 and 4 show this type of cell for the crystals that were just mentioned. If the atoms in the unit cell are not equivalent, as in sodium chloride, the cells may not be chosen

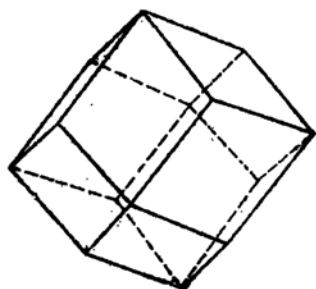


FIG. 1.—The cellular polyhedron for a monatomic face-centered cubic lattice.

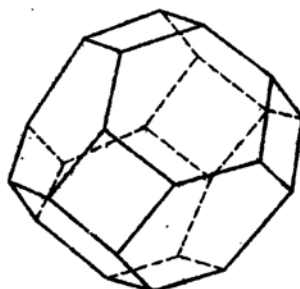


FIG. 2.—The cellular polyhedron for a monatomic body-centered cubic lattice.

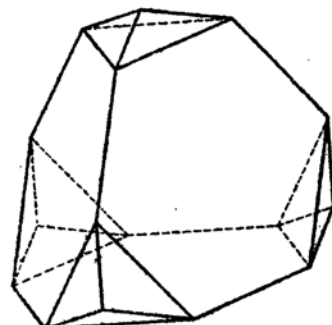


FIG. 3.—The cellular polyhedron for diamond.

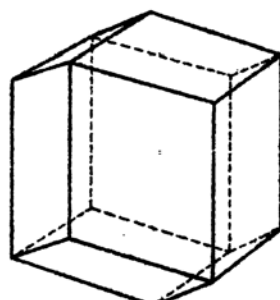


FIG. 4.—The cellular polyhedron for a close-packed hexagonal lattice.

on the basis of symmetry alone. Each case of this type can be handled in many ways.

As we mentioned above, within any cell, the Bloch function $\psi_{\mathbf{k}}$, associated with the energy ϵ and wave number \mathbf{k} , may be expressed¹ in terms of a series of the type

$$\psi_{\mathbf{k}}(\mathbf{r}) = \sum_{l,m} b_{\mathbf{k}}^{l,m} f_l(\mathbf{r}, \epsilon(\mathbf{k})) \Theta_m^l(\theta, \varphi). \quad (6)$$

¹ WIGNER and SMITZ, *op. cit.*; J. C. SLATER, *Phys. Rev.*, **45**, 794 (1934).

The important practical problem associated with this series is that of determining the b from the boundary conditions which the Bloch functions must satisfy. These conditions are the following: (a) $\psi_{\mathbf{k}}$ and its first derivative must be continuous at the boundary points between neighboring polyhedra, and (b) $\psi_{\mathbf{k}}$ must satisfy the relation

$$\psi_{\mathbf{k}}(\mathbf{r} + \boldsymbol{\epsilon}) = \psi_{\mathbf{k}} e^{2\pi i \mathbf{k} \cdot \boldsymbol{\epsilon}} \quad (7)$$

where $\boldsymbol{\epsilon}$ is a translation vector of the lattice. It turns out that these conditions can be satisfied only for discrete values of $\boldsymbol{\epsilon}$ for a given wave-number vector \mathbf{k} . The permissible solutions furnish us the desired relationship between $\boldsymbol{\epsilon}$ and \mathbf{k} . It may easily be seen that the boundary conditions need be satisfied only for points within a single unit cell, for the form of the function at any point outside this cell is connected with a value inside by Eq. (7). Let us suppose that condition a has been satisfied at all the interfaces between polyhedra in a given unit cell. By definition, the vectors that join opposite faces of the remaining surface are primitive translations of the lattice, since these faces constitute the boundary of the unit cell. Moreover, the points on the faces are the only ones in the unit cell to which the condition (7) can apply.

It has not been feasible in any of the work that has been carried through up to the present time to satisfy conditions a and b at all points of the surfaces of the polyhedra. Instead, all but a finite number of the terms in the series (6) are discarded, and boundary conditions are satisfied at just enough points to determine the coefficients of all these terms. The only justification for this procedure lies in the belief that the series (6) should converge rapidly for small values of \mathbf{k} , since then the wave length of the Bloch function is large compared with the dimensions of the cell. Results that have been derived by using this method will be presented in the following chapters.

Shockley¹ made an extensive test of the cellular method in a case in which exact solutions are known, namely, in which $V(r)$ is constant. He found that when a small number of boundary points is used the approximation is satisfactory for zones which normally are occupied but that it usually is very bad for excited states.

Several improvements² on the cellular method have been proposed since Shockley's work; but only one of these, namely, the method of Herring and Hill, has been applied to practical problems. These workers assumed that the $\psi_{\mathbf{k}}$ functions at a given point on the zone may be expressed as a finite sum of free-electron functions of the type

¹ W. SHOCKLEY, *Phys. Rev.*, **52**, 866 (1937).

² J. C. SLATER, *Phys. Rev.*, **51**, 846 (1937); G. WANNIER, *Phys. Rev.*, **53**, 671 (1938); C. C. HERRING and A. G. HILL, *Phys. Rev.* (to appear).

$$\psi_{\mathbf{k}} = \sum_{i=1}^n a_i e^{2\pi i \mathbf{k}_i \cdot \mathbf{r}} \quad (8)$$

in which the constants a_i and \mathbf{k}_i were chosen by use of group theory. The matrix components of the crystalline potential that connect approximate wave functions of the type (8) at the corresponding points in several zones were then computed, and the resulting matrix was diagonalized. In this way, approximate values of $\epsilon(\mathbf{k})$ for the boundary points were obtained. This method, which was applied by Herring and Hill to beryllium (*cf.* Sec. 81), evidently is a special case of the perturbation method described in Sec. 61, the functions (8) being the appropriate linear combinations in the zero-order approximation.

74. The Hartree Field.—In order to determine a self-consistent Hartree field within each of the polyhedra in the cellular approximation, it is necessary, first of all, to adopt a starting field or charge distribution from which wave functions may be computed. This field may be chosen in many ways. For example, when one is dealing with a monatomic solid, the ion-core field plus the field arising from a uniform distribution of valence electrons may be used. In any case, the starting potential V_I and the starting charge distribution ρ_I are related by the equation

$$V_I(\mathbf{r}_1) = \int \frac{\rho_I(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_2. \quad (1)$$

Bloch functions $\psi_{\mathbf{k}}$ may be constructed from the starting field by use of the method described in Sec. 73, which combines the solutions of the equation

$$-\frac{\hbar^2}{2m} \Delta \psi + V_I(\mathbf{r}) \psi = \epsilon \psi \quad (2)$$

where V_I is the spherically symmetric part of (1) in a given polyhedron. A new electronic charge distribution, $e \sum_{\mathbf{k}} |\psi_{\mathbf{k}}(\mathbf{r})|^2$, where \mathbf{k} is summed over all occupied levels, is determined by these $\psi_{\mathbf{k}}$, and a new Hartree field

$$V_{II}(\mathbf{r}_1) = e^2 \int \frac{\sum |\psi_{\mathbf{k}}(\mathbf{r}_2)|^2}{r_{12}} d\mathbf{r}_2 + V_c(\mathbf{r}_1) \quad (3)$$

may be determined from this distribution. In (3), $V_c(\mathbf{r}_1)$ is the total potential from the rigid ion cores; that is,

$$V_c(\mathbf{r}_1) = \sum_{i,\alpha} v_{i,\alpha}(\mathbf{r}_1) \quad (4)$$

where $v_{i,\alpha}(\mathbf{r}_1)$ is the potential at \mathbf{r}_1 that arises from the core of the α th

atom in the i th unit cell. The entire potential (3) may be written in the form of a lattice sum, such as the sum in Eq. (4), by expressing the first term of (3) in the form

$$\sum_{\mathbf{k}} e^2 \int \frac{|\psi_{\mathbf{k}}(\mathbf{r}_2)|^2}{r_{12}} d\mathbf{r}_2 = \sum_{i,\alpha} \int \frac{\rho_{i,\alpha}(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_2$$

where $\rho_{i,\alpha}$ is the value of $\sum_{\mathbf{k}} e^2 |\psi_{\mathbf{k}}|^2$ within the α th polyhedron of the i th unit cell.

Usually, $V_{\text{II}}(\mathbf{r})$ will not be spherically symmetrical within a polyhedron, since it may contain all surface harmonics that are compatible with the symmetry of the polyhedron. Hence, the closest agreement that may be expected, is that V_{I} and the spherically symmetric part of V_{II} should be identical. This will not be the case, unless by fortunate chance. Hence, it usually is necessary to choose a new field V_{III} as the starting field for another computation. There are no general rules for choosing V_{III} in such a way that the field V_{IV} , which is derived from its wave functions in the way in which V_{II} was derived from the solutions for V_{I} , will be closer to V_{III} than V_{II} was to V_{I} . The convergence is often swift in a monatomic lattice of equivalent atoms if V_{III} is taken as the mean of V_{I} and V_{II} , but this scheme does not work very well in solids that contain two or more different kinds of atom. The factors that govern the speed of convergence have not been investigated in any general way.

The final wave functions that are derived from a self-consistent Hartree field may differ appreciably from the solutions of Fock's equations, for exchange terms are neglected in Hartree's equations. Unfortunately, the exchange terms usually cannot be included merely by adding one-electron potential terms (cf. Chap. VI). There are special cases, however, in which they may be included very simply; we shall discuss these in the next section.

75. Exchange Terms*.—There are two cases in which the exchange terms have been handled rigorously, namely, the cases of perfectly bound electrons and perfectly free electrons. In the first case, the atoms are so far apart that the electronic wave functions of separate atoms do not overlap appreciably; in the second, the potential field in which the electrons move is so nearly constant that the one-electron functions have the form $e^{2\pi i \mathbf{k} \cdot \mathbf{r}}$. We shall discuss these two cases in detail.

a. Rigidly Bound Electrons (Narrow Bands).—Atoms and ions when they are far apart affect only the electrostatic field in their own vicinity. Hence, in this case, the atomic or Heitler-London approximation is the most accurate of the one-electron approximations, and the Bloch approximation is as accurate only when it is identical with the Heitler-

London method. We saw in Chap. VIII that the two schemes are equivalent when there are completely closed shells if determinantal eigenfunctions are used. Let us consider the way in which this equivalence appears in Fock's equations for the two systems. For simplicity, we shall deal with a monatomic lattice of atoms whose valence electrons form closed shells of the s^2 type. We shall let $\psi(\mathbf{r} - \mathbf{r}(n))$ represent the Heitler-London wave function of the electron that is centered about the nucleus at the position $\mathbf{r}(n)$. Fock's equation for $\psi(\mathbf{r} - \mathbf{r}(n))$ has the form

$$-\frac{\hbar^2}{2m}\Delta\psi(\mathbf{r}_1 - \mathbf{r}(n)) + \left\{ V(\mathbf{r}_1 - \mathbf{r}(n)) + \int \frac{|\psi(\mathbf{r}_2 - \mathbf{r}(n))|^2}{r_{12}} d\tau_2 \right\} \psi(\mathbf{r} - \mathbf{r}(n)) = e\psi(\mathbf{r}_1 - \mathbf{r}(n)) \quad (1)$$

where $V(\mathbf{r}_1 - \mathbf{r}(n))$ is the ion-core field of the atom at $\mathbf{r}(n)$ and the integral is the coulomb potential of the other electron on this atom. Exchange integrals do not occur because, by assumption, electrons on the same atoms have opposite spin and those on different atoms do not overlap appreciably. Outside a given atom, the coulomb field of the electrons cancels the ion-core field; hence, there are no coulomb terms between different atoms. Since exchange terms are absent, Hartree's and Fock's equations are identical in this particular case.

Now, let us consider Hartree's and Fock's equations in the Bloch approximation. As long as the closed shells do not overlap, we should be able to write the Bloch functions in the form

$$\psi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_n e^{2\pi i \mathbf{k} \cdot \mathbf{r}(n)} \psi(\mathbf{r} - \mathbf{r}(n)) \quad (2)$$

where \mathbf{k} is the wave-number vector, $\psi(\mathbf{r} - \mathbf{r}(n))$ is the normalized one-electron function that is centered about the atom at $\mathbf{r}(n)$, and N is the total number of atoms in the lattice. In a filled zone, each value of \mathbf{k} is associated with a pair of electrons having opposite spin. In the following paragraphs of this section, summations over $2\mathbf{k}$ will imply summation over both types of states associated with the N values of \mathbf{k} in a zone.

Hartree's equations for the $\psi_{\mathbf{k}}$ are

$$-\frac{\hbar^2}{2m}\Delta\psi_{\mathbf{k}}(\mathbf{r}_1) + \sum_{2\mathbf{k}'} e^2 \int \frac{|\psi_{\mathbf{k}'}(\mathbf{r}_2)|^2}{r_{12}} d\tau_2 + \sum_n v_n(\mathbf{r}_1) \psi_{\mathbf{k}}(\mathbf{r}_1) = e\psi_{\mathbf{k}} \quad (3)$$

where $\sum_{2\mathbf{k}'}$ extends over all pairs of electrons in the zone, except one of the pairs having wave number \mathbf{k} , and $v_n(\mathbf{r}_1)$ is the ion-core field of the atom at $\mathbf{r}(n)$. The potential

$$\int \frac{|\psi_k|^2}{r_{12}} d\tau_2, \quad (4)$$

arising from a single electron, is negligible at any point in the lattice as long as the crystal contains a large number of atoms. Hence, it is immaterial whether the sum in (3) extends over all electrons or all electrons except one and the primes may be deleted. Using (2), we have

$$\sum_{2k} |\psi_k(r)|^2 = \sum_{2k} \sum_{n,m} e^{2\pi i k \cdot (r(n) - r(m))} \psi(r - r(n)) \psi^*(r - r(m)). \quad (5)$$

The terms for which $r(n)$ differs from $r(m)$ vanish because the ψ do not overlap. Thus,

$$\sum_{2k} |\psi_k(r)|^2 = \sum_n 2|\psi(r - r(n))|^2 \quad (6)$$

where the factor 2 appears because of spin. Outside the m th atom, the potential arising from the term $2|\psi(r - r(m))|^2$ in (6) and the ion-core field $v_n(r)$ cancel one another. Hence, only the term in (6), arising from $|\psi(r - r(n))|^2$, need be considered in the vicinity of the n th atom, for the other terms are canceled by the ion-core terms. Thus, near the n th atom, (3) reduces to

$$-\frac{\hbar^2}{2m} \Delta \psi(r_1 - r(n)) + \left\{ 2e^2 \int \frac{|\psi(r_2 - r(n))|^2}{r_{12}} d\tau_2 + v_n(r_1) \right\} \psi(r_1 - r(n)) = e\psi(r_1 - r(n)). \quad (7)$$

This equation is not the same as Eq. (1), because of the factor 2 which appears in the coulomb integral. It is easy to trace this spurious screening term to the fact that the electrons are completely uncorrelated in the total wave function on which Hartree's equations are based (*cf.* Chap. VI). In the Hartree approximation of Bloch's scheme, the probability of an electron being at a given atom is determined only by the average charge distribution $2|\psi|^2$ of other electrons on the atom. Actually, other electrons tend to stay away from this atom when the given electron is there both because of the electron repulsion and because of exchange.

Let us consider next the Fock approximation. In this case, we have in addition to the terms on the left-hand side of (3), the exchange terms

$$-\sum_{\substack{k' \\ \parallel \text{ spin}}} e^2 \left[\int \frac{\psi_k^*(r_2) \psi_k(r_2)}{r_{12}} d\tau_2 \right] \psi_{k'}(r_1) \quad (8)$$

where the sum extends only over electrons of one kind of spin. This sum need not be primed if the prime is dropped in (3), since the additional terms just cancel one another. Using (2) once again, we find that (3) is

equal to

$$-\frac{e^2}{N^2} \sum_{\mathbf{k}'} \left\{ \sum_n e^{2\pi i \mathbf{k}' \cdot \mathbf{r}(n)} \psi(\mathbf{r}_1 - \mathbf{r}(n)) \right\} \left\{ \sum_m \int \frac{e^{2\pi i (\mathbf{k} - \mathbf{k}') \cdot \mathbf{r}(m)} |\psi(\mathbf{r}_2 - \mathbf{r}(m))|^2 d\tau_2}{r_{12}} \right\}. \quad (9)$$

The cross terms in the second factor of the sum have been dropped because the ψ on different atoms do not overlap. This equation, in turn, may be written in the form

$$-\frac{e^2}{N^2} \sum_{n,m} \sum_{\mathbf{k}'} e^{2\pi i \mathbf{k}' \cdot [\mathbf{r}(n) - \mathbf{r}(m)]} e^{2\pi i \mathbf{k} \cdot \mathbf{r}(m)} \psi(\mathbf{r}_1 - \mathbf{r}(n)) \int \frac{|\psi(\mathbf{r}_2 - \mathbf{r}(m))|^2}{r_{12}} d\tau_2. \quad (10)$$

The sum $\sum_{\mathbf{k}'} e^{i\mathbf{k}' \cdot [\mathbf{r}(n) - \mathbf{r}(m)]}$ is equal to $N\delta_{n,m}$, however, for \mathbf{k}' is summed over the points of a zone. Hence, (8) reduces to

$$-\frac{e^2}{\sqrt{N}} \sum_n \left\{ e^{i\mathbf{k} \cdot \mathbf{r}(n)} \psi(\mathbf{r}_1 - \mathbf{r}(n)) \int \frac{|\psi(\mathbf{r}_2 - \mathbf{r}(n))|^2}{r_{12}} d\tau_2 \right\}. \quad (11)$$

Thus, Fock's equation for $\psi(\mathbf{r}_1 - \mathbf{r}(n))$ is

$$-\frac{\hbar^2}{2m} \Delta \psi(\mathbf{r}_1 - \mathbf{r}(n)) + \left\{ e^2 \int \frac{|\psi(\mathbf{r}_2 - \mathbf{r}(n))|^2}{r_{12}} d\tau_2 + v_n(\mathbf{r}_1) \right\} \psi(\mathbf{r}_1 - \mathbf{r}(n)) =$$

$$e\psi(\mathbf{r}_1 - \mathbf{r}(n)),$$

which is identical with (1). Thus, the exchange integrals remove a part of the spurious screening that occurs in Hartree's equations (7). This fact shows, however, that we cannot expect exchange to compensate for all the inadequacies of Hartree's scheme, even in the simple example discussed above. The exchange correlation affects only electrons with parallel spin and does not alter the spurious screening of electrons with antiparallel spin. The remaining defect may be corrected only by solving the many-body problem by a method that is more accurate than the one-electron approximation.

Let us consider an example in which the atoms do not have closed shells and the zones are not completely filled. We cannot expect the Heitler-London and Bloch schemes to lead to identical results in this case, but we can examine the relative merits of the two. We shall assume that the atoms have a single valence electron outside the closed shells and shall designate the ion-core field for this electron by $V(r)$. The equation for the Heitler-London function associated with the atom

at $\mathbf{r}(n)$ is

$$-\frac{\hbar^2}{2m}\Delta\psi(\mathbf{r} - \mathbf{r}(n)) + V(\mathbf{r} - \mathbf{r}(n))\psi(\mathbf{r} - \mathbf{r}(n)) = e\psi(\mathbf{r} - \mathbf{r}(n)) \quad (12)$$

as long as the atoms do not overlap. Under these conditions, the total energy of the lattice, relative to a system of N ionized atoms is $N\epsilon$, and is independent of the assignment of spins to the electrons. This solution evidently is as accurate as the ion-core field.

In the Bloch approximation, the wave functions have the form (2) and the \mathbf{k} may be assigned spins in many different ways. The \mathbf{k} range over an entire zone if the spins are parallel, for example, but many other arrangements are possible. The total wave function is a single determinant in two cases, namely, the states of highest multiplicity, in which all spins are parallel, and the state of zero multiplicity, in which all spins are paired. In both these cases, Hartree's equations have the form

$$-\frac{\hbar^2}{2m}\Delta\psi(\mathbf{r}_1 - \mathbf{r}(n)) + \left\{ V(\mathbf{r}_1 - \mathbf{r}(n)) + e^2 \int \frac{|\psi(\mathbf{r}_2 - \mathbf{r}(n))|^2}{r_{12}} d\mathbf{r}_2 \right\} \psi(\mathbf{r}_1 - \mathbf{r}(n)) = e\psi(\mathbf{r}_1 - \mathbf{r}(n)), \quad (13)$$

which differs from (12) by spurious screening terms. The exchange terms of Fock's equations remove this term in the case in which all spins are parallel, for then the \mathbf{k} range over an entire zone and the exchange term is identical with (8). On the other hand, if the spins are paired and only the lower half of a zone is filled, the exchange term for $\psi_{\mathbf{k}}$ is [cf. Eq. (9)]

$$-\frac{e^2}{N^2} \sum_{\mathbf{k}'} \left\{ \sum_n e^{2\pi i \mathbf{k}' \cdot \mathbf{r}(n)} \psi(\mathbf{r}_1 - \mathbf{r}(n)) \right\} \left\{ \sum_m \int \frac{e^{2\pi i (\mathbf{k} - \mathbf{k}') \cdot \mathbf{r}(m)} |\psi(\mathbf{r}_2 - \mathbf{r}(m))|^2 d\mathbf{r}_2}{r_{12}} \right\}$$

where \mathbf{k}' is summed over half a zone. This may be rewritten in the form

$$-\frac{e^2}{N^2} \sum_{n,m} \sum_{\mathbf{k}'} e^{2\pi i \mathbf{k}' \cdot [\mathbf{r}(n) - \mathbf{r}(m)]} e^{2\pi i \mathbf{k} \cdot \mathbf{r}(m)} \psi(\mathbf{r}_1 - \mathbf{r}(n)) \int \frac{|\psi(\mathbf{r}_2 - \mathbf{r}(m))|^2}{r_{12}} d\mathbf{r}_2$$

[cf. Eq. (10)]. The summation

$$\sum_{\mathbf{k}'} e^{2\pi i \mathbf{k}' \cdot [\mathbf{r}(n) - \mathbf{r}(m)]}$$

is not equal to $N\delta_{n,m}$, for \mathbf{k}' does not range over an entire zone. If we assume, however, that the atoms are so far apart that the summation is negligible when $\mathbf{r}(n) \neq \mathbf{r}(m)$, (13) reduces to $(N/2)\delta_{n,m}$. Hence, in this

case, Fock's equation for ψ in the Bloch scheme is

$$-\frac{\hbar^2}{2m}\Delta\psi(\mathbf{r}_1 - \mathbf{r}(n)) + \left\{ V + \frac{e^2}{2} \int \frac{|\psi(\mathbf{r}_2 - \mathbf{r}(n))|^2}{r_{12}} d\tau_2 \right\} \psi(\mathbf{r}_1 - \mathbf{r}(n)) = \epsilon \psi(\mathbf{r}_1 - \mathbf{r}(n)). \quad (14)$$

The exchange term cancels only half the spurious screening term. This fact shows that the Bloch scheme may be very inaccurate¹ for some states when the bands are very narrow. Thus, the energy of the lattice in the Bloch approximation is

$$N \frac{e^2}{4} \int \frac{|\psi(\mathbf{r}_1)|^2 |\psi(\mathbf{r}_2)|^2}{r_{12}} d\tau_{12} \quad (15)$$

higher than the energy in the Heitler-London approximation when spins are paired. The integral (15) is equal to

$$\frac{3}{32} \frac{e^2}{a_h},$$

or about 2.9 ev, for hydrogenic 1s functions.

As the atoms are brought nearer and nearer, the terms in (13) for $\mathbf{r}(n) \neq \mathbf{r}(m)$ may be neglected no longer and the exchange terms reduce part of the screening effect of electrons on different atoms. This type of correlation effect does not occur in the Heitler-London scheme. Thus, a part of the advantage of the Heitler-London scheme over the Bloch scheme begins to disappear as atoms begin to overlap. It is for this reason, among others, that the Bloch scheme may be used in competition with the Heitler-London scheme in the computation of binding energies of actual solids.

b. Perfectly Free Electrons.—As a working model for discussing the case of perfectly free electrons, we shall consider a system of N electrons in a box that contains a uniform positive charge distribution of total magnitude Nc . The positive charge compensates for the over-all repulsion of the electrons and makes the system stable. The exchange terms for this system have not been treated in the Heitler-London approximation. Although this solution probably would be very poor kinematically because the metallic properties of the model are not apparent in the Heitler-London approximation (cf. Sec. 66, Chap. VIII), the cohesive energy probably would compare favorably with that derived on the basis of the Bloch approximation.

Since the positive-ion distribution is uniform in our model, the Bloch functions have the free wave form

¹ Caution must be used in applying Bloch functions to discussions of ferromagnetism; for, as we see here, the Bloch approximation gives a spurious ferromagnetism in a case in which states of all spin actually have the same energy.

$$\psi_{\mathbf{k}} = \frac{1}{\sqrt{V}} e^{2\pi i \mathbf{k} \cdot \mathbf{r}} \quad (16)$$

where V is the volume of the crystal. Let us assume that all the lowest electronic levels are doubly occupied. The value, k_0 , of k at the top of the filled region is given by¹

$$k_0 = \left(\frac{3N}{8\pi V} \right)^{\frac{1}{3}}, \quad (17)$$

and the exchange term in the equation for $\psi_{\mathbf{k}}$ is

$$-\frac{e^2}{V} \sum_{\mathbf{k}'} e^{2\pi i \mathbf{k}' \cdot \mathbf{r}_1} \int \frac{e^{2\pi i (\mathbf{k} - \mathbf{k}') \cdot \mathbf{r}_2}}{r_{12}} d\tau_2, \quad (18)$$

which is identical with

$$-\frac{e^2}{V} e^{2\pi i \mathbf{k} \cdot \mathbf{r}_1} \sum_{\mathbf{k}'} \int \frac{e^{2\pi i (\mathbf{k}' - \mathbf{k}) \cdot (\mathbf{r}_1 - \mathbf{r}_2)}}{r_{12}} d\tau_2. \quad (19)$$

This integral is independent of \mathbf{r}_1 and may be evaluated by direct methods.² The result is that (19) reduces to

$$-C_{\mathbf{k}} \frac{e^{2\pi i \mathbf{k} \cdot \mathbf{r}_1}}{\sqrt{V}}$$

where

$$\begin{aligned} C_{\mathbf{k}} &= e^2 k_0 \left(2 + \frac{k_0^2 - k^2}{k_0 k} \log \left| \frac{k + k_0}{k - k_0} \right| \right) \\ &= 0.306 \frac{e^2}{r_s} \left[2 + \frac{1}{\alpha} (1 - \alpha^2) \log \left| \frac{1 + \alpha}{1 - \alpha} \right| \right] \end{aligned} \quad (20)$$

in which $\alpha = k/k_0$ and r_s is defined by the equation

$$\frac{4\pi}{3} r_s^3 = \frac{V}{N}.$$

Hence, (19) is equal to a constant times $\psi_{\mathbf{k}}$, or

$$A\psi_{\mathbf{k}} = -C_{\mathbf{k}}\psi_{\mathbf{k}} \quad (21)$$

for perfectly free electrons, where A is the Dirac exchange operator. The function $C_{\mathbf{k}} r_s / e^2$ is shown in Fig. 5.

The mean value of $-C_{\mathbf{k}}$ for all electrons is equal to twice the exchange energy per electron. We shall evaluate this energy directly.³ The total exchange energy is

¹ Cf. Eq. 20, Sec. 26; also, Sec. 49.

² P. A. M. DIRAC, *Proc. Cambridge Phil. Soc.*, **26**, 376 (1930); J. BARDEEN, *Phys. Rev.*, **49**, 653 (1936).

³ WIGNER and SEITZ, *op. cit.*

$$\begin{aligned}
& -\frac{1}{V^2} \sum_{\mathbf{k}, \mathbf{k}'} e^2 \int \frac{e^{2\pi i(\mathbf{k}' - \mathbf{k}) \cdot (\mathbf{r}_1 - \mathbf{r}_2)}}{r_{12}} d\tau_1 d\tau_2 \\
& \cong -\frac{e^2}{V^2} \int e^{2\pi i(\mathbf{k}' - \mathbf{k}) \cdot (\mathbf{r}_1 - \mathbf{r}_2)} \rho_{\mathbf{k}} \rho_{\mathbf{k}'} d\mathbf{k}' d\mathbf{k} d\tau_1 d\tau_2 \quad (22)
\end{aligned}$$

where $\rho_{\mathbf{k}} = V$ is the density of points in \mathbf{k} space and the summation has been replaced by an integration. This replacement is allowable for

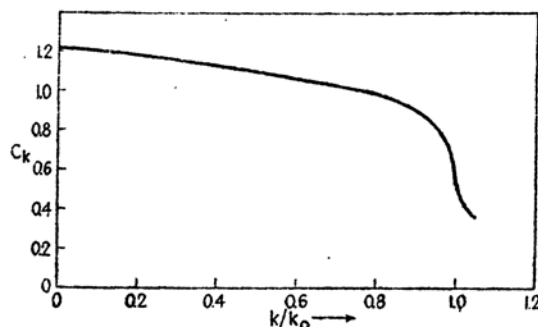


FIG. 5.—The negative of the exchange energy for perfectly free electrons as a function of k/k_0 . The energy units are e^2/r_s .

ordinary-sized crystals since their states are very dense. Integrating over \mathbf{k} and \mathbf{k}' , we find that (22) reduces to

$$-4\pi e^2 \int \frac{(2\pi k_0 r_{12} \cos 2\pi k_0 r_{12} - \sin 2\pi k_0 r_{12})^2}{r_{12}^5} d\tau_1 d\tau_2 = -4\pi e^2 V k_0^4. \quad (23)$$

k_0 may be replaced by its value (17), and then the mean exchange energy per electron is found to be

$$-\frac{e^2}{4\pi^4} n_0^{1/3} \quad (24)$$

where n_0 is the number of electrons per unit volume. In terms of r_s , (24) becomes

$$-0.458 \frac{e^2}{r_s}. \quad (24a)$$

The implications of the exchange terms for perfectly free electrons were discussed in Sec. 49, Chap. VI. We saw there that the exchange terms have the effect of keeping electrons of antiparallel spin apart (cf. Fig. 2, Chap. VI). The absence of such correlational effects for electrons of antiparallel spin in the Fock-Bloch approximation constitutes a large source of error. The Heitler-London approximation will furnish some correlation between electrons of both kinds of spin by keeping them on separate atoms; but as we have seen in molecular problems, such as that of H_2 , this method of introducing correlations is not very accurate. Thus, it

seems reasonable to say that both one-electron schemes have comparable errors when the electronic interactions are appreciable. We shall discuss the correlation terms for the Bloch scheme in the next section.

76. Correlation Correction for Perfectly Free Electrons.—The total correlation correction for the free-electron model in which the positive charge distribution is uniform has been investigated most thoroughly by Wigner.¹ A part of this correction is the exchange energy

$$-0.458 \frac{e^2}{r_s}, \quad (1)$$

which was derived in Sec. 75 (cf. Eq. 24a). We shall be interested here in the additional term that arises primarily from electrons of anti-parallel spin.

The simplest case in which the total correlation may be estimated is that in which the electron density is so low that the electronic kinetic energy is negligible. This is the case in which r_s is very large. The electrons then will form the most stable lattice arrangement, which, according to the Madelung type computations, is a body-centered cubic arrangement. Its energy, relative to the energy of a perfectly uniform negative charge distribution, is

$$-0.746 \frac{e^2}{r_s}. \quad (2)$$

Hence, the correlation correction to the Bloch-Fock scheme, which is the difference between (1) and (2), is

$$-0.288 \frac{e^2}{r_s} \quad (3)$$

for low electron density. The expression that is valid for small values of r_s should approach this asymptotically.

The details of Wigner's calculations for high electron density are too involved for discussion in a book of this type. We shall present only a brief summary of his procedure and results.

In the starting approximation, in which the correlation term (3) is zero, the total electronic wave function may be taken in the form

$$\frac{1}{(N/2)!} \begin{vmatrix} \psi_1(x_1) & \dots & \psi_1(x_{N/2}) \\ \vdots & & \vdots \\ \psi_{N/2}(x_1) & \dots & \psi_{N/2}(x_{N/2}) \end{vmatrix} \begin{vmatrix} \psi_1(y_1) & \dots & \psi_1(y_{N/2}) \\ \vdots & & \vdots \\ \psi_{N/2}(y_1) & \dots & \psi_{N/2}(y_{N/2}) \end{vmatrix} \quad (4)$$

¹ E. WIGNER, *Phys. Rev.*, **46**, 1002 (1934); *Trans. Faraday Soc.*, **34**, 678 (1938).

where the x refer to electrons of one spin, the y to electrons of opposite spin, the ψ are free-electron wave functions, and N is the total number of electrons. This function does not satisfy the Pauli principle, but it has the same energy as though it did. Exchange effects, which are the principal consequence of the Pauli principle in the one-electron approximation, are given correctly by (4) since all terms of given spin are contained in the same determinant.

Wigner considers the following modified form of (4)

$$\begin{vmatrix} \psi_1(x_1; y_1, \dots, y_N) \dots \psi_1(x_N; y_1, \dots, y_N) \\ \vdots \\ \psi_N(x_1; y_1, \dots, y_N) \dots \psi_N(x_N; y_1, \dots, y_N) \end{vmatrix} \begin{vmatrix} \psi_1(y_1) \dots \psi_1(y_N) \\ \vdots \\ \psi_N(y_1) \dots \psi_N(y_N) \end{vmatrix} \quad (5)$$

in order to obtain functions from which to construct a better total wave function. Here, the $\psi(y_i)$ are the wave functions for free electrons. The $\psi(x_i, y_1, \dots, y_N)$ are to be determined by the condition that the mean energy of (5) should be a minimum. The ψ are then used in a new total wave function from which a new total energy may be computed. This new total wave function evidently will not be a rigorous solution of the complete Schrödinger equation, but it is a closer approximation than the function based on (4).

The correlation energy obtained by Wigner in this way is accurate only for high electron densities because of his approximate computational methods. He found that this result may be joined to (3) by the function

$$-\frac{0.288e^2}{r_s + 5.1a_k}, \quad (6)$$

which he estimates is accurate to within 20 per cent.

An expression similar to (20) in the preceding section that will give correlation energy as a function of k has not been developed. However, we may compute the correlation energy of the uppermost electron in the filled band. Suppose that a number of electrons are removed from the top of the band so that the total number is equal to N_0 instead of N , the total number in a neutral lattice. The correlation energy will change as a result, and the new value may be derived by taking into account the change in density, that is, by changing r_s to the value $(N/N_0)^{1/3}r_s$. Hence, if the total correlation energy is expressed in the form

$$E_c = N_0 g\left(\left(\frac{N}{N_0}\right)^{1/3} r_s\right) \quad (7)$$

where, according to (6), g is

$$g(r) = -e^2 \frac{0.288}{r + 5.1a_A}, \quad (8)$$

the correlation energy of the upper electrons is

$$\left(\frac{\partial E_c}{\partial N_s} \right)_{N_s=N} = g(r_s) - \frac{1}{3} g'(r_s) r_s. \quad (9)$$

This expression may be used in considering energy changes during a process in which an electron is removed or added to a system, such as during thermionic or photoelectric emission or when an electron jumps from the conduction band to a vacant inner-shell level during X-ray emission.

CHAPTER X

THE COHESIVE ENERGY

77. Introduction*.—The degree to which the computed energy of a system agrees with an accurate observed value is a measure of the accuracy of the wave functions that are used in the theoretical computations, because of the variational theorem. For this reason, computations of the cohesive energies of solids occupy an important position in the development of the theory of solids. The existing calculations deal with simple substances, such as the monovalent metals, the alkali halides and hydrides, and rare gas solids, all of which will be discussed in this chapter under three headings: metals, ionic crystals, and molecular crystals. There have been no accurate computations that deal with valence crystals, such as diamond.

Before beginning the detailed discussion of cases, we shall derive several useful equations. The cohesive energy of a solid is defined as the difference between the energy of the crystal in the normal, bound state, at absolute zero of temperature, and the energy of the isolated atoms or molecules of which it is composed. If the surface energy is neglected, the cohesive energy is proportional to the total number of atoms or molecules in the lattice and may be expressed conveniently in units of electron volts per molecule or in the thermochemist's unit of kilogram-calories per mol. This energy is equal to the actual heat of sublimation only when the substance evaporates into the atomic or molecular constituents to which the separated system is referred.

Let us derive an approximate expression for the cohesive energy in the general case in which there are ν atoms per molecule and m molecules per unit cell of the crystal. Then, if E_A^β ($\beta = 1, \dots, \nu$) is the energy of the β th neutral atom and Ψ is the complete electronic wave function for N unit cells of the crystal, the cohesive energy E_o , relative to a system of separated atoms, is

$$E_o = mN \sum_{\beta=1}^{\nu} E_A^\beta - \frac{1}{Nm} \int \Psi^* H \Psi d\tau(x_1, \dots, x_n, \xi_1, \dots, \xi_n) \quad (1)$$

where H is the complete Hamiltonian of the crystal and n is the total number of electrons. We shall develop this in the important case in which Ψ is expressed in terms of the solutions of Fock's or Hartree's equations.

If the one-electron functions for the β th atom are $\varphi_k^\beta(x, y, z)$, where k ranges over the n_β electrons of the atom, and if the total wave function of the atom is $\Psi_A^\beta(x_1, \dots, z_n, \zeta_1, \dots, \zeta_{n_\beta})$, the energy of the atom is

$$E_A^\beta = \sum_{k=1}^{n_\beta} \int \varphi_k^{*\beta} \left(-\frac{\hbar^2}{2m} \Delta - \frac{Z_\beta e^2}{r_{\beta k}} \right) \varphi_k^\beta d\tau(x, y, z, \zeta) + \frac{1}{2} \sum_{k,l} \int \Psi_A^{*\beta} \frac{e^2}{r_{kl}} \Psi_A^\beta d\tau(x_1, \dots, \zeta_{n_\beta}). \quad (2)$$

The notation that is used in this equation is the same as that which was introduced in Chap. VI. $-Z_\beta e^2/r_{\beta k}$ is the potential of an electron in the field of the nucleus, whose charge is $Z_\beta e$, and e^2/r_{kl} is the interaction potential of the k th and l th electrons. The last term in (2) may be expressed as a sum of coulomb and exchange integrals of the type

$$I_{kl}^\beta = e^2 \int \frac{|\varphi_k(\mathbf{r}_1)|^2 |\varphi_l(\mathbf{r}_2)|^2}{r_{12}} d\tau_1' d\tau_2' \quad (3a)$$

and

$$K_{kl}^\beta = e^2 \int \frac{\varphi_k^*(\mathbf{r}_1) \varphi_l^*(\mathbf{r}_2) \varphi_k(\mathbf{r}_2) \varphi_l(\mathbf{r}_1)}{r_{12}} d\tau_1' d\tau_2' \quad (3b)$$

which have coefficients depending upon the multiplicity of Ψ_A . We shall not be concerned with the numerical values of these coefficients at present but shall express the two sums in the form

$$\frac{1}{2} \sum_{k,l} \alpha_{kl}^\beta I_{kl}^\beta \quad \text{and} \quad -\frac{1}{2} \sum_{k,l} \beta_{kl}^\beta K_{kl}^\beta. \quad (4)$$

The wave functions of those g_β electrons on the β th atom that belong to rigid closed shells are practically unchanged in passing from the free to the bound state. We shall designate the terms in (2) and (4) that involve only these g_β electrons by the symbol \sum_g^β . Similarly, the terms that involve these electrons and the remaining $n_\beta - g_\beta$ electrons on the atom will be designated by $\sum_{g, n-g}^\beta$, and the terms that involve only the $n_\beta - g_\beta$ electrons outside closed shells will be designated by \sum_{n-g}^β . The first set of terms \sum_g^β is canceled by the similar term of Eq. (1) that appears in the expression for the total energy of the solid and may be dropped from consideration. The two other terms are

$$\sum_{g, n-g}^{\beta} = \sum_{k=1}^g \sum_{l=g_{\beta}}^{n_g} (\alpha_{kl}^{\beta} I_{kl}^{\beta} - \beta_{kl}^{\beta} K_{kl}^{\beta}), \quad (5)$$

$$\sum_{n-g}^{\beta} = \sum_{k=g}^{n_g} \int \varphi_k^{\beta*} \left(-\frac{\hbar^2}{2m} \Delta - \frac{Z_{\beta} e^2}{r_{\beta k}} \right) \varphi_k^{\beta} d\tau(x, y, z) + \frac{1}{2} \sum_{k, l=g}^{n_g} (\alpha_{kl}^{\beta} I_{kl}^{\beta} - \beta_{kl}^{\beta} K_{kl}^{\beta}). \quad (6)$$

The first of these expressions is the energy of the interaction between the valence and closed-shell electrons; the second is the total energy of the valence electrons, minus this interaction. In the two cases in which the one-electron functions are solutions of either Hartree's or Fock's equations, namely,

$$\begin{aligned} H^H \varphi_i^{\beta} &= e^{\beta}(H) \varphi_i^{\beta}, \\ H^F \varphi_i^{\beta} &= e^{\beta}(F) \varphi_i^{\beta}, \end{aligned} \quad (7)$$

respectively, it may be verified readily that the sum of (5) and (6) reduces to

$$\begin{aligned} E_A^{\beta}(H) &= \sum_{k=g}^{n_g} e_k^{\beta}(H) - \frac{1}{2} \sum_{k, l=g}^{n_g} (\alpha_{kl}^{\beta} I_{kl}^{\beta} - \beta_{kl}^{\beta} K_{kl}^{\beta}), \\ E_A^{\beta}(F) &= \sum_{k=g}^{n_g} e_k^{\beta}(F) - \frac{1}{2} \sum_{k, l=g}^{n_g} (\alpha_{kl}^{\beta} I_{kl}^{\beta} - \beta_{kl}^{\beta} K_{kl}^{\beta}). \end{aligned} \quad (8)$$

Let us now designate the one-electron wave functions of the valence electrons in the crystal by φ_i , for which i ranges over values from 1 to n' , where n' is the total number of valence electrons. We shall assume that the total wave function Ψ has unit multiplicity. The total energy of the crystal then is (Chap. VI)

$$\begin{aligned} E &= \int \Psi^* H \Psi d\tau = \sum_{i=1}^n \int \varphi_i^*(\mathbf{r}_1) \left[-\frac{\hbar^2}{2m} \Delta + V(\mathbf{r}_1) \right] \varphi_i(\mathbf{r}_1) d\tau_1 + \\ &\quad \frac{1}{2} \sum_{i, j=1}^n \int \frac{|\varphi_i(\mathbf{r}_1)|^2 |\varphi_j(\mathbf{r}_2)|^2}{r_{12}} d\tau_{12} - \frac{1}{2} \sum_{i, j}^n e^2 \int \frac{\varphi_i^*(\mathbf{r}_1) \varphi_j^*(\mathbf{r}_2) \varphi_i(\mathbf{r}_2) \varphi_j(\mathbf{r}_1)}{r_{12}} d\tau_{12} + \\ &\quad \frac{1}{2} \sum_{\alpha, \beta} \frac{Z_{\alpha} Z_{\beta} e^2}{r_{\alpha\beta}} + \frac{1}{2} \sum_{\alpha, \beta} V_{\alpha\beta} + \sum \epsilon_{\alpha i} \end{aligned} \quad (9)$$

in which the self-energy of the rigid ion cores is neglected. Here, V is the coulomb ion-core field, $V_{\alpha\beta}$ is the exchange interaction between ion α

and ion β , and $\epsilon_{\alpha i}$ is the exchange interaction between φ_i and the α th ion. Just as in Eqs. (8), the first term in (9) may be expressed in terms of the energy parameters of Hartree's or Fock's equations and the exchange and coulomb integrals. The energy per molecule of the crystal may then be obtained by dividing (9) by Nm , and an approximate expression for the cohesive energy per molecule may be obtained by subtracting (9) from the expression for the total energy of the free atoms that is derived by adding terms of the type (8).

Thus, the cohesive energy may be expressed in terms of energy parameters and coulomb and exchange integrals for the atomic and crystalline systems in the Hartree or Fock approximations. Since the practical difficulties of evaluating these parameters and integrals are very large, progress has been made only in those cases in which simple approximations methods, such as those described in the last chapter, can be used. There is a tendency for the errors that are introduced by use of the one-electron approximation to compensate one another when this approximation is employed in both the atomic and crystalline states, for the computed energies are too high in both cases. Thus, the computed cohesive energy may be larger or smaller than the correct value, depending upon whether or not the correlational error in the atomic state is larger than that of the solid. In solids such as ionic and molecular crystals in which the Heitler-London approximation can be used for both the crystalline and the atomic states, the correlation error is nearly the same in both cases, and we may expect to obtain a good value of the cohesive energy. On the other hand, the correlation energy of the valence electron on a free alkali atom is much smaller than that of the valence electrons in the corresponding metal. Hence, the one-electron approximation cannot be expected to give a good value of the cohesive energy in this case.

A. METALS

78. The Alkali Metals.—The cohesive energies of the three alkali metals lithium, sodium, and potassium have been computed¹ to a similar degree of accuracy by using the approximate methods that were described in the last chapter. We shall discuss the three alkali metals at the same time.

a. The Ion-core Field.—The first step in computing the cohesive energy of any substance is to obtain an ion-core field that takes into account the

¹ Li: F. SEITZ, *Phys. Rev.*, **47**, 400 (1935); J. BARDEEN, *Jour. Chem. Phys.*, **6**, 367 (1938).

Na: E. WIGNER and F. SEITZ, *Phys. Rev.*, **43**, 804 (1933); **46**, 509 (1934). E. WIGNER, *Phys. Rev.*, **46**, 1002 (1934). BARDEEN, *op. cit.*

K: E. GORIN, *Physik. Z. Sowj.*, **9**, 328 (1936).

interaction between the closed-shell electrons and the valence electrons. In both lithium and sodium, it was found possible to construct a radial potential field $v_c(r)$ having the property that the eigenvalues of the equation

$$-\frac{\hbar^2}{2m}\Delta\psi + v_c(r)\psi = E\psi \quad (1)$$

closely reproduce the observed atomic-term values. From the standpoint of Fock's equations, it is possible to say that in these cases the exchange interaction between core and valence electrons may be replaced by an ordinary potential term. In lithium, for example, the best field that could be obtained by a trial-and-error method duplicated the atomic values to within a few tenths of a per cent. The field for sodium, which was derived by Prokofjew¹ for another purpose, leads to term values that agree with the observed ones to within about 1 per cent. Thus, the computed atomic 3s function has an energy of 0.381 Rydberg unit, whereas the observed value is 0.378. Gorin attempted to construct a similar field for potassium, but he found that this could not be done with sufficient accuracy. Presumably, in this case the exchange terms cannot be replaced even approximately by ordinary potential terms. As a result, Gorin used a Hartree ion-core field and evaluated the exchange integrals between the valence and core electrons by direct methods. The ionization potential that he obtained in this way is 0.2934 Rydberg unit, which should be compared with the measured value of 0.3190 Rydberg unit.

b. Application of the Cellular Approximation.—All the alkali metals form body-centered cubic lattices for which the polyhedron that should be used in the cellular approximation is the truncated octahedron shown in Fig. 2, Chap. IX. It may be assumed, for simplicity, that these polyhedra can be replaced by spheres of equal volume. The error that is made in doing so can be shown to be negligible and will be discussed below. Since each of the spheres is electrostatically neutral, the coulomb potential in a given cell that arises from any other cell is zero. Hence, all that is necessary is to consider the coulomb field arising from the charge in a given cell in the sphere approximation. For this reason, we shall restrict the following discussion to a single sphere.

When deriving the electronic wave function within a sphere, we may neglect the potential of the valence-electron distribution in the first approximation. This procedure is permissible because the electronic charge distribution turns out to be very nearly constant and its potential is a slowly varying function that may be included readily by perturbation methods in a later approximation.

¹ W. PROKOFJEW, *Z. Physik*, **58**, 255 (1929).

Thus, the first problem to be solved in determining electronic wave functions is that of finding solutions of Eq. (1) that satisfy boundary conditions implied by the Bloch form

$$\psi_{\mathbf{k}} = \chi_{\mathbf{k}} e^{2\pi i \mathbf{k} \cdot \mathbf{r}} \quad (2)$$

(cf. Chap. VIII). We may expect the lowest eigenfunction to be one for which \mathbf{k} is zero and which has cubic symmetry relative to any nucleus. The lowest order surface harmonics that possess this symmetry are s and g functions of the type

$$s = 1$$

$$g = \frac{1}{r^4} \left[(x^2 y^2 + y^2 z^2 + z^2 x^2) - \frac{1}{3} (x^4 + y^4 + z^4) \right]. \quad (3)$$

Hence, ψ_0 should have approximately the form

$$\psi_0 = f_s(r) + g f_g(r) \quad (4)$$

where f_s and f_g are radial functions. The equations for continuity of the normal gradient of ψ at the points (100) and (111) on the sphere of radius r_s are

$$\begin{aligned} f'_s(r_s) - \frac{1}{3} f'_g(r_s) &= 0, \\ f'_s(r_s) + \frac{2}{3} f'_g(r_s) &= 0, \end{aligned} \quad (5)$$

where r_s is the radius of the sphere and the prime indicates differentiation with respect to r . The solutions of (5) are either

$$\psi_0 = f_s(r) \quad \text{with} \quad f'_g(r_s) = 0 \quad (6)$$

or

$$\psi_0 = g f_g(r_s) \quad \text{with} \quad f'_s(r_s) = 0. \quad (7)$$

Since an s function should have lower energy than a g function that satisfies the same boundary conditions, we should use the first of these conditions for the lowest state. If the actual polyhedron were used instead of a sphere, each of the two equations (6) would involve different values of r_s , one of which would be the distance from the center of the polyhedron to a point on the surface in the (100) direction, and the other of which would be the distance to a point in the (111) direction. We should then obtain two solutions that involve both the s and g functions. One of these, however, would be predominantly an s function and the other predominantly a g function.

We may conclude that within the sphere ψ_0 is that s function which goes into the lowest atomic s function when the lattice is expanded and which satisfies the condition $\psi'_s(r_s) = 0$. A relative scale plot of this function for the value of r_s corresponding to the actual lattice constant is shown in Fig. 1 for sodium. The energy of ψ_0 as a function of r_s ,

$\epsilon_0(r_s)$, is shown in Fig. 2 for the three alkali metals. The full curve for potassium includes the exchange interaction between the valence and closed-shell electrons. These energy curves resemble very closely the

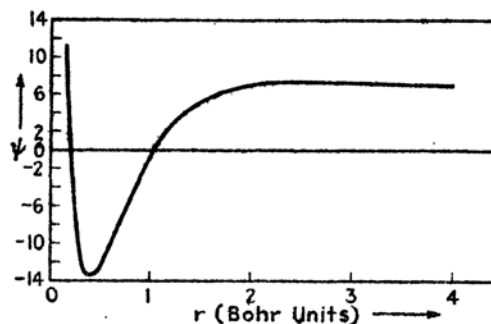


FIG. 1.—The lowest wave function of metallic sodium. It should be noted that this function is practically constant for over ninety per cent of the atomic volume.

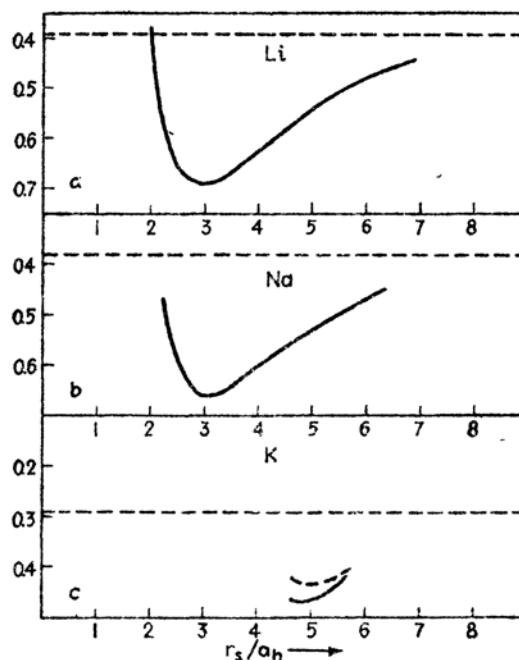


FIG. 2.—The $\epsilon(r_s)$ curves for *a* lithium, *b* sodium, and *c* potassium. The energy scale is in Rydberg units (one Rydberg unit equals 13.54 ev). The horizontal dotted lines represent the normal-state atomic energies. The dashed curve for potassium represents the case in which the valence-electron closed-shell exchange energy is neglected. Exchange is included in the full curve.

characteristic energy versus internuclear distance plots for diatomic molecules and show that the stability of metals is related to the fact that the spatial distribution of potential in a solid allows some of the electrons

to have a lower energy than they have in the free atom. A part of this decrease in energy is connected with the decrease in kinetic energy that may be associated with increased smoothness of the wave functions, and part is connected with the fact that the center of gravity of the electronic charge in a given cell is nearer the nucleus.

Only two electrons are in the state going with $\mathbf{k} = 0$. In order to find the other wave functions in first approximation, it is necessary to solve (1) for those cases in which \mathbf{k} is not zero. Since only half of the first zone is filled in the normal state of the alkali metals, it is not necessary¹ to find the exact form of the $\epsilon(\mathbf{k})$ curve near the boundary of this zone in order to compute the cohesive energy. With this in mind, we shall compute $\psi_{\mathbf{k}}$ and $\epsilon(\mathbf{k})$ by a method that is accurate near the center of the zone but not accurate near the boundary.

If we substitute (2) in Eq. (1), we find that $\chi_{\mathbf{k}}$ must satisfy the equation

$$-\frac{\hbar^2}{2m}\Delta\chi_{\mathbf{k}} + v_0\chi_{\mathbf{k}} - \frac{\hbar^2}{m}2\pi i\mathbf{k} \cdot \text{grad } \chi_{\mathbf{k}} = \epsilon'(\mathbf{k})\chi_{\mathbf{k}} \quad (8)$$

where

$$\epsilon'(\mathbf{k}) = \epsilon_0(\mathbf{k}) - \frac{\hbar^2}{2m}\mathbf{k}^2. \quad (9)$$

The solution of (8), as given by the Rayleigh-Schrödinger perturbation method,² is

$$\epsilon'(\mathbf{k}) = \epsilon_0 - \frac{\hbar^2}{m}2\pi i \int \psi_0^* \mathbf{k} \cdot \text{grad } \psi_0 d\tau + \frac{\hbar^4}{m^2}4\pi^2 \sum_{\nu} \frac{|\int \psi_{\nu}^* \mathbf{k} \cdot \text{grad } \psi_0 d\tau|^2}{\epsilon_0 - \epsilon_{\nu}} + \dots, \quad (10)$$

$$\chi_{\mathbf{k}} = \psi_0 - \frac{\hbar^2}{m}2\pi i \sum_{\nu} \psi_{\nu} \frac{\int \psi_{\nu}^* \mathbf{k} \cdot \text{grad } \psi_0 d\tau}{\epsilon_0 - \epsilon_{\nu}} + \dots, \quad (11)$$

where the ψ_{ν} are the solutions of (8) for $\mathbf{k} = 0$ and ψ_0 is the lowest s function. The first integral in (10) is zero; the second term may be simplified in the following way. The only function ψ_{ν} for which $\int \psi_{\nu}^* \mathbf{k} \cdot \text{grad } \psi_0 d\tau$ does not vanish is one having p symmetry, that is, one having any one of

¹ As we see from Fig. 17, p. 300, the closest point of approach to the zone boundary occurs in the (110) direction of wave-number space. Since the ratio of k_0 to the value of k at the zone is only 0.88 in this direction, it is possible that a small correction for deviations from the free-electron $\epsilon(\mathbf{k})$ curve should be made (cf. footnote 2, p. 366).

² Bardeen, *op. cit.*, has developed an alternative method for solving Eq. (7) which involves the assumption that the solution may be expressed in the form

$$\chi_{\mathbf{k}} = f(r) + \mathbf{k} \cdot \mathbf{r}g(r)$$

where $f(r)$ and $g(r)$ are radial functions within each sphere. Evidently, $f(r)$ is ψ_0 and $g(r)$ is $f_p(r)$ in Eq. (18).

the forms

$$\frac{x}{r}\varphi_r(r), \quad \frac{y}{r}\varphi_r(r), \quad \frac{z}{r}\varphi_r(r), \quad (12)$$

where $\varphi_r(r)$ is a radial function. In each of these three cases, the matrix components are, respectively,

$$k_x \int \frac{x^2}{r^2} \varphi_r(r) \psi_0(r) d\tau, \quad k_y \int \frac{y^2}{r^2} \varphi_r(r) \psi_0(r) d\tau, \quad k_z \int \frac{z^2}{r^2} \varphi_r(r) \psi_0(r) d\tau. \quad (13)$$

Since the integrals in these terms are equal, Eqs. (9) and (10) may be written in the form

$$\epsilon'(\mathbf{k}) = \epsilon_0 + \frac{\hbar^4}{m^3} \mathbf{k}^2 \sum_r \frac{64\pi^4}{9} \frac{|\int \varphi_r \psi_0' r^2 dr|^2}{\epsilon_0 - \epsilon_r} \quad (14)$$

$$\chi_{\mathbf{k}} = \psi_0 - \frac{\hbar^2}{m} 2\pi i \sum_r \mathbf{k} \cdot \mathbf{r} \varphi_r \frac{4\pi}{3} \frac{\int \varphi_r \psi_0' r^2 dr}{\epsilon_0 - \epsilon_r} \quad (15)$$

The p functions that are satisfactory solutions of (1) in the sphere approximation are those which satisfy the boundary condition $\varphi_r(r_s) = 0$. This fact may be proved by setting up equations similar to (5). When these p functions have been computed, both $\epsilon'(\mathbf{k})$ and $\psi_{\mathbf{k}}$ may be determined by evaluating the integrals in (14) and (15). It is not evident from what has been said that the higher order perturbation terms are negligible, but a practical examination of these for all three metals shows that they actually are so for values of \mathbf{k} in the first zone that are not too near the zone boundary.

It is worth noting that, in the approximation in which Eq. (14) is valid, $\epsilon(\mathbf{k})$ may be written as

$$\epsilon(\mathbf{k}) \equiv \epsilon'(\mathbf{k}) + \frac{\hbar^2}{2m} \mathbf{k}^2 = \epsilon_0 + \frac{\hbar^2}{2m^*} \mathbf{k}^2 \quad (16)$$

where¹

$$\frac{1}{m^*} = \frac{1}{m} \left(1 + \frac{\hbar^2}{m} \sum_r \frac{32\pi^2}{9} \frac{|\int \varphi_r \psi_0' r^2 dr|^2}{\epsilon_0 - \epsilon_r} \right). \quad (17)$$

Similarly, $\psi_{\mathbf{k}}$ may be expressed in the form

$$\psi_{\mathbf{k}} = e^{2\pi i \mathbf{k} \cdot \mathbf{r}} (f_s(r) + 2\pi i \mathbf{k} \cdot \mathbf{r} f_p(r)) \quad (18)$$

¹ We shall see in Chap. XVII that the terms in the coefficient of $1/m$ in Eq. (17) are related to the f factors of radiation theory.

where f_s is the lowest s function and

$$f_p(r) = -\frac{\hbar^2}{2m} \sum_p \varphi_p(r) \frac{4\pi}{3} \frac{\int \varphi_p \psi_0' r^2 dr}{\epsilon_0 - \epsilon_p}. \quad (19)$$

The radial integrals that appear in the preceding equations were determined for several values of r , for each of the three metals. The values of the quantity m/m^* that were computed¹ from these [cf. Eq. (17)] are listed in Table LII. It turns out in the case of sodium that the

TABLE LII.—VALUES OF m/m^* FOR THE ALKALI METALS
(The values for the observed values of r , are given in boldface type.)

Li		
r/a_h		m/m^*
3.00		0.584
3.21		0.653
3.32		0.684
Na		
3.80		1.079
3.96		1.069
4.12		1.059
K		
4.82		1.72
5.06		1.59
5.34		1.48
5.47		1.44

summation in (17) "accidentally" vanishes for values of r , in the vicinity of the observed one, so that this ratio is practically unity. The ratio is less than unity in lithium and greater than unity in potassium. A small term that takes into account the variation with k of the interaction between the valence electrons and the closed-shell electrons has been included in the case of potassium.

The expression (16) for $\epsilon(k)$ is the same as the expression for the energy of free electrons that was used in the theory of metals discussed in Chap. IV, with the difference that ϵ_0 replaces $-W_a$. The quantity m^* is, as before, the effective electron mass. Since this mass is practically equal to the ordinary electronic mass in the case of sodium, this metal should behave more like an ideal metal than either lithium or potassium. The same point is indicated by the fact that ψ_0 for sodium (cf. Fig. 1) is almost constant in about 90 per cent of the volume of the sphere, a fact which shows that the wave functions are closely equal to $Ae^{2\pi i k r}$, where A is a constant.

As long as Eq. (16) is valid, the electrons in their normal state completely occupy a sphere centered about the origin of k space, just as in the

¹ The values for Na and Li were computed by Bardeen using the method discussed in footnote 2, p. 352.

Sommerfeld theory. The radius k_0 of the occupied sphere of points in \mathbf{k} space is given by the relation

$$\frac{4\pi}{3}k_0^3V = \frac{N}{2} \quad (20)$$

where V is the volume of the crystal, N is the total number of electrons, and the 2 in the denominator of the right-hand side arises from spin degeneracy. The mean value of

$$\frac{\hbar^2}{2m^*}\mathbf{k}^2 \quad (21)$$

for these electrons will be called the mean Fermi energy in the following discussion, since it corresponds to the mean kinetic energy $\bar{\epsilon}$ that the electrons would possess if they were distributed according to Fermi-Dirac statistics. Actually, (21) is a combination of kinetic and potential energy so that the mean value of this quantity is not the mean value of $-\hbar^2\Delta/2m$. The mean Fermi energy is easily found to be

$$\epsilon_F = \frac{3}{5} \frac{\hbar^2}{2m^*} k_0^2 = \frac{3}{10} \frac{\hbar^2}{m^*} \left(\frac{3n_0}{8\pi} \right)^{\frac{1}{3}} \quad (22)$$

where $n_0 = N/V$ is the number of electrons per unit volume. This is identical with the expression for $\bar{\epsilon}$ that was derived in Sec. 26, Chap. IV. Since $n_0 = 1/(4\pi r_s^3/3)$, Eq. (22) is

$$\epsilon_F = \frac{3}{10} \frac{\hbar^2}{m^*} \left(\frac{9}{32\pi^2} \right)^{\frac{1}{3}} \frac{1}{r_s^2} = \frac{1.105}{r_s^2} \frac{e^2}{a_h} \left(\frac{m}{m^*} \right). \quad (23)$$

Values of this energy are listed in Table LIII.

It is interesting to note that the quantity

$$-(\epsilon_0 + \epsilon_F + \epsilon_i), \quad (24)$$

where ϵ_i is the negative of the atomic energy, that is, the ionization energy, agrees closely with the cohesive energies in the cases of lithium and sodium. In the first case, (24) is 39 kg cal/mol, and the observed value is also 39; in the second case, (24) is 24.4 kg cal/mol, and the observed value is 26. These computed values are given for the observed values of r_s at which (24) actually does have a minimum. The agreement is not so close in the case of potassium, which we shall discuss separately below.

In taking (24) to be the cohesive energy per atom, it is effectively assumed that the field acting upon an electron is essentially that of the ion core in the polyhedron in which the electron is momentarily found. Thus, it is assumed that no more than one electron can be in a given cell at one time. The agreement between (24) and the observed cohesive

energies in the cases of lithium and sodium suggests that on the average the electrons actually do avoid one another in this way.

The cohesive energy of potassium, computed from Eq. (24), is 6 kg cal/mol, whereas the observed value is 22.6. This discrepancy is surprising at first, for we might expect the properties of the alkali metals to vary continuously as we pass down the periodic chart from lithium to cesium. Gorin believes that the error in potassium is related to flaws in the Hartree field on which the computations are based, for this field does not reproduce the atomic energy levels with the same accuracy as the fields used for lithium and sodium (*cf.* part *a* of this section) even when exchange terms are included. Thus, the error in the lowest level of the atom is 0.735 ev without exchange and 0.347 ev with exchange.

Since the exchange terms are larger in the solid than in the free atom, because the center of gravity is nearer the nucleus in the solid, it seems

TABLE LIII.—VALUES OF $\epsilon_0 + \epsilon_I$ AND ϵ_F FOR THE ALKALI METALS
(The values of ϵ_I , the ionization energy, are theoretical values.)

r_s	$\epsilon_0 + \epsilon_I$	ϵ_F	$-(\epsilon_0 + \epsilon_I + \epsilon_F)$	Observed cohesive energy
Li($\epsilon_I = 5.365$ ev = 123.4 kg cal/mol)				
3.00	-84.6 kg cal/mol	44.7	39.9	39 kg cal/mol
3.21	-82.6	43.6	39.0	
3.32	-81.5	42.7	38.8	
Na($\epsilon_I = 5.159$ ev = 118.7 kg cal/mol)				
3.80	-75.3	51.4	23.9	26
3.96	-71.3	46.9	24.4	
4.12	-67.2	42.9	24.3	
K($\epsilon_I = 3.973$ ev = 91.4 kg cal/mol)				
4.82	22.6
5.06	-41.0	42.7	-1.7	
5.34	-40.6	35.7	4.9	
5.47	-39.7	33.1	6.6	

reasonable to expect that the entire correlational effect between the valence and closed-shell electrons increases in passing from the free atom to the solid. Now the exchange and correlation interaction energies of the valence and closed-shell electrons in the free atom are, respectively, 0.388 ev and 0.347 ev. Gorin attempted to correct the absolute error of the lowest level in the solid by multiplying the closed-shell valence-electron exchange energy of the solid by a factor

$$\frac{(0.388 + 0.347)}{0.388} = 1.89.$$

The correction induced by this is listed in Table LIV and increases the binding energy, as computed by Eq. (24), to about 14.5 kg cal/mol for the

TABLE LIV.—CORRECTED VALUES OF $\epsilon_0 + \epsilon_I$ FOR POTASSIUM OBTAINED BY INCREASING CLOSED-SHELL VALENCE-ELECTRON EXCHANGE BY A FACTOR 1.89
($\epsilon_I = 4.333$ ev = 99.7 kg cal/mol)

	$\epsilon_0 + \epsilon_I$	ϵ_F	$-(\epsilon_0 + \epsilon_I + \epsilon_F)$	Observed cohesive energy
4.82	22.6
5.06	-51.5	42.5	9.0	
5.34	-48.1	34.5	13.6	
5.47	-46.2	31.9	14.3	

point $r_s = 5.4$ where the total energy is a minimum. The agreement between the observed and calculated energies is now comparable with that found above for lithium and sodium. The value of the lattice parameter at which the minimum occurs is much too large, however, a fact showing that the corrections of Table LIV are not adequate for small interatomic distances.

c. The Influence of Coulomb Terms.—We shall now proceed to correct the equations in part *b* by considering the effects of the coulomb term

$$\sum_{\mathbf{k}} e^2 \int \frac{|\psi_{\mathbf{k}}(\mathbf{r}_2)|^2 d\mathbf{r}_2}{r_{12}}. \quad (25)$$

We could evaluate this term, using the wave function (9), and determine a new set of functions to replace (9) by placing the result in Eq. (1). The new solutions then could be employed in a reiteration of this procedure and the process repeated until a self-consistent set of functions is found. Actually, the change in the wave functions is negligibly small at the end of the first step in this procedure. The charge distribution

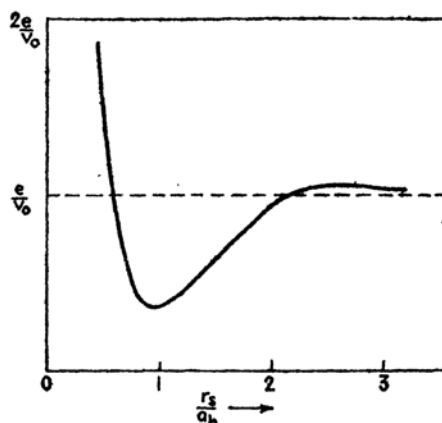


FIG. 3.—The charge distribution in the unit sphere of lithium. The ordinates are expressed in units of e/v_0 , where v_0 is the atomic volume.

$$\sum_{\mathbf{k}} |\psi_{\mathbf{k}}|^2 = \sum_{\mathbf{k}} |\chi_{\mathbf{k}}|^2, \quad (26)$$

obtained from the functions (10), is shown in Fig. 3 for lithium when $r_s = r_0$. It should be noted that this function is practically constant over the major part of the unit sphere. This condition is satisfied even better in sodium and not quite so well in potassium. Since the potential that arises from a constant charge distribution turns out to have a negligible effect on the wave functions (10), as we shall see, we may conclude that the effect will also be negligible when (26) is used. The electronic potential at a point r , arising from a spherical charge distribution of constant density $-e/v_0$ where v_0 is the volume of the sphere, is

$$V(r) = \frac{3}{2} \frac{e^2}{r_s} - \frac{1}{2} \frac{e^2 r^2}{r_s^3} \quad (r \leq r_s). \quad (27)$$

The constant term does not affect the wave function, and so we need consider only the term in r^2 . The correction to ψ_0 arising from this term is

$$-\frac{e^2}{2r_s^3} \sum_k' \psi_k \frac{\int \psi_k^* r^2 \psi_0 d\tau}{\epsilon_0 - \epsilon_k} \quad (28)$$

where the ψ_k obviously form the family of s functions that satisfy the condition $\psi'(r_s) = 0$. An upper limit to the value of the integrals $\int \psi_k^* r^2 \psi_0 d\tau$ may be obtained by evaluating the quantity

$$\sum_k' \left| \int \psi_0^* r^2 \psi_k d\tau \right|^2 = \int \psi_0^4 r^4 d\tau - \left| \int \psi_0^2 r^2 \psi_0 d\tau \right|^2.$$

Since ψ_0 is practically constant, this is

$$\left(\frac{3}{4} - \frac{3}{32}\right)r_s^4 = \frac{1^2}{176}r_s^4.$$

Thus, the maximum value of any of the integrals in (28) is $0.13e^2/r_s$. The difference between the first two s levels in the alkali metals is of the order of $e^2/2a_0$. Hence, the ratio of the coefficient of any ψ_k to that of ψ_0 is at most $0.26/r_s$, which varies from 0.08 to 0.05 for the three alkali metals. Actually, only the lowest of the ψ_k has a coefficient of this magnitude; the others play a less important role because the energy denominators increase and the value of the integral $\int \psi_k^* r^2 \psi_0 d\tau$ decreases with increasing k . In addition, it seems likely that the upper limit is too high, for it is determined by adding the absolute values. On the whole, then, we may conclude that the correction term to ψ_0 changes the function by only a few per cent. The same conclusion can be drawn for the other functions, for they are practically equal to $\psi_0 e^{2\pi i \mathbf{k} \cdot \mathbf{r}}$. The total error made by neglecting the change in wave functions can be shown to be less than 1 kg cal/mol, which is less than the computational error of the present work.

By way of summary, it may be said that the solutions of (1) that are given by (2) and (11) in the sphere approximation also are solutions of Hartree's equations for the lattice.

d. The Influence of Exchange Terms.—It was demonstrated in Sec. 75, Chap. IX, that the exchange operator A is diagonal in the special case in which the eigenfunctions are free-electron waves $e^{2\pi i \mathbf{k} \cdot \mathbf{r}}$. In this case, the solutions of Hartree's and Fock's equations are identical. Now, we saw in part *c* that the solutions of Hartree's equations are very nearly free-electron waves for all three alkali metals under consideration. Hence, the exchange operator should be almost diagonal in these cases, and we may assume that the solutions of Fock's equations are identical with (10). This fact is a very fortunate one, for unless it were true the problem of treating the alkali metals would involve many more difficulties.

The influence on the exchange energy of the small p term in (18) may be estimated in the following way. For simplicity, we shall assume that $\psi_{\mathbf{k}}$ has the form

$$\psi_{\mathbf{k}} = e^{2\pi i \mathbf{k} \cdot \mathbf{r}} (\alpha + i\gamma \mathbf{k} \cdot \mathbf{r}) \quad (29)$$

where both α and γ are constants. This is equivalent to assuming that f_s and f_p are constant. In this case, the quantity $A \cdot \psi_{\mathbf{k}}$, where A is the exchange operator, is

$$-e^2 \sum_{\mathbf{k}'} \int \frac{(\alpha - i\gamma \mathbf{k}' \cdot \mathbf{r}_2)(\alpha + i\gamma \mathbf{k}' \cdot \mathbf{r}_1)(\alpha + i\gamma \mathbf{k} \cdot \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} e^{2\pi i[(\mathbf{k} - \mathbf{k}') \cdot \mathbf{r}_2 + \mathbf{k}' \cdot \mathbf{r}_1]} d\mathbf{r}_2.$$

If we make the transformation $\mathbf{r}_3 = \mathbf{r}_2 - \mathbf{r}_1$, we may,¹ to terms in γ^2 change this to

$$-e^2(\alpha + i\gamma \mathbf{k} \cdot \mathbf{r}_1) e^{2\pi i \mathbf{k} \cdot \mathbf{r}_1} \sum_{\mathbf{k}'} \int \frac{(\alpha + i\gamma \mathbf{k}' \cdot \mathbf{r}_3)(\alpha - i\gamma \mathbf{k}' \cdot \mathbf{r}_3)}{|\mathbf{r}_3|} e^{2\pi i(\mathbf{k} - \mathbf{k}') \cdot \mathbf{r}_3} d\mathbf{r}_3.$$

Hence, the functions (18) are eigenfunctions of A to terms in γ^2 . We shall find that the γ^2 terms are very small when we discuss the value of the exchange energy in part *e*.

e. The Energy in the One-electron Approximation.—We may now evaluate the energy of the crystal in the one-electron approximation. This energy is

$$E = \int \Psi^* H \Psi d\mathbf{r}$$

¹ This approximation is equivalent to assuming that the wave functions have the form

$$\psi_{\mathbf{k}} = \alpha e^{i\gamma \mathbf{k} \cdot \mathbf{r}}.$$

where Ψ is the determinantal eigenfunction, formed of the functions (2), and

$$H = \sum_i \left(-\frac{\hbar^2}{2m} \Delta_i + V_i \right) + \frac{1}{2} \sum_{i,j} \frac{e^2}{r_{ij}} + \frac{1}{2} \sum_{\alpha,\beta} \frac{e^2}{r_{\alpha\beta}} \quad (30)$$

is the total Hamiltonian of the lattice. In Eq. (30),

$$V_i = \sum_n v_c(|\mathbf{r}_i - \mathbf{r}(n)|) \quad (31)$$

where $v_c(|\mathbf{r}_i - \mathbf{r}(n)|)$ is the potential at \mathbf{r}_i arising from the ion at the position $\mathbf{r}(n)$, and the last term is the interaction potential of the ions. The latter has the same form as for point charges, for the ion cores are so far apart that they do not overlap appreciably.¹

The mean value of the operator (30) is

$$\begin{aligned} & 2 \sum_{\mathbf{k}} \int \psi_{\mathbf{k}}^* \left(-\frac{\hbar^2}{2m} \Delta + V \right) \psi_{\mathbf{k}} d\tau + 4 \sum_{\mathbf{k}, \mathbf{k}'} \frac{e^2}{2} \int \frac{|\psi_{\mathbf{k}}(\mathbf{r}_1)|^2 |\psi_{\mathbf{k}'}(\mathbf{r}_2)|^2}{r_{12}} d\tau - \\ & 2 \sum_{\mathbf{k}, \mathbf{k}'} \frac{e^2}{2} \int \frac{\psi_{\mathbf{k}}^*(\mathbf{r}_1) \psi_{\mathbf{k}'}^*(\mathbf{r}_2) \psi_{\mathbf{k}}(\mathbf{r}_2) \psi_{\mathbf{k}'}(\mathbf{r}_1)}{r_{12}} d\tau + \frac{1}{2} \sum_{\alpha,\beta} \frac{e^2}{r_{\alpha\beta}} \end{aligned} \quad (32)$$

where the sums over \mathbf{k} and \mathbf{k}' extend over the occupied values of \mathbf{k} , excluding spin, and the factors 2, 4, and 2 in the first three terms give the results of spin summation. We may split the first integral into integrals over each of the N cells of the lattice. Since the component integrals must be the same for every cell, they are equal to

$$N2 \sum_{\mathbf{k}} \int \psi_{\mathbf{k}}^* \left(-\frac{\hbar^2}{2m} \Delta + V \right) \psi_{\mathbf{k}} d\tau, \quad (33)$$

in which the integration extends over a single cell. From (31), we may derive the relation

$$\sum_{\mathbf{k}} 2 \int \psi_{\mathbf{k}}^* V \psi_{\mathbf{k}} d\tau = \sum_{\mathbf{k}} 2 \int |\psi_{\mathbf{k}}|^2 v_c(|\mathbf{r} - \mathbf{r}(n)|) d\tau + \sum_{n' \neq n} \sum_{\mathbf{k}} 2 \int |\psi_{\mathbf{k}}|^2 v_c(|\mathbf{r} - \mathbf{r}(n')|) d\tau \quad (34)$$

where the integration extends over the n th cell. In the second term, $v_c(|\mathbf{r} - \mathbf{r}(n')|)$ may be replaced by $-e^2/|\mathbf{r} - \mathbf{r}(n')|$ since the field is coulombic outside the n' th cell. Hence, if the fact that the distribution $\sum_{\mathbf{k}} 2|\psi_{\mathbf{k}}|^2$

¹ The contribution to the energy from the exchange and van der Waals terms is less than 0.2 kg cal/mol so that these terms are negligible for the cohesive energy. They are important, however, when the elastic constants are computed.

is spherically symmetric within a given cell is taken into account, (34) reduces to

$$\sum_{\mathbf{k}} 2 \int_{v_n} |\psi_{\mathbf{k}}(\mathbf{r})|^2 v_c(|\mathbf{r} - \mathbf{r}(n)|) d\tau = \sum_{n'}' \frac{e^2}{|\mathbf{r}(n) - \mathbf{r}(n')|}, \quad (35)$$

so that (33) may be replaced by

$$N \sum_{\mathbf{k}} 2 \int_{v_n} \psi_{\mathbf{k}}^* \left(-\frac{\hbar^2}{2m} \Delta + v_c(|\mathbf{r} - \mathbf{r}(n)|) \right) \psi_{\mathbf{k}} d\tau = \frac{1}{2} \sum_{\alpha, \beta}' \frac{e^2}{r_{\alpha\beta}}, \quad (36)$$

since

$$N \sum_{n'}' \frac{e^2}{|\mathbf{r}(n) - \mathbf{r}(n')|} = \frac{1}{2} \sum_{\alpha, \beta}' \frac{e^2}{r_{\alpha\beta}}. \quad (37)$$

The coulomb term in (32) may be split into two parts, namely, one for which \mathbf{r}_1 and \mathbf{r}_2 lie in the same cell, and another in which they lie in different cells. In the sphere approximation, the second term is equal to the mutual potential of a set of point charges, that is, to

$$\frac{1}{2} \sum_{\alpha, \beta}' \frac{e^2}{r_{\alpha\beta}}, \quad (38)$$

which cancels the last term in (36).

Upon combining these results, we find that (32) reduces to

$$2 \sum_{\mathbf{k}} \epsilon(\mathbf{k}) + 4N \sum_{\mathbf{k}, \mathbf{k}'} \frac{e^2}{2} \int_{\sigma} \frac{|\psi_{\mathbf{k}}(\mathbf{r}_1)|^2 |\psi_{\mathbf{k}'}(\mathbf{r}_2)|^2}{r_{12}} d\tau_{12} - 2 \sum_{\mathbf{k}, \mathbf{k}'} \frac{e^2}{2} \int \frac{\psi_{\mathbf{k}}^*(\mathbf{r}_1) \psi_{\mathbf{k}'}^*(\mathbf{r}_2) \psi_{\mathbf{k}}(\mathbf{r}_2) \psi_{\mathbf{k}'}(\mathbf{r}_1)}{r_{12}} d\tau_{12}. \quad (39)$$

Thus, the total energy differs from the sum of the energy parameters for all doubly occupied states by the self-potential of the charge distribution in each of the polyhedra and the exchange energy of all electrons. That the self-potential and exchange terms would enter in just this way might easily have been predicted on the basis of the discussions of parts *c* and *d*.

The self-energy of the charge distribution within a unit cell has been evaluated numerically for several values of r_s in each of the three metals under discussion; the results are listed in Table LV. It should be noted that the actual energy is very close to $0.6e^2/r_s$, which is the self-energy of a constant charge distribution, for the observed lattice constant in lithium and sodium. There is a considerable deviation, however, in the case of potassium, which probably is related to the errors in the potassium field that were discussed in part *b*.

TABLE LV.—COMPARISON OF COULOMB AND EXCHANGE ENERGIES FOR ACTUAL ELECTRONS AND FOR FREE ELECTRONS AT PARTICULAR VALUES OF r_s
(In kg cal/mol)

	Coulomb		Exchange	
	Actual	$0.6e^2/r_s$	Actual	$-0.458e^2/r_s$
Li($r_s = 3.21a_h$)	114.8	116.3	-90.2	-88.9
Na($r_s = 3.96a_h$)	93.8	94.4	-72.0	-72.0
K($r_s = 5.06a_h$)	82.2	74.3	-57.5	-56.6

The exchange terms are complicated by the fact that the integrals cannot be broken up into integrals over single polyhedra in any simple manner. They may be evaluated, however, by replacing the functions (18) with the simpler functions (29). This approximation actually is a good one in the alkali metals, for f_s is very nearly constant and f_p is small. The exchange energy has been evaluated by direct computation in this case and the result is

$$N\left(-0.458\frac{e^2}{r_s} - 1.05\eta\frac{e^2}{r_s} - 1.09\eta^2\frac{a_h^4e^2}{r_s^5}\right) \quad (40)$$

where η is the constant appearing in ψ_k when it is expressed in the form

$$\psi_k = \frac{e^{2\pi i \mathbf{k} \cdot \mathbf{r}}}{\sqrt{v}} \left(1 + i\frac{\sqrt{5}\eta}{r_s} \mathbf{k} \cdot \mathbf{r}\right). \quad (41)$$

When η is zero, (41) is a free-electron wave function and the expression (40) reduces to

$$-0.458N\frac{e^2}{r_s},$$

which, as we have seen in Sec. 75 of the last chapter, is the exchange energy for free electrons. Numerical values of the quantity (40) appear in Table LV. The terms in η^2 actually are very small, a fact showing that the exchange operator is almost constant (cf. part d).

f. Justification of the Sphere Approximation.—Before discussing the cohesive energy, we shall justify the approximation in which the polyhedron surrounding each atom is replaced by a sphere. It will be shown that the expression for the total energy does not differ appreciably from that derived in part e, if the wave functions (18) are used at points inside the polyhedron instead of at those in the sphere.

We may dispose of the exchange terms at once by observing that the principal term of (40), namely, $-0.458e^2/r_s$, was derived by using func-

tions which extend throughout the lattice. The terms in η may be influenced slightly by the sphere approximation, but they are so small that this variation cannot be important.

Next, let us discuss Eq. (32). V may be expressed rigorously in the form (31), so that we are interested in the sum

$$\sum_{\mathbf{k}} 2 \int_{\tau} \psi_{\mathbf{k}}^* \left(-\frac{\hbar^2}{2m} \Delta + v_c(|\mathbf{r} - \mathbf{r}(n)|) \right) \psi_{\mathbf{k}} d\tau \quad (42)$$

where the integration now extends over the polyhedron.

Although $\psi_{\mathbf{k}}$ satisfies Eq. (1), we cannot set the sum (42) equal to $2 \sum \epsilon_{\mathbf{k}}$, for the normal gradient of $\psi_{\mathbf{k}}$ is not continuous at the boundary of the polyhedron. In other words, since $\Delta \psi_{\mathbf{k}}$ is singular at the boundary, it is necessary to determine the contribution to the integral that is associated with the singularity. It may be shown (see the papers on sodium, footnote 1, page 348) that the necessary addition to the energy is

$$-\frac{\hbar^2}{2m} \sum_{\mathbf{k}} \sum_{\tau} \int_{\tau} \psi_{\mathbf{k}}^* \text{grad } \psi_{\mathbf{k}} \cdot d\sigma \quad (43)$$

where the integrals extend over the surface of a polyhedron and the second sum extends over all cells in the lattice. This integral is less than 0.005 ev for ψ_0 and may be neglected. Since $f_{\mathbf{p}}$ in (13) is small near the boundary, the integral is also negligible in other cases.

The remaining terms in the expression for the total energy, namely,

$$N \sum_{\mathbf{k}} \sum_{n' \neq n} 2 \int_{v_n} |\psi_{\mathbf{k}}|^2 v_c(|\mathbf{r} - \mathbf{r}(n')|) d\tau + \frac{e^2}{2} \int \frac{\left[2 \sum_{\mathbf{k}} |\psi_{\mathbf{k}}(\mathbf{r}_1)|^2 \right] \left[2 \sum_{\mathbf{k}'} |\psi_{\mathbf{k}'}(\mathbf{r}_2)|^2 \right]}{r_{12}} d\tau_1 d\tau_2 + \frac{1}{2} \sum_{\alpha, \beta} \frac{e^2}{r_{\alpha\beta}}, \quad (44)$$

may be broken into two expressions, namely, the for the self-energy of the charge in a polyhedron, and one for the mutual interaction energy of the set of polyhedra. In order to do so, all that is necessary is to write the second term of (44) in the form

$$N \sum_n \frac{e^2}{2} \int_{v_n} \frac{\left[2 \sum_{\mathbf{k}} |\psi_{\mathbf{k}}(\mathbf{r}_1)|^2 \right] \left[2 \sum_{\mathbf{k}'} |\psi_{\mathbf{k}'}(\mathbf{r}_2)|^2 \right]}{r_{12}} d\tau_1 d\tau_2 + \sum_{n \neq n'} \frac{e^2}{2} \int_{v_n} \int_{v_{n'}} \frac{\left[2 \sum_{\mathbf{k}} |\psi_{\mathbf{k}}(\mathbf{r}_1)|^2 \right] \left[2 \sum_{\mathbf{k}'} |\psi_{\mathbf{k}'}(\mathbf{r}_2)|^2 \right]}{r_{12}} d\tau_1 d\tau_2. \quad (45)$$

The first term is the sum of the self-energies of the electronic charge distribution in each cell, and the second is the interaction energy of the electronic charges in different cells. The sum of the second term in (45) and the first and last terms in (44) is the total electrostatic interaction energy of the polyhedra. In the sphere approximation, this sum was assumed to be zero, and (44) was evaluated by computing the self-energy of the charge in the sphere.

The electronic charge density in each of the three alkali metals is very nearly equal to e/v throughout all parts of the polyhedron, except in the region S near the nucleus where the eigenfunctions have nodes. Since the distribution is spherically symmetrical in this region, the potential outside S is the same as though the polyhedron contained a positive point charge at its center and a uniform negative distribution of density e/v . Hence, except for the value of the self-energy of the electronic charge in S , (44) has the same value as for a lattice of positive point charges that contains a uniform distribution of negative charge. The self-energy of this lattice may be computed by the methods discussed in Chap. II and is

$$-0.8958 \frac{e^2}{r_s}$$

per ion. The interaction energy of the positive and negative charge in a given polyhedron of this lattice, as determined by direct numerical calculation, is

$$-1.4939 \frac{e^2}{r_s}$$

If we designate by $\Delta\epsilon$ the difference between the self-energy of the actual electronic charge distribution in the region S and the self-energy of a constant electronic distribution in the same region, (44) is equal to

$$N \left(0.598 \frac{e^2}{r_s} + \Delta\epsilon \right) \quad (46)$$

where N is the number of atoms. The result corresponding to (46) is, in the sphere approximation,

$$N \left(0.6 \frac{e^2}{r_s} + \Delta\epsilon \right) \quad (47)$$

where $0.6e^2/r_s$ is the self-energy of a sphere of constant charge distribution. The difference, namely, $-N0.002e^2/r_s$, is never more than 0.2 kg cal/mol and may be neglected.

The close agreement between (46) and (47) is not a fortuitous coincidence but rests upon the fact that the field outside a polyhedron actually is practically zero, as assumed in the sphere approximation.

g. Energy in the Fock Approximation.—The complete expression for the cohesive energy of the alkali metals in the one-electron approximation is

$$N\left(\epsilon_0 + \frac{1.105}{r_s^2} \frac{e^2}{a_h} + \frac{0.598e^2}{r_s} + \Delta\epsilon - \frac{0.458e^2}{r_s} + \epsilon_f\right), \quad (48)$$

TABLE LVI.—CONTRIBUTIONS TO THE COHESIVE ENERGY OF THE ALKALI METALS
(The values for K contain the corrected values of $\epsilon_0 + \epsilon_f$ of Table LIV.)

r_s	$\epsilon_0 + \epsilon_I + \epsilon_F$	Cou- lomb term	Ex- change energy (40)	Cohesive energy in Fock approximation	Corre- lation energy	Final cohe- sive energy	Ob- served
Li							
3.00	-39.9 kg cal/mol	124.6	-95.1	10.4 kg cal/mol	-22.3	32.7	39.0
3.21	-39.0	114.8	-90.2	14.4	-21.7	36.1	
3.32	-38.8	110.9	-87.1	14.0	-21.5	35.5	
Na							
3.80	-23.9	98.3	-75.1	0.7	-20.3	21.0	26.0
3.96	-24.4	93.8	-72.0	2.6	-19.9	23.5	
4.12	-24.3	89.6	-69.2	3.9	-19.6	24.5	
K							
4.82	22.6
5.06	- 9.0	81.7	-57.3	-15.4	-17.8	2.4	
5.34	-13.6	69.7	-54.1	- 2.0	-17.3	15.3	
5.47	-14.3	67.6	-52.5	- 0.8	-17.1	16.3	

which includes only the principal term in the expression (4) for the exchange energy. This result is listed in Table LVI in the column headed "Cohesive energy in Fock approximation." The minimum values of the cohesive energies and the corresponding values of the lattice constant appear in Table LVII. A striking feature of these results is

TABLE LVII.—COMPARISON OF OBSERVED AND CALCULATED VALUES OF THE COHESIVE ENERGY AND LATTICE CONSTANT IN THE FOCK APPROXIMATION AND FINAL APPROXIMATION

	Cube-edge distance, Å			Cohesive energy, kg cal/mol		
	Observed	Fock	Final	Observed	Fock	Final
Li	3.46	3.50	3.50	39	14.6	36.2
Na	4.25	4.53	4.51	26	4.1	24.5
K	5.20	5.86	5.82	23	- 0.7	16.5

the fact that the cohesive energy determined by the one-electron approximation is smaller than the observed value by about the amount that one would expect from consideration of the atomic and molecular problems for which correspondingly accurate solutions have been found. It seems reasonable to ascribe most of the error, which is about 1 ev per electron, to the neglect of correlations between electrons having opposite spin.¹

Since the one-electron wave functions are nearly the same as those for free electrons, we may use Wigner's expression for the correlation energy, which was discussed in Sec. 76, namely,

$$-\frac{0.288}{r_s + 5.1a_h} e^2. \quad (49)$$

The final energies are given in the next to the last column of Table LVI and in Table LVII.

The calculations for lithium probably are the most significant since they are the most accurate. It is not easy to trace the source of the error of 3 kg cal/mol, but it probably arises from an error in the expression (49) for the correlation energy.² It is possible, however, that the effective ion-core field also contributes to this error, for it may not adequately represent the valence-electron and closed-shell interaction in the solid. The error in sodium probably has the same origin as that in lithium, whereas a large part of the error in potassium, which was discussed in part b, undoubtedly is connected with the inaccuracies in Hartree's closed-shell wave functions.

79. Metallic Hydrogen.—Although a metallic modification of hydrogen is unknown, Wigner and Huntington³ have made a computation of the properties of this hypothetical substance in order to estimate the conditions under which it should be stable. This computation was carried through on the assumption that the metallic lattice would be

¹ Tables LVI and LVII contain a summary of values that the writer regards as the "best" results of the papers listed in footnote 1, p. 348.

² C. Herring has pointed out to the writer that two types of correction to the equation $\epsilon = \hbar^2 k^2 / 2m^*$ may be expected in the case of lithium. In the first place, there may be a slight downward curvature in the (110) direction because of the proximity of the zone boundary. This effect presumably will also occur in sodium and probably is not large. In addition, since the $\epsilon(k)$ curve for lithium is well below the free-electron curve near $k = 0$ and we may expect it to rise in higher zones, there is probably a positive term of the order k^4 in a more accurate representation of $\epsilon(k)$. This term probably does not affect the Fermi energy appreciably but may be responsible for a decrease in level density near the edge of the filled region that is important for other properties, such as the conductivity and paramagnetic susceptibility (see Chaps. XV and XVI).

³ E. WIGNER and H. B. HUNTINGTON, *Jour. Chem. Phys.*, **3**, 764 (1935).

body-centered, so that the computational procedure closely resembles that used for the alkali metals. The principal difference is that the ion core is a proton which rigorously has a coulomb field. This simplification makes it possible to evaluate the wave functions and energies analytically when the sphere approximation is used. The coulomb, exchange, and correlation energies were estimated by the same methods that were used for the alkali metals. Figure 4 shows the electronic energy per electron as a function of r_s . The binding energy, which involves a small correction for the zero-point vibrational energy, is found to be 10.6 kg cal/mol for a density 0.59. The corresponding values for the molecular forms are 52.4 kg cal/gram-atom and 0.087. The difference in energy shows clearly why the ordinary form is not metallic.

If it is assumed that the observed compressibility of the molecular form, namely, $3 \cdot 10^{-9}$ cm²/dyne, is constant for a large change in volume, it is found that the energy of the molecular form would be increased by only 0.92 kg cal/mol when the density is changed from 0.087 to 0.59, so that there would be no tendency to change to a metallic form. Actually, the compressibility decreases with decreasing volume. Even if it became large enough to make the change from the molecular to the metallic phase possible, however, the pressure required would be at least 400,000 atmospheres which is not attainable at present.

80. Monovalent Noble Metals.—Fuchs¹ treated the cohesion of copper along the lines developed in the preceding sections and found that the interaction between closed shells and the exchange and correlation interaction between valence and closed-shell electrons play a much more important role in this metal than in the alkali metals. The reasons for this may be found from an investigation of Hartree's wave functions for atomic copper, for these show that about 0.4e of the 10e charges of the newly completed 3d shell lie outside the sphere whose volume is equal to the volume of the unit cell. This means that the effect of binding on the d-shell electrons is nearly as important as the effect on the valence electrons and that the discrete atomic d levels are split into bands in the solid. Since it would be difficult to treat all eleven electrons by the

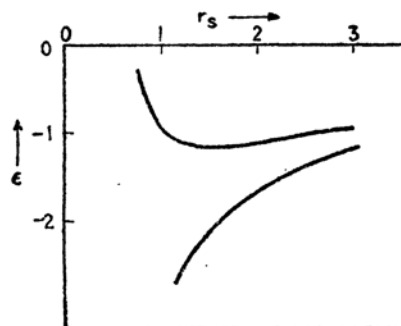


FIG. 4.—The lower curve is the $\epsilon_s(r_s)$ curve for metallic hydrogen. The upper curve is the total energy per electron. The origin of the energy scale is the energy of an ionized hydrogen atom, so that only the values of the upper curve relative to one Rydberg unit are of interest for cohesion. Abscissa is expressed in Rydberg units.

¹ K. FUCHS, *Proc. Roy. Soc.*, **151**, 585 (1935); **153**, 622 (1936); **157**, 444 (1936).

Bloch scheme in a computation of the cohesive energy, Fuchs assumed that the d shell is nearly rigid.

The closed-shell one-electron functions may be expressed in the form

$$\psi_{nl,m} = f_{nl}(r)O_l^m(\theta, \varphi) \quad (1)$$

where $f_{nl}(r)$ is a radial function associated with the radial and orbital

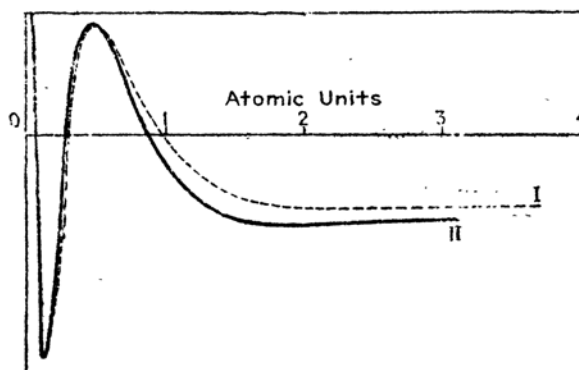


FIG. 5.—The lowest energy wave functions of metallic copper. The dotted curve corresponds to a case in which exchange is neglected and the full curve to one in which it is included. (After Fuchs.)

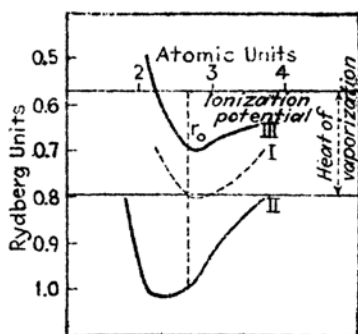


FIG. 6.—The $e_0(r_s)$ curves of metallic copper without exchange (curve I) and with exchange (curve II). Curve III represents the mean energy per electron after adding the Fermi energy. (After Fuchs.)

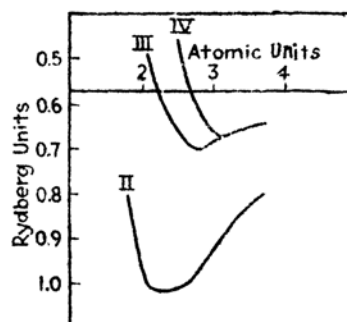


FIG. 7.—Curves II and III are the same as in Figure 6. Curve IV includes the ion-exchange repulsion. (After Fuchs.)

angular momentum quantum numbers n and l , respectively, and O_l^m is a surface harmonic. In a given closed shell, the z component of angular momentum quantum number m ranges over all integer values from $-l$ to l , and each ψ occurs once with each of the two possible spin orientations. If it is assumed that the valence electrons in the metal are so well correlated that only one appears in a given cell at a time, Fock's equation for the valence-electron wave function $\phi(\mathbf{r}_1)$ is

$$-\frac{\hbar^2}{2m}\Delta\varphi(\mathbf{r}_1) + v_c\varphi(\mathbf{r}_1) - r_s^2 \sum_{n,l} \sum_{m=-l}^l \int \frac{\psi_{nl,m}^*(\mathbf{r}_2)\psi_{nl,m}(\mathbf{r}_1)\varphi(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_2 = \epsilon\varphi(\mathbf{r}_1). \quad (2)$$

Here, v_c is the coulomb potential of the ion core, and the summation extends over those ψ in the closed shells which have the same spin as φ .

Fuchs solved Eq. (2) for the radial function of the 4s type that satisfies the boundary conditions $\varphi'(r_s) = 0$. In doing so, he evaluated the coulomb and exchange terms by the use of the Hartree one-electron wave functions for atomic copper. The solutions that were obtained with and without the exchange terms are shown in Fig. 5 along with Hartree's 3d function. In Fig. 6, $\epsilon_0(r_s)$ is represented for both approximations. The exchange energy is 2.5 ev for the value of r_s corresponding to the observed lattice constant. Since this is a large fraction of the cohesive energy of 3.1 ev, we may conclude that the correlation interaction between closed-shell electrons and valence electrons cannot be neglected in an accurate computation. This term could be included roughly in the manner developed by Gorin for potassium, but the labor is not justified in the present case because of other approximations that are made.

The Fermi energy was not determined by use of the perturbation method described in Sec. 78. Instead, it was evaluated more roughly by means of a perturbation scheme in which the entire periodic potential field of the ion cores $V(\mathbf{r})$ is treated as a perturbation. The starting wave functions in this case are free waves, and the energy of the perturbed functions $\chi_{\mathbf{k}}e^{2\pi i\mathbf{k}\cdot\mathbf{r}}$ is

$$\epsilon'(\mathbf{k}) = \frac{\hbar^2}{2m}\mathbf{k}^2 + V_{000} + \frac{1}{v^2} \sum_{\mathbf{k}'} \frac{|\int e^{-2\pi i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} V(\mathbf{r}) d\mathbf{r}|^2}{\epsilon(\mathbf{k}) - \epsilon(\mathbf{k}')} \quad (3)$$

(cf. Sec. 61). Here V_{000} is the integral of $V(\mathbf{r})$ over a unit cell,

$$\mathbf{k}' = \mathbf{k} + \mathbf{K},$$

where \mathbf{K} is any lattice vector in the reciprocal lattice, and v is the volume of the unit cell. Fuchs retained only that term of the second-order sum in (3) that belongs to the lowest zone and derived the expression

$$\frac{1}{r_s^2}(2.21 - 0.10075V_0^2r_s^2) \quad (4)$$

for the mean Fermi energy. This result is expressed in Rydberg units when r_s is given in Bohr units. V_0 is the integral:

$$V_0 = |\int e^{-4\pi i\mathbf{g}\cdot\mathbf{r}} V(\mathbf{r}) d\mathbf{r}|$$

where \mathbf{g} is the vector joining the origin of \mathbf{k} space to the nearest point of the first zone boundary for the face-centered lattice. Fuchs assumed that V_g , which is approximately equal to one-half the energy gap at the first zone boundary, is of the order of magnitude 2 ev. The second term in (4) is then practically negligible, and the Fermi energy is the same as for free electrons. This conclusion is not fully justified, for the same line of reasoning would lead to a Fermi energy less than the value for free electrons in the case of potassium, whereas Gorin's work shows that the opposite is true. We may conclude that the other terms in (3) usually are not negligible,¹ so that the approximation (4) is not reliable. The sum of $\epsilon_0(r_s)$ and the free-electron Fermi energy is shown in Fig. 6.

Fuchs assumed that the valence-electron self-energy is exactly balanced by the exchange and correlation energies, just as in the alkali metals. The error made in doing this undoubtedly is smaller than that introduced in estimating the Fermi energy from (4).

Finally, the very important ion-ion repulsion term was estimated² by means of a modified Fermi-Thomas method. When this result is added to the previous results, the energy curve shown in Fig. 7 is obtained. The closed-shell interaction correction is of the order of 0.5 ev when r_s is equal to r_0 and rises very rapidly as the lattice constant decreases. As we shall see in Sec. 82, the fact that the compressibility of the monovalent noble metals is less than that of the alkali metals can be associated with this interaction term.

The cohesive energy and the computed lattice constant of copper are listed in Table LVIII. This energy was obtained by subtracting the

TABLE LVIII.—THE OBSERVED AND CALCULATED COHESIVE ENERGY AND LATTICE CONSTANT OF COPPER

	Cube edge, Å	Cohesive energy, kg cal/mol
Calculated	4.2	33
Observed	3.6	81

observed ionization energy of atomic copper from the energy curve of Fig. 7. The large error in the calculated value cannot be associated with an error in a single term in this case, as it could in the case of lithium and sodium. In spite of this fact, the computations do give an indication

¹ Except in the case of lithium, the actual free-electron band does not correspond to the lowest zone, so that there are terms in (3) with positive as well as negative energy denominators.

² In subsequent work on the elastic constants, which is discussed in Sec. 82, this interaction was also used.

of the relative importance of the various terms that influence the cohesive energy in the monovalent noble metals.

81. Metallic Beryllium.—Herring and Hill¹ have given a detailed treatment of metallic beryllium along the lines of the preceding sections. When compared with the monovalent metals, this case presents two additional complicating features, namely, the fact that the lattice is close-packed hexagonal instead of cubic, so that there are two atoms per unit cell, and the fact that two zones are nearly completely occupied, so that the free-electron approximations cannot be applied without careful investigation. The atomic cell for beryllium is shown in Fig. 4 of Chap. IX, and the first two zones are shown in Fig. 18 of Chap. VIII. Since the computations contain more details than it seems advisable to present

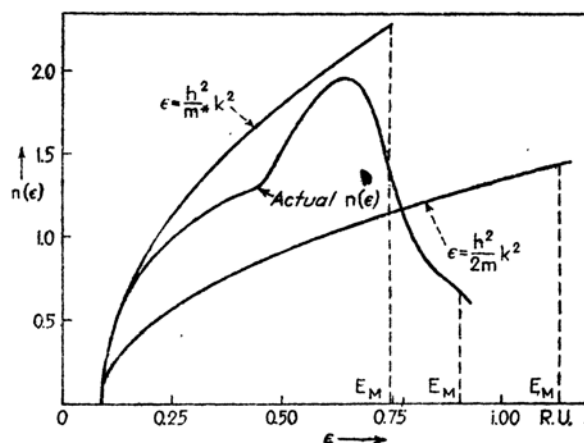


FIG. —The $n(\epsilon)$ curve for beryllium. This is compared with the free-electron $n(\epsilon)$ curves for $m/m^* = 1$ and 0.62 .

here, we shall survey only the general outline of their work and compare their results with experimental material.

To begin with, they obtained a self-consistent Hartree field for the valence electrons in the metal, using Hartree's atomic field for the $(1s)^2$ core. It should be mentioned at this point that all computations were carried out for values of r_s , the radius of the atomic sphere, both larger and smaller than the observed value, as well as for the observed value ($2.37a_h$). Wave functions for a number of points in k space were then computed, the functions for the center points of the zone being determined by the method used for ψ_0 for the alkali metals, and the functions at the zone boundaries being determined by the use of the free-electron perturbation scheme discussed in Sec. 73. From the energies of these functions, a level-density curve was obtained and the mean Fermi energy computed.

¹ C. C. HERRING and A. G. HILL, *Phys. Rev.* (to appear).

A complicating feature of this process is the fact that the exchange interaction between valence and core electrons had to be included, as in the case of potassium. A comparison of the actual distribution curve with that for perfectly free electrons and that for free electrons having the effective mass m^* determined from the curvature of the $\epsilon(k)$ curve near $k = 0$ is shown¹ in Fig. 8. The vertical lines represent the top of the occupied regions of levels in the three cases. It may be seen that the actual density function has a strong minimum near the top of the filled region—a fact that seems to occur generally among the alkaline earth metals. Values of m/m^* for several values of r_s are given in Table LIX. It may be noted that the ratio is less than unity as in lithium.

TABLE LIX.—VALUES OF m/m^* FOR THE VALENCE ELECTRONS OF BERYLLIUM DETERMINED FROM THE CURVATURE OF $\epsilon(k)$ NEAR THE BOTTOM OF THE FILLED BAND

r_s, a_0	m/m^*
2.07	0.422
2.37	0.616
2.67	0.697

Next, Herring and Hill attempted to make a more accurate estimate of exchange than would correspond to the use of the free-electron value. This proved to be very difficult, but they came to the conclusion that the exchange probably does not deviate by more than about 6 per cent from the free-electron value.

In lieu of a better alternative, they employed the free-electron correlation energy. The correlation energy is not greatly larger than the uncertainty in the exchange, so that this procedure probably does not introduce an important new error.

The results of the computation are listed in Table LX and are compared with computed quantities. The theoretical values are expressed

TABLE LX.—COMPARISON OF OBSERVED AND COMPUTED VALUES OF THE COHESIVE ENERGY, LATTICE PARAMETER, AND COMPRESSIBILITY OF BERYLLIUM

	Calculated	Observed
Cohesive energy (kg cal/mol).....	53 to 36	75
Equilibrium value of $r_s(a_0)$	2.23 to 2.57	2.37
$1/\beta \cdot 10^{-12}$ (cgs).....	0.87 to 1.32	1.25

in terms of the limiting values obtained as a result of several methods of approximating the various quantities.

To explain the discrepancy in cohesive energy, Herring and Hill suggest the possibility that exchange and correlation energies for the

¹ Figure 8 is not the final level-density curve obtained by Herring and Hill, but resembles it closely.

electrons near the top of the filled region, where the free-electron approximation undoubtedly is worst, may be considerably larger than their computations indicate. In connection with this, they note that the work function computed from their results by a method to be discussed in the next chapter is in very bad agreement with the best observed value. This discrepancy would also be decreased if the exchange and correlation energies of the uppermost electrons were increased.

82. The Elastic Constants of Metals.—It is pointed out in footnote 1, page 94, that the elastic constants c_{ij} that enter in the relation

$$X_i = \sum_j c_{ij} x_j \quad (i, j = 1, \dots, 6),$$

between the six independent components X_i of the stress tensor and the six independent components x_j of the symmetric strain tensor are given by the equation

$$c_{ij} = \frac{\partial^2 E(x_1, \dots, x_6)}{\partial x_i \partial x_j} \quad (1)$$

where E is the energy per unit volume of the crystal as a function of the strains. Thus, the elastic constants may be computed if the energy of the crystal as a function of homogeneous atomic displacements is known.

a. Compressibility.—The simplest energy change to compute is that accompanying a uniform compression in all directions. In this case,

$$\begin{aligned} x_1 = x_2 = x_3 &\equiv \xi, \\ x_4 = x_5 = x_6 &= 0, \end{aligned} \quad (2)$$

so that the change in energy per unit volume is

$$\delta E = \frac{1}{3}(c_{11} + c_{22} + c_{33})\xi^2 + (c_{12} + c_{23} + c_{31})\xi^2. \quad (3)$$

Since the relative change in volume $\delta V/V$ is 3ξ for small displacements, Eq. (3) may be placed in the form

$$\delta E = \frac{1}{2\beta} \left(\frac{\delta V}{V} \right)^2, \quad (4)$$

in which the compressibility β is related to the c by the equation

$$\frac{1}{\beta} = \frac{1}{9}[(c_{11} + c_{22} + c_{33}) + 2(c_{12} + c_{23} + c_{31})]. \quad (5)$$

In cubic crystals,

$$\begin{aligned} c_{11} &= c_{22} = c_{33}, \\ c_{12} &= c_{23} = c_{31}, \end{aligned}$$

so that

$$\frac{1}{\beta} = \frac{1}{3}(c_{11} + 2c_{12}), \quad (6)$$

whereas, in hexagonal crystals,

$$c_{11} = c_{22},$$

$$c_{23} = c_{31},$$

whence

$$\frac{1}{\beta} = \frac{1}{9}(2c_{11} + c_{33} + 2c_{12} + 4c_{31}). \quad (7)$$

Now, $\delta V/V$ is simply related to the relative change $\delta r_s/r_s$ in the radius of the atomic sphere by the equation

$$\frac{\delta V}{V} = 3 \frac{\delta r_s}{r_s}.$$

Hence, if $E(r_s)$ is the energy per unit volume of the crystal as a function of r_s ,

$$\frac{1}{\beta} = \frac{r_s^2}{9} \frac{\partial^2 E}{\partial r_s^2},$$

or if $\epsilon(r_s)$ is the energy per atom,

$$\frac{1}{\beta} = \frac{1}{12\pi r_s} \frac{\partial^2 \epsilon(r_s)}{\partial r_s^2}. \quad (8)$$

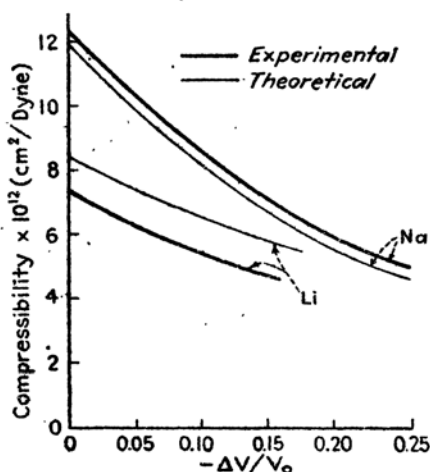


FIG. 9.—Comparison of the observed and calculated compressibilities of lithium and sodium as functions of the relative change in volume $-\Delta V/V_0$. (After Bardeen.)

Bardeen¹ has used the computed $\epsilon(r_s)$ curves for lithium and sodium to compute the compressibilities for a range of values of r_s and has compared these with values obtained from Bridgman's room-temperature results² by extrapolation to absolute zero of temperature. The computed and observed curves are given in Fig. 9. The pressure required to produce the maximum change of volume in these cases is of the order of magnitude 40,000 kg/cm². Bardeen suggests that the disagreement in the case of lithium arises from neglect of the effect of the discontinuity of the $\epsilon(k)$ curve at the first zone boundary (cf. footnote 2, page 366).

b. Other Relations.—In order to compute all the elastic constants, it is necessary to determine the energy change of the crystal for deforma-

¹ BARDEEN, *op. cit.*, 372.

² P. W. BRIDGMAN, *Proc. Am. Acad. Sci.*, **72**, 207 (1938).

tions other than a homogeneous compression. The most extensive work of this kind has been carried out by Fuchs¹ for the alkali metals. He determined the energy change for the following two additional types of dilation:

1. *Equal contraction and expansion, respectively, along two cube edges, which leave the volume unchanged.*—This deformation may be expressed in terms of the fractional displacements along the three axes which will be designated by α_x , α_y , and α_z . If the z direction is that along which the distances are unchanged, we have

$$\begin{aligned}\alpha_z &= x_3 = 0, \\ \alpha_x &= x_1 = \xi, \\ \alpha_y &= x_2 = -\xi, \\ x_4 &= x_5 = x_6 = 0.\end{aligned}$$

Thus, in this case,

$$\delta E = (c_{11} - c_{12})\xi^2,$$

and

$$c_{11} - c_{12} = \frac{1}{2} \frac{\partial^2 E}{\partial \xi^2}. \quad (9)$$

2. *Shearing strain in a plane parallel to two cube edges.*—In this case, the compressional strains are zero, and only the shear strains are finite. Thus the dilatation is described by the relations

$$x_1 = x_2 = x_3 = 0; \quad x_4 = \xi; \quad x_5 = x_6 = 0.$$

Hence,

$$\delta E = \frac{1}{2} c_{44} \xi^2.$$

It is evident that these deformations, unlike the deformation that determines the compressibility, distort the cells into noncubic forms, thus deforming the spheres of the sphere approximation into ellipsoids. Fuchs treated the changes in each of the contributions to the total energy that were discussed in the preceding sections in the following way:

i. Since the lowest wave function ψ_0 is practically constant near the boundary of the polyhedron, Fuchs assumed that it is not appreciably changed by a dilatation that does not alter the volume. Thus, ϵ_0 was regarded as a function of r_s alone.

ii. Fermi energy. Since the Fermi energy of a free-electron gas depends only upon the volume, this assumption was retained in treating the alkali metals.

¹ FUCHS, *op. cit.*

iii. The coulomb energy. We saw in part *f*, Sec. 78, that the coulomb interaction of the electrons and ions is the same as the self-energy of a set of positive point charges in a uniform cloud of negative charge, except for regions very near to the nuclei. Since the form of the wave functions in these regions is not affected by distortions that do not change the volume, the change in coulomb energy for the distortions 1 and 2 is the same as the change in the electrostatic self-energy of the simple lattice. The methods for computing this energy were discussed in Chap. II for the case in which the atoms are at lattice positions. The changes for the distortions 1 and 2 are simply related to second derivatives of these expressions for the appropriate lattice.

iv. Exchange energy and correlation energy. It was assumed that these depend only on the volume, as for perfectly free electrons.

v. Noncoulomb ion-ion interaction. The interaction energy of the closed shells is very sensitive to the interatomic distance and, hence, affects the elastic constants appreciably even in cases, such as the alkali metals, in which its contribution to the cohesive energy is small. Fuchs determined the effect of this interaction in the alkali metals by use of a repulsive term of the Born-Mayer type

$$C_{12}be^{\frac{r_1+r_2-r}{\rho}}$$

which was discussed in Chap. II, and a van der Waals term of the type

$$\frac{A}{r^6},$$

which was discussed in Chap. VII. The methods used for determining C_{12} and A need not be discussed again here. In the case of copper, the closed-shell interaction was taken from the work discussed in Sec. 80. This interaction leads to a slightly expanded lattice.

To summarize, the elastic constants are determined by terms *iii* and *v*.

In the alkali metals, the coulomb contribution is two to three times larger than the contribution from ion-ion interaction, whereas the situation is reversed in the noble metals, as may be seen in Table LXI. A comparison of observed and computed constants is given in Table LXVII. It may be seen from Table LXI that the comparatively high rigidity of the noble metals arises from the closed-shell interactions.

It is interesting to examine the extent to which these constants satisfy the conditions¹ for isotropy and the Cauchy-Poisson relations, which are, respectively,

$$2c_{44} = c_{11} - c_{12}$$

¹ See footnote 1, p. 94 and footnote 1, p. 106.

TABLE LXI.—COMPOSITION OF THE ELASTIC CONSTANTS OF MONOVALENT METALS
(AFTER FUCHS)
(In units of 10^{11} dynes/cm²)

	Coulomb contribution	Ion-ion interaction
Li		
$c_{11} - c_{12}$	0.339	-0.02 _{1/2}
c_{44}	1.263	0.086
Na		
$c_{11} - c_{12}$	0.143	-0.02
c_{44}	0.532	0.048
K		
$c_{11} - c_{12}$	0.0644	0.02
c_{44}	0.240	0.020
Cu		
$c_{11} - c_{12}$	0.573	4.53
c_{44}	2.57	6.4

TABLE LXII.—COMPARISON OF OBSERVED AND CALCULATED VALUES OF THE ELASTIC
CONSTANTS OF MONOVALENT METALS
(In units of 10^{11} dynes/cm²)

	$1/\beta$	$c_{11} - c_{12}$	c_{44}	c_{11}	c_{12}
Li					
Calculated.....	1.30	0.341	1.349	1.53	1.19
Na					
Calculated.....	0.88	0.141	0.580	0.97	0.83
Observed.....	~0.85	0.145	0.59	~0.95	~0.80
K					
Calculated.....	0.41	0.062	0.260	0.45	0.39
Cu					
Calculated.....	14.1	5.1	8.9	17.5	12.4
Observed.....	13.9	5.1	8.2	18.6	13.5

and

$$c_{12} = c_{44}.$$

It is readily seen from the results of Table LXII that neither condition is closely satisfied and that the alkali metals¹ are very far from isotropic. As was pointed out in Sec. 19, this fact accounts for a large part of the anomalies in the specific-heat curves of these metals.

83. Cohesion of Alloys.—It was seen in Sec. 3 that alloys usually have a small heat of formation. There have been no extensive computations of these heats. Mott,² however, has attempted to estimate the difference in energy between completely ordered and completely disordered β brass, which is a body-centered metal containing equal numbers of copper and zinc atoms (*cf.* Sec. 3). He assumed that the additional valence electrons of zinc cluster mainly about the zinc ions and that the potential near a zinc ion is greater than that near a copper ion by an amount

$$\varphi(r) = \frac{e}{r} e^{-qr} \quad (1)$$

where q is a constant. Since this potential vanishes as e^{-qr} at large distances, it follows that its use is equivalent to the assumption that the zinc atoms are neutral. The screening constant q was evaluated by comparing the observed resistivity of the alloy with that computed on the assumption that the difference between the potential of the two ions is given by (1) (*cf.* Sec. 130) and was found to be $2.7 \cdot 10^8 \text{ cm}^{-1}$, or

$$\frac{1}{q} = 0.37 \text{ \AA},$$

which is only about one-quarter of the radius of the zinc atom. When this value of q is used, the charges in the zinc and copper polyhedra of the body-centered lattice are $0.075e$ and $-0.075e$, respectively, which corresponds to an electrostatic interaction energy per atom of

$$-\frac{1.017}{a}(0.075)^2 e^2 \quad (2)$$

where a is the cube-edge distance and 1.017 is the appropriate Madelung constant. This energy is 0.027 ev for β brass. Since the mean potential at an ion arising from neighbors would be zero in a perfectly disordered lattice, it follows that (2) is the electrostatic ordering energy. In addition, Mott estimated the decrease in exchange repulsive energy in going

¹ The experimental values for sodium were obtained by Quimby and Siegel (*see* Sec. 19).

² N. F. Mott, *Proc. Phys. Soc.*, **49**, 258 (1937).

from the disordered to the ordered state and obtained a value of 0.013 ev per atom. Thus the total ordering energy is 0.04 ev per atom. The total change in energy in the transition from order to disorder as obtained by integrating the specific-heat curve of Fig. 43, Chap. I, is about 0.043 ev per atom, which is the same order of magnitude as the computed value. This computation suggests that the largest source of ordering energy is the Madelung term, as in ionic crystals.

84. Simplified Treatments of Cohesion.¹—In addition to the preceding work, in which the computation of cohesive properties is based entirely on the Schrödinger equation, there have been several treatments of cohesion that start from other points. Among these treatments, the two most important are those which start from semiempirical equations of state and interrelate measured quantities and those which use the Fermi-Thomas statistical equation. We shall discuss these briefly.

a. The Semiempirical Method.—Perhaps the most extensive work of this kind is that of Grüneisen,² who assumed, following a suggestion of Mie, that the atoms in monatomic substances interact in pairs with a potential energy relation of the type

$$\epsilon(r) = -\frac{a}{r^m} + \frac{b}{r^n} \quad (1)$$

in which r is the interatomic distance and a , b , m , and n are positive constants, n being larger than m . This assumption is analogous to that of the Born theory of ionic crystals in which $m = 1$ and $a = e_1 e_2$. According to (1), the total energy of the crystal at absolute zero of temperature is

$$E = -\frac{a}{2} N \sum_i \frac{1}{r_i^m} + \frac{b}{2} N \sum_i \frac{1}{r_i^n} \quad (2)$$

in which the sum extends over all values of the distance r_i between a given atom and the others.

Three relations among the four parameters in (2) were determined by the condition that this expression give the observed values of the atomic volume, cohesive energy, and compressibility of the solid at absolute zero.

The temperature-dependent free energy of the lattice was obtained by adding to (2) the free-energy function corresponding to Debye's specific-heat law. When this was done, it was found that

¹ This type of work is extensively discussed in the book by J. C. Slater, *Introduction to Chemical Physics* (McGraw-Hill Book Company, Inc., New York, 1939).

² G. GRÜNEISEN, see *Handbuch der Physik*, vol. X.

$$\frac{n+2}{6} = -\frac{d \log \Theta_D}{dV} \quad (3)$$

where Θ_D is the characteristic temperature and V is the molar volume. It may be shown that the right-hand side of (3) is equal to the quantity

$$\frac{-V(\partial V/\partial T)_p}{C_p(\partial V/\partial p)_T} \quad (4)$$

Thus a fourth relation among the parameters was determined by the condition that $(n+2)/6$ be equal to the measured values of (4).

Using the resulting total free-energy function, Grüneisen was able to correlate a number of properties of metals and of monatomic insulators such as diamond. For example, Table LXIII gives a comparison of

TABLE LXIII.—COMPARISON OF OBSERVED VALUES OF EXPANSION COEFFICIENTS WITH THOSE COMPUTED BY THE USE OF GRÜNEISEN'S THEORY
(The values of α_l are given in cgs units.)

Temperature interval, °K	$\alpha_l \cdot 10^6$	
	Calculated	Observed
Diamond		
84.8–194.1	0.16	0.18
194.1–273.2	0.61	0.58
273.2–296.2	0.97	0.97
296.2–328	1.17	1.17
328–351	1.37	1.45
Copper		
20.4–80.5	4	3.8
82–289	14.0	14.2
289–523	17.4	17.2
523–648	18.7	18.6
648–773	19.5	19.6

observed mean values of the expansion coefficients of diamond and copper, taken for a range of temperature ΔT , and the mean values computed from Grüneisen's equation of state.

It is clear that the function (2) cannot be expected to give the proper elastic constants since it would lead to the Cauchy-Poisson relations, which are not usually satisfied in metals (*cf.* Sec. 82).

Modifications of Grüneisen's plan that are based on more accurate information of the cohesive forces in metals have been developed by a

number of people, principally Rice¹ and Bardeen.² These methods have in common the property that they express the absolute zero energy of the entire metal as a sequence of terms which vary inversely as different powers of the volume, semitheoretical and empirical data being used to evaluate the constants. This type of modification has been pushed furthest by Bardeen, who used it to discuss the behavior of alkali metals at very high pressures and to correlate a number of Bridgman's measurements.

It was seen in Sec. 78 that the cohesive properties of the alkali metals are given closely by the quantity $-(\epsilon_0 + \epsilon_f + \epsilon_p)$, in which $\epsilon_0 + \epsilon_f$ is the energy of the electron of zero wave number relative to the energy of the free atom, in the sphere approximation, and ϵ_f is the Fermi energy. This result depends upon the fact that the exchange and correlation effects combine in such a way as to allow on the average only one electron in a given polyhedron at a given time. Now, by integrating Eq. (1), Sec. 78, with appropriate simplifying assumptions, it is possible to show that $\epsilon_0(r_s)$ has the approximate form

$$\epsilon_0(r_s) = \frac{a}{r_s^3} - \frac{3}{r_s} \quad (5)$$

where a is a constant that varies from solid to solid. In order to obtain a slightly more general result, Bardeen assumed that ϵ_0 actually can be expressed in the form

$$\epsilon_0(r_s) = \frac{a}{r_s^3} - \frac{c}{r_s} \quad (6)$$

to a higher degree of accuracy. We shall not consider a proof of (5) necessary, for (6) is a valid assumption if taken in the same spirit as Grüneisen's relation (1). Now, we saw in part f, Sec. 78, that the Fermi energy varies as

$$2.21\alpha \frac{e^2}{r_s^2} \quad (7)$$

in which $\alpha = m/m^*$. Since r_s is proportional to $v^{1/3}$, where v is the atomic volume, the total energy of the crystal in the sphere approximation may be expressed in the form

$$E(v) = A\left(\frac{v_0}{v}\right) + B\left(\frac{v_0}{v}\right)^{2/3} - C\left(\frac{v_0}{v}\right)^{1/3} \quad (8)$$

where A , B , and C are constants for a given metal and v_0 is the observed atomic volume.

¹ O. K. RICE, *Jour. Chem. Phys.*, 1, 649 (1933).

² BARDEEN *p. cit.*, 372.

Bardeen determined the parameters A , B , and C by adjusting them so that (8) would give the empirical values of $E(v_0)$, v_0 , and the compressibility. Thus, since E must have a minimum when $v = v_0$, we obtain

$$\frac{C}{3} = A + \frac{2}{3}B. \quad (9)$$

Moreover, if the value of E at absolute zero is E_0 and if the value of the compressibility is β_0 , we find in addition

$$\begin{aligned} -E_0 &= 2A + B, \\ \frac{3}{\beta_0} &= 2A + \frac{2B}{3}. \end{aligned} \quad (10)$$

Values of A , B , and C determined from these equations by the use of empirical data are given in Table LXIV for all the alkali metals. The

TABLE LXIV.—COMPARISON OF EMPIRICAL AND THEORETICAL VALUES OF THE CONSTANTS IN BARDEEN'S EMPIRICAL EQUATION OF STATE
(The constants are expressed in units of 10^{-12} erg per atom.)

	Li	Na	K	Rb	Cs
A (empirical).....	1.4	4.3	5.1	4.2	3.7
B (empirical).....	8.4	1.2	- 2.1	- 0.6	0.0
B (theoretical).....	4.5	3.0	2.3	1.7	1.5
C (empirical).....	20.8	15.4	11.2	11.4	11.2
C (theoretical).....	20.9	16.2	13.2	12.3	11.4

values of B and C are compared with the values obtained using the value $c = 3$, corresponding to Eq. (5), and the free-electron value of α in (7). It may be seen that (5) is a good approximation, whereas (7) is not good if α is assumed to be a constant. We have seen that α actually varies appreciably with v in lithium and potassium [see Eq. (17), Sec. 78; and Table LII]; hence, this discrepancy in the value of B is not surprising. Bardeen suggests that the true Fermi energy probably varies in the manner

$$B\left(\frac{v_0}{v}\right)^{\frac{1}{2}} + B'\left(\frac{v_0}{v}\right) \quad (11)$$

in which the second term takes into account the variation of α with v .

As a test of the validity of the relation (8), Bardeen compared the volume-pressure curves obtained from this equation with those obtained by extrapolating Bridgman's measurements to absolute zero of temperature. The theoretical volume at pressure p is determined by the equation

$$\left(\frac{\partial E}{\partial V}\right)_p = -p, \quad (12)$$

which leads to the relation

$$pv_0 = y^4(y-1)[2A + \frac{2}{3}B + A(y-1)] \quad (13)$$

in which

$$y = \left(\frac{v_0}{v}\right)^{\frac{1}{3}}.$$

The comparison is shown in Fig. 10. It may be observed that the agreement is fairly good in all the alkalis, the largest deviation occurring for rubidium. In addition, the experimental curve for cesium shows a break indicative of a phase change.

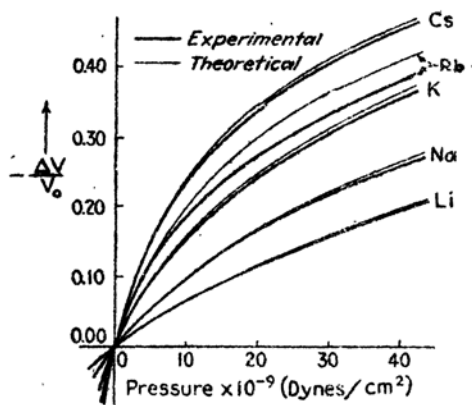


FIG. 10.—Comparison of the observed and calculated relative changes in the volume of the alkali metals as functions of pressure. The break in the experimental curve for cesium is discussed in the text. (After Bardeen.)

The effect of ion-ion exchange interaction is neglected in Eq. (8). This varies as

$$Ae^{-\frac{r}{\rho}} \quad (14)$$

and should contribute higher power terms to (8). One might expect these higher power terms to enter first for cesium, since it has the highest compressibility. Bardeen assumed that the expression (8) is valid for both a face-centered and a body-centered lattice with given values of the constants and that additional different terms should be added in the two cases in order to include the ion-ion interaction. He determined ρ in Eq. (14) from the results of Mayer and Bleick's computation of the exchange interaction between neon atoms, and he determined A by the methods used in the Born-Mayer theory (*cf.* Sec. 11). He found that a polymorphic change from a body-centered to a face-centered lattice should occur at about the same pressure as the observed change indicated in Fig. 10, which suggests that the change actually is of this type.

b. *The Fermi-Thomas Method.*—In the Fermi-Thomas¹ statistical treatment of the many-electron problem, it is assumed that the electrons are effectively free at each point, so that the mean kinetic energy of the electrons at a point r is related to the density $n(r)$ by the equation for perfectly free electrons, namely,

$$\epsilon_k(r) = \frac{3}{10} \frac{h^2}{m} \left[\frac{3n(r)}{8\pi} \right]^{\frac{2}{3}} \quad (15)$$

[cf. Eq. (20), Sec. 26.]. This assumption clearly is rigorous only if the change in potential is very small over the distance of the electronic wave length, a condition that is not satisfied near the nuclei of atoms. In their original treatments of neutral atoms, Fermi and Thomas assumed that the kinetic energy of the fastest electrons, namely,

$$\epsilon_m(r) = \frac{5}{2} \epsilon_k(r), \quad (16)$$

is equal to the negative of the potential energy $-e\varphi(r)$ at r . Thus, $n(r)$ and $\varphi(r)$ are related by the equation

$$e\varphi(r) = \frac{h^2}{2m} \left[\frac{3n(r)}{8\pi} \right]^{\frac{2}{3}}$$

or

$$n(r) = \frac{8\pi}{3} \left[\frac{2me}{h^2} \varphi(r) \right]^{\frac{3}{2}} \quad (17)$$

Now, φ satisfies Poisson's equation

$$\Delta\varphi = 4\pi en(r); \quad (18)$$

and if $n(r)$ is eliminated from (18) by means of Eq. (17),

$$\Delta\varphi = \frac{32\pi^2}{3} e \left(\frac{2me}{h^2} \right)^{\frac{3}{2}} \varphi^{\frac{3}{2}}. \quad (19)$$

This equation is solved for neutral atoms with the boundary conditions that φ be zero at infinity and vary as Ze/r near the origin. It yields reasonably good qualitative distribution functions for the electrons in heavy atoms.

Modifications of Eq. (19) that are valid for systems more general than neutral atoms have been developed by Dirac² and by Lenz and Jensen.³ The Lenz-Jensen scheme, which is a variational one, is formally equivalent to the original Fermi-Thomas scheme, inasmuch as the Eulerian

¹ See, for example, L. BRILLOUIN, *Die Quantenstatistik* (Julius Springer, Berlin, 1930), for a survey of early work on the Fermi-Thomas theory.

² P. A. M. DIRAC, *Proc. Cambridge Phil. Soc.*, **123**, 714 (1929).

³ H. JENSEN and W. LENZ, *Z. Physik*, **77**, 713, 722 (1932).

equation of their variational principle reduces to Eq. (19) when applied to a neutral atom. Dirac's scheme, however, is more general, for it contains an additional term that decreases the energy at a point r by an amount equal to the exchange energy of one of a system of electrons in a region where the density is n , namely,

$$-0.916e^2 \left[\frac{4\pi n(r)}{3} \right]^{\frac{1}{2}}. \quad (20)$$

Thus, Lenz and Jensen's method is equivalent to Hartree's when applied to perfectly free electrons, whereas Dirac's scheme is equivalent to Fock's.

Slater and Krutter¹ applied Dirac's method to a system of electrons that are in a lattice of point positive charges but did not obtain a minimum in the energy versus interatomic distance curve. This result is not surprising; for the correlation energy, which is neglected in this method, is a large fraction of the cohesive energy in the alkali metals, which correspond most closely to Slater and Krutter's model.

Gombas² has applied Lenz and Jensen's method to lattices that correspond to the alkali and alkaline earth metals. The lattices are even less stable in this approximation than in Dirac's; however, Gombas added a number of correction terms in order to compensate for the errors of the method. Thus, he added the free-electron exchange and correlation energies, which are sufficient to make the lattices stable. In addition, he added ion-ion interaction correction terms and valence-electron ion-core exchange terms. In this way, he has obtained energies that approximate the observed ones closely. The success of this procedure in the case of the alkali metals undoubtedly lies in the fact that the valence electrons are very nearly free so that the results obtained from the Fermi-Thomas method are nearly the same as those obtained from Hartree's method.

B. IONIC CRYSTALS

85. Sodium Chloride.—The most significant, purely quantum mechanical computations of the cohesive energies of ionic crystals are those which have been carried out on sodium chloride and on lithium hydride by Landshoff³ and Hylleraas,⁴ respectively. The first of these will be discussed in the present section. Landshoff based his work on a

¹ J. C. SLATER and H. KRUTTER, *Phys. Rev.*, **47**, 559 (1935).

² P. GOMBAS, *Z. Physik*, **95**, 687 (1936); **99**, 729 (1936); **100**, 599 (1936); **104**, 592 (1937); **108**, 509 (1938).

³ R. LANDSHOFF, *Z. Physik*, **102**, 201 (1936); *Phys. Rev.*, **52**, 246 (1937). The writer is indebted to Landshoff for the values given in Table LVIII.

⁴ E. A. HYLLEAAS, *Z. Physik*, **63**, 771 (1930).

Heitler-London approximation in which he used the solutions of Fock's equations for Na^+ and Cl^- that were determined by Fock and Petrashen and by Hartree and Hartree, respectively. For simplicity, he did not attempt to compute the absolute energy of the lattice. Instead, he determined the energy of the crystal relative to the theoretical energy of the free ions as obtained from the one-electron functions. The advantage of this procedure lies in the fact that the internal energies of the ions do not appear in the final expression for the cohesive energy and do not need to be evaluated. Since the absolute accuracy of the solutions of Fock's equations for the ions has not been determined, Landshoff's results cannot furnish us with an estimate of the absolute accuracy of the Heitler-London approximation when applied to the solid. Cohesive energies computed by Landshoff's method might turn out to be larger than the observed value, as we have seen in Sec. 77.

It should be recalled that the Heitler-London and Bloch schemes are identical in cases such as the present one in which the Heitler-London scheme contains only closed shells. Thus, the Bloch scheme should lead to results of comparable accuracy.

The one-electron functions associated with neighboring ions are not orthogonal for the observed lattice spacing, for they overlap appreciably. Since it is convenient to use an orthogonal set of one-electron functions, Landshoff orthogonalized the free-ion functions in the following approximate manner: Let ψ_μ designate the free-ion wave functions that are centered about different nuclei. Landshoff showed that the following linear combinations of the ψ ,

$$\chi_\mu = \psi_\mu(1 + \frac{3}{2}S_{\mu\mu}) - \frac{1}{2}\sum_{\eta}' S_{\mu\eta}\psi_\eta \quad (1)$$

where

$$\begin{aligned} S_{\mu\nu} &= \int \psi_\nu^* \psi_\mu d\tau, \\ S_{\mu\mu} &= 0, \\ S_{\mu\mu}^2 &= \sum_{\eta} S_{\eta\mu}^* S_{\mu\eta}, \end{aligned} \quad (2)$$

satisfy the conditions

$$\begin{aligned} \int \chi_\mu^* \chi_\nu d\tau &= -\frac{3}{4} \sum_{\eta} S_{\mu\eta} S_{\eta\nu} + O(S^3), \\ \int |\chi_\mu|^2 d\tau &= 1 + O(S^2), \end{aligned} \quad (3)$$

in which $O(S^3)$ designates terms that contain the "overlap" integrals $S_{\mu\nu}$ to the third and higher powers. Thus, the χ are orthogonal to the

approximation in which $-\frac{1}{2}\sum_{\lambda} S_{\mu\lambda} S_{\lambda\mu}$, and higher order terms are negligible.

Landshoff proceeded on the assumption that these terms are negligible and treated the φ as though they were orthogonal. It is not possible to tell from his results to what extent this assumption actually is justifiable.

Under these conditions, the mean total energy of the Hamiltonian operator,

$$H = -\frac{\hbar^2}{2m}\sum_i \Delta_i - \sum_{i,\alpha} \frac{e^2 Z_\alpha}{r_{i\alpha}} + \frac{1}{2}\sum_{i,j} \frac{e^2}{r_{ij}} + \frac{1}{2}\sum_{\alpha,\beta} \frac{Z_\alpha Z_\beta}{r_{\alpha\beta}}, \quad (4)$$

is

$$\begin{aligned} E = \sum_{\mu} 2 \int \chi_{\mu}^*(\mathbf{r}_1) \left(-\frac{\hbar^2}{2m} \Delta - \sum_{\alpha} \frac{e^2 Z_{\alpha}}{r_{1\alpha}} \right) \chi_{\mu}(\mathbf{r}_1) d\tau + \\ \sum_{\mu,\nu} 2e^2 \int \frac{|\chi_{\mu}(\mathbf{r}_1)|^2 |\chi_{\nu}(\mathbf{r}_2)|^2}{r_{12}} d\tau_{12} - \\ \sum_{\mu,\nu} e^2 \int \frac{\chi_{\mu}^*(\mathbf{r}_1) \chi_{\nu}^*(\mathbf{r}_2) \chi_{\mu}(\mathbf{r}_2) \chi_{\nu}(\mathbf{r}_1)}{r_{12}} d\tau_{12} + \frac{1}{2} \sum_{\alpha,\beta} \frac{e^2 Z_{\alpha} Z_{\beta}}{r_{\alpha\beta}}. \end{aligned} \quad (5)$$

The factor 2 enters in the first two terms because of spin.

If the χ are expressed in terms of the ψ by means of Eq. (1), Eq. (5) becomes

$$\begin{aligned} E = \sum_{\mu} 2(\mu|h|\mu) + \sum_{\mu,\nu} [2(\mu\nu|g|\mu\nu) - (\mu\nu|g|\nu\mu)] + \sum_{\mu} S_{\mu\mu}^2 \{ 2(\mu|h|\mu) + \\ 2 \sum_{\nu} [2(\mu\nu|g|\mu\nu) - (\mu\nu|g|\nu\mu)] \} - \sum_{\mu,\eta} S_{\mu\eta} \{ 2(\eta|h|\mu) + \sum_{\nu} [2(\eta\nu|g|\mu\nu) - \\ (\eta\nu|g|\nu\mu)] \} + \sum_{\alpha,\beta} \frac{Z_{\alpha} Z_{\beta} e^2}{r_{\alpha\beta}} \end{aligned} \quad (6)$$

where

$$\begin{aligned} (\eta|h|\mu) &= \int \psi_{\eta}^* \left(-\frac{\hbar^2}{2m} \Delta - \sum_{\alpha} \frac{e^2 Z_{\alpha}}{r_{1\alpha}} \right) \psi_{\mu} d\tau, \\ (\eta\nu|g|\mu\lambda) &= e^2 \int \frac{\psi_{\eta}^*(\mathbf{r}_1) \psi_{\nu}^*(\mathbf{r}_2) \psi_{\mu}(\mathbf{r}_1) \psi_{\lambda}(\mathbf{r}_2)}{r_{12}} d\tau_{12}. \end{aligned} \quad (7)$$

All second-order terms of the type

$$\sum_{\eta,\lambda} S_{\mu\eta} S_{\nu\lambda} (\eta\lambda|g|\mu\nu)$$

have been discarded in deriving Eq. (6).

Expression (6) may be simplified considerably if the fact that the ψ are solutions of Fock's equations is used, for then the integrals involving Δ may be expressed in terms of the energy parameters of Fock's equations and coulomb and exchange integrals. Many of these terms cancel when the resulting equation is subtracted from the expression for the energy of the system of free ions. The final equation for the cohesive energy is

$$E_c = \sum_{\mu} \left[\int |\psi_{\mu}(r_1)|^2 \left(- \sum_{\alpha \neq \mu} \frac{Z_{\alpha} e^2}{r_{\alpha 1}} \right) d\tau_1 + \sum_{\nu}' 2(\mu\nu|q|\mu\nu) + \frac{1}{2} \sum_{\alpha, \beta}' \frac{Z_{\alpha} Z_{\beta} e^2}{r_{\alpha\beta}} \right] + \sum_{\mu} \left\{ \sum_{\nu}' (\mu\nu|g|\nu\mu) + \sum_{\eta} S_{\mu\eta} \left[\int \psi_{\mu}^* \psi_{\eta} \left(\sum_{\beta \neq \mu} - \frac{Z_{\beta} e^2}{r_{\beta 1}} \right) d\tau_1 + 4 \sum_{\nu \neq \eta}' (\mu\nu|g|\mu\nu) \right] \right\}. \quad (8)$$

The first, or coulomb, term is the electrostatic energy of a lattice of ions of which the electronic charge distribution is given by the functions $|\psi_{\mu}|^2$. This differs from the value $-1.748Ne^2/r_0$, corresponding to the Madelung energy (cf. Chap. II), because neighboring ions overlap. Landshoff found that only the overlapping of neighboring Na^+ and Cl^- ions is important. There are two terms in the correction to the Madelung value, namely, a positive term I_1 , which arises from the repulsion between the electrons, and a negative one $-I_2$, which arises from the attraction between the electrons and the nuclei. Although I_1 and I_2 turn out to be of the order of magnitude of several electron volts per ion pair at the observed value of r_0 , the anion-cation distance, Landshoff found that they nearly cancel one another so that the total coulomb correction is small.

The exchange terms were evaluated in a straightforward manner. Only the exchange terms between nearest neighbors are important, but for these ions both the conventional exchange terms

$$C = \sum_{\mu} \sum_{\nu}' (\mu\nu|g|\nu\mu)$$

and the term, which we shall call B , that includes the factors $S_{\mu\nu}$, are very large. For example, C is -0.348 ev per ion pair, and B is 0.645 ev for $r_0 = 5a_k$. The two terms compensate for one another, however, and their sum is much smaller.

The coulomb and exchange terms appear in Table LXV along with the Madelung energy. The cohesive energy has a maximum of 179.3 kg cal/mol at $r = 5.34a_k$ which should be compared with the observed value of 183 ± 10 kg cal/mol at $r_0 = 5.4a_k$.

TABLE LXV.—CONTRIBUTIONS TO THE COHESIVE ENERGY OF SODIUM CHLORIDE (r_0 is the nearest-neighbor distance. In kg cal/mol)

r_0/a_h	Madelung energy	Coulomb correction	Exchange interaction	Cohesive energy	$I_{v.d.w.}$	Final cohesive energy
5.0	-217.7	0.3	41.1	176.3	-4.2	180.5
5.1	-213.5	0.1	35.5	177.9	-3.7	181.6
5.2	-209.4	-0.1	30.6	178.9	-3.3	182.2
5.3	-205.4	-0.2	26.4	179.2	-3.0	182.2
5.4	-201.6	-0.3	22.8	179.1	-2.7	181.8
6.0	-181.5	-0.3	9.5	172.3	-1.5	173.8

r_0/a_h	Final values	
	Calculated	Observed
	5.25	5.4
E_c	183	183

It is difficult to estimate the absolute accuracy of this one-electron approximation; for on the one hand the absolute accuracy of Fock's approximation for the free ions is not known, and on the other Landshoff does not give a numerical estimate of the magnitude of the neglected terms in E_c . If Landshoff's approximations are valid, the agreement between his results and experiment indicates that the Heitler-London method that is based on solutions of Fock's equations for the free ions is fairly good. As the one-electron functions become more accurate, the center of gravity of the electronic charge on an ion approaches the nucleus, and the wave functions of different ions overlap less. This in turn decreases the repulsive terms and increases the computed value of E_c . Since Landshoff's result leaves little room for improvement, the one-electron functions are probably very good.

In higher approximations, both van der Waals and polarization terms should be added to the preceding results. Landshoff has evaluated the van der Waals term by means of the expressions developed in Chap. II and has obtained the following equation:

$$E_{v.d.w.} = \frac{17.6}{(r_0/a_h)^6} \frac{e^2}{a_h}$$

This adds about 3 kg cal to the result of Table LXV and decreases r_0 to $5.25a_h$.

86. Lithium Hydride.—Hylleraas¹ has treated lithium hydride in essentially the same way that Landshoff has treated sodium chloride. He employed one-electron wave functions of hydrogenic type with nuclear screening in order that the entire computation might be performed analytically. The principal objection to this procedure is that his approximate hydrogenic wave functions do not lead to very good binding energies for the free H^- and Li^+ states. The one-electron wave functions are

$$\psi = e^{-\left(Z - \frac{5}{16}\right)\frac{r}{a_0}} \quad (1)$$

where Z is 1 for hydrogen and 3 for lithium. The ionic energies derived from these functions are

	Observed, ev	Calculated, ev
H^-	- 0.718	0.745
Li^+	75.28	73.72

In other words, H^- is not stable in the approximation in which the functions (1) are employed. This fault is reflected in the fact that the electronic distribution of H^- which is obtained from the hydrogenic functions is not very accurate. This error would not be very serious if it were not for the fact that Hylleraas found the H^-H^- interaction to be

TABLE LXVI.—CONTRIBUTIONS TO THE COHESIVE ENERGY OF LITHIUM HYDRIDE (In kg cal/mol)

r_0/a_0	Madelung energy	H-Li interaction		H-H, coulomb plus exchange	Total cohesive energy
		Coulomb	Exchange		
3.84	-283.1	33.8	67.4	-34.0	215.9
4.16	-261.3	21.9	46.9	-26.0	218.5
4.56	-238.5	15.0	26.0	-14.0	211.5

	Final values	
	Calculated	Observed
r_0/a_0	4.08	3.84
E_c	218	218.5

¹ E. A. HYLLERAAS, *Z. Physik*, **63**, 771 (1930).

large and to favor binding. The contribution from this interaction term probably would be much smaller if more accurate wave functions were used.

We shall not dwell on the details of Hylleraas' computations since they involve exactly the same approximations as those of the preceding case. Although his results were given in analytical form we shall list numerical values. Table LXVI contains the values of the quantities which were discussed in the last section. The H-Li and H-H interaction terms are listed separately.

The principal contribution to the H-H interaction is the coulomb correction, for the exchange term is practically negligible. As we mentioned above, this correction undoubtedly would be less if more accurate wave functions had been used.

87. The Elastic Constants of Ionic Crystals.—The methods discussed in Sec. 82 evidently could be applied to compute the elastic constants of sodium chloride and lithium hydride. Actually, only the compressibility of sodium chloride has been evaluated. It is given in Table LXVII.

TABLE LXVII.—THE RECIPROCAL OF THE COMPRESSIBILITY OF SODIUM CHLORIDE
(AFTER LANDSHOFF)
(In units of 10^{12} dynes/cm²)

	Calculated	Observed
$1/\beta$	4.35	4.16

C. MOLECULAR CRYSTALS

88. Computations of Cohesive Energy.—The principal source of intermolecular cohesion in nonpolar molecular crystals and in many polar molecular crystals is the van der Waals force. The quantum mechanical methods of computing this force were described in Chap. VII. We saw there that the first approximation term in the expression for the van der Waals energy of two molecules varies as A/r^6 , where r is the interatomic distance, and that the next term varies as B/r^8 . This interaction term is not the only one, however, for just as in ionic crystals (*cf.* Chap. II) there are other sources of interaction energy. Those which are most important in the simple cases to be considered here are the following: (a) The electrostatic interaction term which arises from the "static" charge distributions on the molecules. Although the molecules considered here are neutral, they are not spherically symmetrical and, for this reason, have an electrostatic interaction. (b) The repulsive term, which in the Born-Mayer theory varies as $be^{-r/\rho}$ where b and ρ are constants. The force arising from this term varies more rapidly with dis-

tance than the van der Waals force, so that the van der Waals energy is larger than the repulsive energy at the observed intermolecular distance.

There is a continuous gradation between those molecules which rigidly retain the electronic structure of the free molecule in entering the solid and those which become as highly deformed as the constituents of valence crystals. For this reason, there is no sharp dividing line between valence and molecular types, as we have seen in Sec. 8. Since the cohesion of valence compounds is characterized by exchange energies that favor cohesion in the Heitler-London approximation, it follows that the repulsive term changes its sign to favor binding in the course of the transition from ideal molecular crystals to valence types. Only the more ideal molecular types, which are characterized by very low heats of sublimation, have been considered in any detail up to the present time.

a. London's Calculations.—London¹ first suggested that the electrostatic and repulsive energy terms are very small in comparison with van der Waals terms, and he computed the cohesive energies simply by evaluating the van der Waals term for the observed interatomic distance. As we shall see from the more accurate work described in parts *b* and *c*, this hypothesis sometimes is valid within the comparatively large error of computations of the van der Waals energy but is often very inaccurate. In addition, London treated diatomic and triatomic molecules as though they were spherically symmetrical, in order that he might use the equations that were derived in Sec. 58. We shall see in part *c* that this probability is only a fair approximation.

According to Eq. (19), Sec. 58, the van der Waals interaction energy ϵ_v of two molecules is, in first approximation,

$$\epsilon_v = -\frac{A}{r^6} \quad (1)$$

where

$$A = \frac{3}{4}h\nu_0\alpha^2, \quad (2)$$

in which α is the polarizability and ν_0 is a mean excitation frequency. The sum of terms of type (1) for a face-centered cubic lattice containing N molecules is

$$E = -N\frac{59A}{d^6} \quad (3)$$

where d is the cube-edge distance. In evaluating A , London assumed that ν_0 should be closely equal to the principal oscillator frequencies that appear in empirical equations for the refractive index of the gas of each kind of molecule (*cf.* Sec. 148, Chap. XVII). He used these frequencies

¹ F. LONDON, *Z. physik. Chem.*, **11B**, 222 (1930).

in the cases in which they have been measured and spectroscopically determined series-limit frequencies in the other cases. In addition, he employed measured polarizabilities.

Computed values of E are given in Table LXVIII for a number of molecular crystals. Not all these have face-centered cubic crystals, but London assumed that the error made in assuming they have is small.

TABLE LXVIII.—THE VAN DER WAALS ENERGIES OF A NUMBER OF MOLECULAR CRYSTALS AS DETERMINED BY LONDON'S APPROXIMATE EQUATION

Molecule	Lattice	Calculated E , kg cal/mol		Measured E
		ν_e from refraction	ν_e from spect.	
Ne	f.c.c.	0.47	0.40	0.52
A	f.c.c.	2.08	1.83	1.77
N ₂	h.c.p.	1.64	1.61	1.50
O ₂	h.c.p.	1.69	1.48	1.74
CO			1.86	
CH ₄	f.c.c.	2.42	2.47	2.40
NO		2.89	2.04	
HCl	f.c.c.		4.04	4.34
HBr	f.c.c.		4.53	4.79
HI			6.50	
Cl ₂	Fig. 78, Chap. I		7.18	6.0

b. *More Accurate Calculations for the Rare Gases.*—It was mentioned in Sec. 58 that the repulsive-energy terms have been computed in the case of helium and neon. Using Mayer and Bleick's equation for the neon repulsive energy and assuming that this term is valid only for nearest neighbors, Deitz¹ found that the total repulsive energy is 0.35 kg cal/mol which, in absolute magnitude, is almost as large as London's value of the van der Waals energy, namely, 0.47 kg cal. This suggests that London's value of the van der Waals term is about half as large as the true value and that most of the agreement in Table LXVIII is fortuitous. The discrepancy presumably arises from the fact that Margenau's higher order term is neglected, for this may be as much as half as large as London's term.

c. *Carbon Dioxide.*—Sponer and Bruch-Willstätter² have computed the cohesive energy of solid carbon dioxide, taking into account all three of the terms that were mentioned in the introductory paragraphs. It may be recalled from the discussion of Sec. 7 that this solid, which is a typical molecular crystal, consists of a face-centered cubic arrangement

¹ V. DEITZ, *Jour. Franklin Inst.*, **219**, 459 (1935).

² H. SPONER and M. BRUCH-WILLSTÄTER, *Jour. Chem. Phys.*, **5**, 745 (1937).

of CO_2 molecules that is similar to the lattice of FeS_2 (Fig. 64, Chap. I). The cohesive energy is 8.24 kg cal/mol.

In first approximation, the molecule was treated as a spherically symmetric unit. The van der Waals energy that was obtained for the observed lattice distance by use of the London-Margenau equations of Sec. 58 and the observed polarizability and ionization energies is 6.0 kg cal/mol of which 3.6 arises from London's term and 2.4 from Margenau's. In the next approximation, the molecule was treated as though it were composed of two centers, which in practice are regarded as the effective centers of charge of the two oxygen ions. Since the electronic distributions on these ions are presumably distorted from spherical form, it was assumed that the positions of the centers lie between the carbon and oxygen ions. Using several reasonable values of these positions, Sponer and Bruch-Willstätter obtained values of the van der Waals energy ranging between 9.6 and 7.6 kg cal/mol.

The electrostatic energy was computed on the assumption that two excess electrons of the oxygen ions are localized at the centers mentioned in the preceding paragraph and that the carbon ion, which is midway between the centers, has a positive charge of $4e$. In this way, it was found that the electrostatic energy varies between -0.1 and -0.5 kg cal/mol.

The constants in the repulsive term were determined from measured values of the compressibility and expansion coefficient by the method described in Sec. 9, Chap. II. The resulting value of the repulsive energy is about -1.1 kg cal. A summary of the computed quantities is given in Table LXIX.

TABLE LXIX

	Contribution to Cohesive Energy, kg cal/mol
Van der Waals (one center)	6.0
Van der Waals (two centers)	Between 9.6 and 7.6
Electrostatic energy	Between -0.1 and -0.5
Repulsive term	-1.1
Total result (two center)	~ 7
Observed	8.24

CHAPTER XI

THE WORK FUNCTION AND THE SURFACE BARRIER

89. The Principles Involved in Computing the Work Function.—

There is a close correlation between the work function of a clean metal surface and the volume properties of the metal. In general, the work function is high¹ if the cohesive energy is high, and vice versa. On the other hand, the work function may be appreciably affected if the surface is altered by oxidation or by the deposition of a fraction of an atomic layer of another metal. These facts indicate that the work function is determined both by the binding properties and by the surface structure. The correct relationship between these factors was first pointed out by Wigner and Bardeen.² We shall begin by discussing the principles involved in their work and then shall present detailed computations.

Let us consider a semiinfinite crystal that is bounded by a plane, as shown in Fig. 1. The distribution of ions in regions of the main body of the crystal far from the surface is not influenced by the presence of the surface and may be computed under the assumption that the lattice is perfectly periodic. On the other hand, the distribution in the cells near the surface should be different from the distribution in internal cells because the potential changes rapidly near the boundary. The type of difference depends upon the kind of crystal, the orientation of the surface, and the kind of adsorbed atoms or ions; for this reason, it must be investigated separately in each case. Fortunately, this surface-sensitive region does not affect the bulk properties of the solid in any

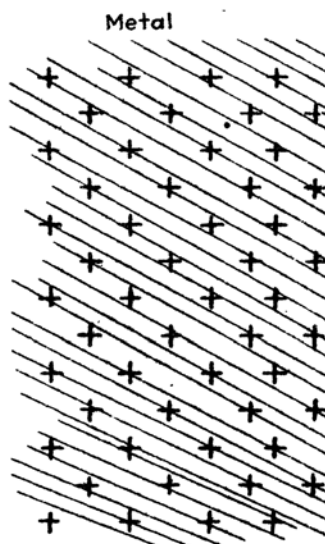


FIG. 1.—A semiinfinite crystal. The distribution of ions in internal regions is the same as that determined by X rays. The surface distribution, however, may be different.

¹ This correlation was first pointed out by A. Sommerfeld, *Naturwissenschaften*, **15**, 825 (1927); **16**, 374 (1928). See also J. Frenkel, *Z. Physik*, **49**, 31 (1928).

² E. WIGNER and J. BARDEEN, *Phys. Rev.*, **48**, 84 (1935); J. BARDEEN, *Phys. Rev.*, **49**, 653 (1936).

practically, for its volume is of the order of $1/N$ times the volume of the crystal if there are N^3 cells in the lattice, and N is of the order of 10^7 for a single crystal of ordinary size.

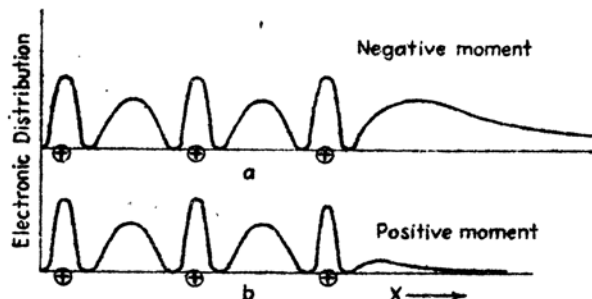


FIG. 2.—Schematic diagram showing cases in which the electronic distribution may lead to negative and positive dipole layers. In case *a* the electronic distribution extends beyond the surface a good deal; in case *b* the extension is small.

The only effect of the surface cells in which we need be interested at present is the way in which they influence the dipole moment of the

(100) surface. If the crystal is a monatomic cubic metal, such as sodium, the dipole moment of interior cells is zero. The dipole moment of cells near the surface usually is not zero, however, for the potential field in this region is not cubically symmetrical. Hence, these cells effectively give the surface a dipole moment (*cf.* Fig. 2). In polar crystals, such as sodium chloride, in which the unit cell may be chosen in such a way that it has a dipole moment, the surface moment depends both upon the way in which the surface cuts through the lattice and upon the distortion of the surface cells (*cf.* Fig. 3). In any case, we shall designate the component of the dipole moment per unit area in the direction normal to the surface by P_n . The important property of this moment for our purposes is the fact that it raises the coulomb potential inside the lattice by a term $-4\pi P_n$.

FIG. 3.—The intersection of different surface planes with a (010) plane. If the ions are not displaced relative to the positions in an ideal lattice, the (100) surface has zero dipole moment because there are alternate positive and negative charges. In the (110) surface the surface charges shown are positive, however, those surface charges in (010) planes above and below the one illustrated are negative so that there is no surface dipole in this case. On the other hand, the charges in a (111) surface are either all positive or all negative; hence, there is a dipole layer in this case.

Let us consider the influence of the surface on the work function from the standpoint of Koopmans' theorem (*cf.* Sec. 67, Chap. VIII).

This theorem states that the energy required to remove an electron in the state ψ_k from the crystal is equal to the negative of the parameter $\epsilon(k)$ in Fock's equation.

$$-\frac{\hbar^2}{2m}\Delta\psi_{\mathbf{k}}(\mathbf{r}_1) + \left[V(\mathbf{r}_1) + e^2 \int \frac{\rho(\mathbf{r}_2)}{r_{12}} d\tau_2 + A(\mathbf{r}_1) \right] \psi_{\mathbf{k}}(\mathbf{r}_1) = \epsilon(\mathbf{k})\psi_{\mathbf{k}}(\mathbf{r}_1). \quad (1)$$

Here, $V(\mathbf{r}_1)$ is the total ion-core potential, $\rho(\mathbf{r}_2)$ is the valence-electron distribution, and A is the exchange operator. The quantity

$$V(\mathbf{r}_1) + e^2 \int \frac{\rho(\mathbf{r}_2)}{r_{12}} d\tau_2$$

is the Hartree potential of the crystal, which, at points inside the lattice, differs from the potential for a crystal in which the surface dipole is zero by the term $4\pi P_n$. $A(\mathbf{r}_1)$ arises from the exchange correlation hole, which is confined to the vicinity of the electron, and is not affected by the surface as long as the electron is inside the crystal. $\psi_{\mathbf{k}}$ has the form $\chi_{\mathbf{k}}e^{2\pi i\mathbf{k}\cdot\mathbf{r}}$ inside the lattice if we employ the Bloch scheme; and although it is different in the cells near the surface, the total volume of these cells is so small that they may be neglected in computing integrals involving ψ that extend over the entire lattice.

Multiplying (1) by $\psi_{\mathbf{k}}^*$ and integrating, we find

$$\epsilon(\mathbf{k}) = -\frac{\hbar^2}{2m} \int \psi_{\mathbf{k}}^* \Delta\psi_{\mathbf{k}} d\tau_1 + \int |\psi_{\mathbf{k}}|^2 \left[V'(\mathbf{r}_1) + e^2 \int \frac{\rho(\mathbf{r}_2)}{r_{12}} d\tau_2 \right] d\tau_1 + \int \psi_{\mathbf{k}}^* A\psi_{\mathbf{k}} d\tau_1 + 4\pi e P_n. \quad (2)$$

Here,

$$V'(\mathbf{r}_1) + \int \frac{e^2 \rho(\mathbf{r}_2)}{r_{12}} d\tau_2$$

is the Hartree potential for a lattice having no surface dipole. Thus, except for the term $4\pi e P_n$, the work function $-\epsilon(\mathbf{k})$ is determined by volume integrals. The introduction of correlation effects does not alter this conclusion, for correlation terms, like exchange terms, arise from a hole in the vicinity of the electron.

When an electron is removed from a metal, the remaining electrons concentrate in the interior of the solid in order to keep this region electrostatically neutral. At first sight, it appears as though this effect might invalidate the previous conclusions. This is not so, however, for the energy change accompanying the concentration process is equal to the difference between the electrostatic energy of a volume and a surface distribution of 1 electronic unit. This difference, which is of the order e^2/L , where L is the diameter of the crystal, is about 10^{-6} ev for ordinary specimens.

90. The Internal Contribution to the Work Function.—Wigner and Bardeen have evaluated the volume part of the expression (2) for the

uppermost electrons in the filled levels of the metals discussed in Chap. X. It may be assumed that the metals are uncharged without introducing an error; for, in the first place, additional charge would accumulate at the surface, leaving the interior of the metal neutral, and, in the second place, this surface charge is never high enough in actual cases to alter the surface dipole layer appreciably. An ordinary specimen of metal has about 10^{16} surface cells which contain about 10^6 esu of electronic charge. If one per cent of this charge were removed, the field near the metal would be raised to about one million volts, which is as high as practical values ordinarily go; however, the dipole layer would not be altered by more than a few per cent.

According to the sphere approximation, which is reliable in the alkali metals (*cf.* Chap. X), the Hartree potential within any cell may be determined by the charge inside that cell, for the electronic and ionic charges in other cells cancel one another. Thus, the first two terms in Eq. (2) of the preceding section, namely,

$$-\frac{\hbar^2}{2m} \int \psi_k^* \Delta \psi_k d\tau_1 + \int |\psi_k|^2 \left[V'(r_1) + e^2 \int \frac{\rho(r_2)}{r_{12}} d\tau_2 \right] d\tau_1,$$

are equal to

$$\frac{V}{v} \left\{ \int \psi_k^* \left(-\frac{\hbar^2}{2m} \Delta + v_c \right) \psi_k d\tau_1 + \frac{V}{v} e^2 \int |\psi_k|^2 \left(\int \frac{\rho(r_2)}{r_{12}} d\tau_2 \right) d\tau_1 \right\} \quad (1)$$

where the integrals extend over a single cell, V/v is the ratio of the volume of the crystal to that of a cell, v_c is the ion-core field inside the cell, and ρ_e is the electronic distribution in the cell. The first term in this equation is equal to

$$\epsilon_0 + \frac{\hbar^2}{2m^*} k^2 \quad (2)$$

where ϵ_0 is the energy parameter in the equation

$$-\frac{\hbar^2}{2m} \Delta \psi_0(r) + v_c(r) \psi_0(r) = \epsilon_0 \psi(r)$$

(*cf.* Sec. 78, Chap. X) and $(\hbar^2/2m^*)k^2$ is the additional energy of ψ_k . If we replace the polyhedral cell by a sphere and assume that both $e\rho_e$ and $eV|\psi_k|^2/v$ are constant and equal to e/v , the second integral is simply twice the self-energy of a spherical charge distribution, namely, $1.2e^2/r_s$, where r_s is the radius of the sphere.

The exchange integral in Eq. (2) of the preceding section was evaluated in Sec. 75, Chap. IX, for plane waves and is equal to

$$-0.306 \frac{e^2}{r_s} f\left(\frac{k}{k_0}\right)$$

where

$$f\left(\frac{k}{k_0}\right) = 2 + \frac{k_0}{k} \left[1 - \left(\frac{k}{k_0}\right)^2 \right] \log \left| \frac{k_0 + k}{k_0 - k} \right|.$$

An explicit equation for the correlation energy of any given electron has not been derived; however, the correlation energy of the electrons at the top of the band is (cf. Sec. 76, Chap. IX)

$$g(r_s) = \frac{1}{3}g'(r_s)r_s$$

where

$$g(r_s) = -e^2 \frac{0.288}{r_s + 5.1a_h}.$$

If we add these results, we find that the energy ϵ_w required to remove an uppermost electron is given by the equation

$$-\epsilon_w = \epsilon_0 + \frac{\hbar^2}{2m^*}k_0^2 + 1.2\frac{e^2}{r_s} - 0.612\frac{e^2}{r_s} + g(r_s) - r_s\frac{g'(r_s)}{3} + 4\pi eP_n \quad (3)$$

where $4\pi eP_n$ is the surface dipole term.

Now, $(\hbar^2/2m^*)k_0^2$ is equal to five-thirds of the mean Fermi energy ϵ_F of the crystal and may be replaced by this quantity. If we then replace ϵ_0 by the value obtained from the expression for the cohesive energy per atom ϵ_c , namely,

$$-\epsilon_c = Z \left[\epsilon_0 + \epsilon_F + \frac{0.6e^2}{r_s} - \frac{0.458e^2}{r_s} + g(r_s) \right] + I(Z), \quad (4)$$

where Z is the number of valence electrons per atom and $I(Z)$ is the ionization potential of the free atom, we find

$$\epsilon_w = \frac{1}{Z} \left[\epsilon_c + I(Z) \right] + \left[-\frac{2}{3}\epsilon_F - 0.6\frac{e^2}{r_s} + \frac{0.458e^2}{3r_s} + \frac{r_sg'(r_s)}{3} - 4\pi eP_n \right]. \quad (5)$$

Wigner and Bardeen derived this equation by another method, namely, by computing the energy of the crystal as a function of the number of electrons and ions, N_e and N_i , respectively. The work function ϵ_w is then the derivative of this energy with respect to N_e . The advantage of this procedure is that the work required to remove an ion or a neutral atom may be computed from the same expression.¹

All the quantities in (5) except P_n have been computed for lithium, sodium, and potassium in Chap. X, so that it is possible to determine $\epsilon_w + 4\pi eP_n$ for these metals. The values are given in Table LXX.

¹The expression for the energy required to remove an ion contains the surface dipole with opposite sign. Hence, the work required to remove a neutral atom does not depend on the dipole moment.

TABLE LXX

	r_s , CA	$\epsilon_w + 4\pi e P_n$, ev	Experimental value
Li	3.28	2.19	2.28
Na	4.00	2.15	2.25
K	4.97	1.87	2.24

The close agreement between the observed values of ϵ_w and the computed values of $\epsilon_w + 4\pi e P_n$ for lithium and sodium suggests that the surface dipole moment is very small for clean metal surfaces. This conclusion is borne out by explicit computations of P_n that Bardeen has made for sodium.¹

On the basis of the work discussed in Sec. 81 and Eq. (5), Herring and Hill have found the work function of beryllium to be negative by about

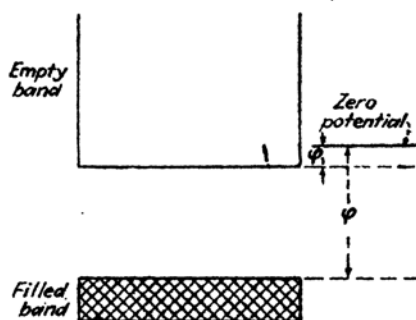


FIG. 4.—Schematic energy-level diagram of an insulator. The zero of potential is assumed to be slightly above the bottom of the empty band. The work function ϕ_1 for inserting an electron is smaller than that for removing an electron from the filled band ϕ by the energy difference of the filled and empty band.

1.7 ev, if it is assumed that the dipole layer is zero.* This result is in disagreement with the best observed value, which is about 4 ev. The discrepancy suggests either that the width of the occupied region of levels is much less than these investigators have found, or that beryllium usually has a tightly bound surface layer of electronegative atoms.

91. The Work Function in Non-metallic Crystals.—All the fundamental principles used in the previous section in discussing the work function of metals can also be applied to nonmetals. In general, the energy required to remove an electron from

the solid or to put one in depends both upon the volume characteristics of the substance and upon the surface dipole layer. As an example, let us consider the energy required to put an electron into a neutral sodium chloride crystal. This energy is less than the energy required to remove an electron from the uppermost levels of the filled band by an amount equal to the gap between the filled and unfilled bands (cf. Fig. 4).

We shall consider a somewhat idealized case in which the surface is a (100) plane and in which the ions near the surface retain the same relative positions as those in the interior. Under these conditions, the ordinary lattice potential along the dotted line in Fig. 5 is zero because

¹ BARDEEN, *op. cit.*

points on this line are equidistant from positive and negative ions. This statement is rigorously true inside the crystal only if we assume that the ions act as point charges. We shall see in Chap. XIII that the valence electron is distributed principally about the sodium ions, very much as in metallic sodium. For this reason, we may assume that the energy of the electron is equal to the energy of a sodium ion in the Madelung field at the position of a sodium ion plus the additional energy by which this level is lowered by the development of band structure. The Madelung potential at a sodium ion is $1.74e^2/r_0$, and the ionization potential of a sodium atom is $-0.19e^2/a_h$. Since the sodium-sodium distance in the salt is about the same as in the metal, we shall assume that the depression due to band formation is also the same. According to Fig. 2b, Chap. X, this is $0.11e^2/a_h$. Hence, the work function is¹

$$\begin{aligned}\varphi_1 &\cong 1.74 \frac{e^2}{r_0} - 0.19 \frac{e^2}{a_h} - 0.11 \frac{e^2}{a_h} \\ &= 0.03 \frac{e^2}{a_h} \cong 0.8 \text{ ev.}\end{aligned}$$

This actually is the sum of the internal and surface contributions; however, there is no surface dipole layer in the present case because the surface is a (100) plane that contains equal numbers of positive and negative charges. If we were to deal with another plane, such as a (111) plane, or were to alter the interionic distances near the surface, we could compute the surface dipole term by computing the Madelung potential at a sodium ion. The difference between this value and $1.74e^2/r_0$ would then be $-4\pi eP_n$.

It is doubtful whether the energy gained by an electron on entering an insulator is always as small as the value computed above for a typical alkali halide. The photoelectric work function of insulators, such as cuprous oxide, that absorb in the visible is² of the order of 5 ev, a fact indicating that the width of the forbidden region is about 2 ev and the work function φ_1 , discussed above, about 3 ev.

¹ A similar value has been obtained by N. F. Mott, *Trans. Faraday Soc.*, **34**, 500 (1938) using slightly different reasoning.—

² See, for example, A. L. HUGHES and L. A. DuBRIDGE, *Photoelectric Phenomena* (McGraw-Hill Book Company, Inc., New York, 1932); R. FLEISCHMANN, *Ann. Physik*, **5**, 73 (1930); R. J. CASHMAN (paper 179, program Washington Meeting, American Physical Society, 1940).

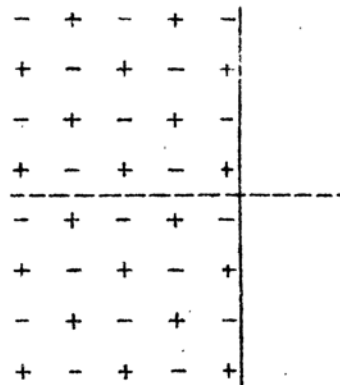


FIG. 5.—The potential along the dotted line is zero because points on this line are equidistant from positive and negative charges.

92. Thermionic Emission and the Temperature Dependence of the Work Function*.—In Sec. 30, Chap. IV, we derived the Richardson-Dushman equation for the thermionic electron emission from unit area of a metal

$$I = A(1 - r)T^2 e^{-\frac{W}{kT}}. \quad (1)$$

Here, W is the work function, which was assumed to be a constant for the entire surface, A is a universal constant, $120 \text{ amp/cm}^2\text{-deg}^2$, and r is the electronic reflection coefficient. In this section, we shall reexamine the relations that enter into Eq. (1) in the light of the previous work of this chapter.

Let us suppose that the metal is at temperature equilibrium with an external electron atmosphere. If we may assume that the electron cloud behaves as a perfect gas, which is a reasonable assumption as long as the density is small, the number of electrons that pass from the outside to the inside per unit area in unit time is

$$\frac{p(1 - r)}{(2\pi mkT)^{\frac{1}{2}}}, \quad (2)$$

where p is the external pressure, m is the electron mass, and k is Boltzmann's constant. This should be equal to the number that evaporates from a unit area of the surface since the system is at equilibrium. Hence, the thermionic current is

$$I = \frac{ep(1 - r)}{(2\pi mkT)^{\frac{1}{2}}}. \quad (3)$$

The equilibrium pressure $p(T, V)$, which is a function of the temperature T and volume V of the crystal, may be related to the heat ΔH required to sublime 1 mol of electrons from the metal at constant pressure by the Clausius-Clapeyron equation

$$\left(\frac{\partial \log p}{\partial T}\right)_V = \frac{\Delta H}{RT^2}, \quad (4)$$

where R is the gas constant. If the specific heat of the electrons inside the metal is neglected,

$$\Delta H = N_A W(V, T) + \frac{5}{2}RT \quad (5)$$

where $W(V, T)$ is the work function of the metal when the volume is V and the temperature is T , and $5R/2$ is the molar heat at constant pressure of the electron gas. The integral of (4) is

$$\begin{aligned} \log p &= \int \frac{\Delta H}{RT^2} dT + C = -\frac{\Delta H}{RT} + \int \left(\frac{d\Delta H}{dT}\right)_V \frac{dT}{RT} + C' \\ &= -\frac{\Delta H}{RT} + \frac{\Delta S}{R} \end{aligned} \quad (6)$$

where

$$\Delta S = \int \left(\frac{d\Delta H}{dT} \right)_v \frac{dT}{T} + C'R$$

is the entropy change that accompanies evaporation of a mol of electrons. We may substitute ΔH from Eq. (5). According to the third law of thermodynamics,¹ the constant term in ΔS should be chosen in such a way that the entropy change associated with $W(V, T)$ is

$$\int_0^T \left[\frac{dW(V, T)}{dT} \right]_v \frac{dT}{T}$$

and the entropy change of the gas is

$$R\left(\frac{5}{2} \log T + \frac{5}{2} + j\right)$$

where

$$j = \log \frac{2(2\pi m)^{3/2}}{h^3} k^2$$

is the chemical constant. Thus, (6) may be written

$$\log p = -\frac{W(V, T)}{kT} + \frac{1}{k} \int_0^T \left[\frac{dW(V, T)}{dT} \right]_v \frac{dT}{T} + \frac{5}{2} \log T - j. \quad (7)$$

Substituting this in (3), we obtain

$$I = A(1 - r)T^2 e^{-\frac{W(V, T)}{kT} + \frac{1}{k} \int_0^T \left(\frac{dW(V, T)}{dT} \right)_v \frac{dT}{T}} \quad (8)$$

where A is the coefficient that occurs in Eq. (1).

It should be noted that Eq. (8) differs from Eq. (1) by the factor

$$e^{\frac{1}{k} \int_0^T \left(\frac{dW}{dT} \right)_v \frac{dT}{T}} \quad (9)$$

which is unity for all temperatures only if $(dW/dT)_v = 0$. The appearance of this term indicates that the method used to derive Eq. (1) is faulty whenever the work function is temperature-dependent. Since the correction term (9) arises from the entropy of electrons in the solid, we see that the simple model used to derive (1) is in error because we neglected interactions between the electrons and the solid that cannot be described adequately by a simple potential barrier. Equation (8) could be derived on the basis of statistical mechanics, but it would be necessary to consider the entire solid in doing so.

Before investigating the importance of the temperature dependence of W , we shall introduce convenient definitions of the work function and

¹ See P. S. EPSTEIN, *Textbook of Thermodynamics* (John Wiley & Sons, Inc., New York, 1937).

of the thermionic coefficient that were developed by Becker and Brattain.¹ In ordinary thermionic experiments, the emitted current I is measured over a certain temperature range, and the quantity $\log (I/T^2)$ is then plotted as a function of $1/T$. The experimental work function and thermionic coefficient, $W^*(T)$ and $A^*(T)$, are defined, respectively, as the slope of this curve and intercept of the tangent on the $\log (I/T^2)$ axis (cf. Fig. 6); that is,

$$\begin{aligned} W^* &= kT^2 \left[\frac{d \log (I/T^2)}{dT} \right]_p, \\ \log A^* &= \log \frac{I}{T^2} + T \left[\frac{d \log (I/T^2)}{dT} \right]_p, \\ &= \log \frac{I}{T^2} + \frac{W^*}{kT}. \end{aligned} \quad (10)$$

These derivatives are taken at constant pressure because the specimen is kept in a vacuum during the experiments.

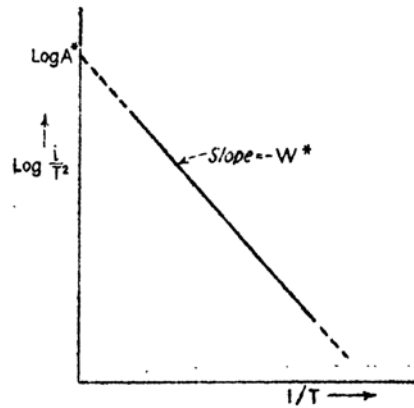


FIG. 6.—Diagrammatic representation of the definitions of A^* and W^* .

If Eq. (8) is substituted in these equations and if the relation

$$\begin{aligned} \left(\frac{\partial}{\partial T} \right)_p &= \left(\frac{\partial}{\partial T} \right)_v + \left(\frac{\partial V}{\partial T} \right) \left(\frac{\partial}{\partial V} \right)_T \\ &= \left(\frac{\partial}{\partial T} \right)_v + \alpha V \left(\frac{\partial}{\partial V} \right)_T \end{aligned} \quad (11)$$

is used, where $\alpha = (\partial V / \partial T) / V$ is the coefficient of volume expansion, it is found that

¹ J. A. BECKER and W. H. BRATTAIN, *Phys. Rev.*, **45**, 694 (1934).

$$W^* = T^2 \left[\left(\frac{\partial}{\partial T} \right)_V + \alpha V \left(\frac{\partial}{\partial V} \right)_T \right] \left[-\frac{W(V, T)}{T} + \int_0^T \left(\frac{\partial W}{\partial T} \right)_V \frac{dT}{T} \right]$$

$$= W(V, T) + \alpha T^2 V \left(\frac{\partial}{\partial V} \right)_T \left[-\frac{W(V, T)}{T} + \int_0^T \left(\frac{\partial W}{\partial T} \right)_V \frac{dT}{T} \right], \quad (12)$$

$$\log \frac{A^*}{(1-r)A} = \frac{1}{k} \int_0^T \left(\frac{\partial W}{\partial T} \right)_V \frac{dT}{T} + \frac{\alpha V}{k} \left(\frac{\partial}{\partial V} \right)_T \left[-\frac{W(V, T)}{T} + \int_0^T \left(\frac{\partial W}{\partial T} \right)_V \frac{dT}{T} \right]. \quad (13)$$

Wigner¹ has attempted to estimate some of the terms in (13), using the expression for the work function of sodium developed in Sec. 90. His treatment is not rigorous, but it shows that the terms on the right-hand side of (13) are not negligible even for simple metals and that the value of $\log \frac{A^*}{A(1-r)}$ is of the order unity, that is, lies between +6 and -5. $W(V, T)$ may be related to the work function $W(V, 0)$ computed in the previous sections in the following way: $W(V, T)$ is defined as the energy required to remove an electron from the metal at temperature T . The total energy of the crystal when no electrons have been removed is

$$E = E_D \left(\frac{\Theta_D}{T} \right) + E_e$$

where E_D is the vibrational energy of the lattice, which we shall express in terms of Debye's characteristic temperature Θ_D , and E_e is the electronic energy, which is not temperature-dependent if the small electronic specific heat is neglected. If n electrons are removed at temperature T , the new energy E' is

$$E' = E_D + E_e + nW(V, T). \quad (14)$$

Suppose that n electrons are removed at absolute zero of temperature instead and that the metal is then raised to temperature T . The resulting energy, which is again E' , is

$$E_D \left(\frac{\Theta_D + \Delta\Theta_n}{T} \right) + E_e + nW(V, 0) \quad (15)$$

where $\Delta\Theta_n$ is the change in the characteristic temperature that results from the removal of n electrons. Equating (14) and (15), we obtain

$$W(V, T) = \frac{1}{n} \frac{\Delta\Theta_n}{T} E_D \left(\frac{\Theta_D}{T} \right) + W(V, 0). \quad (16)$$

¹ E. WIGNER, *Phys. Rev.*, **49**, 696 (1936). See also K. F. HERZFELD, *Phys. Rev.*, **35**, 248 (1930).

Wigner finds that the contribution to (13) from the last term¹ in Eq. (16) is about -3.6 for sodium. It turns out that this result is balanced somewhat by the contribution from the first term in (16). There is, however, no reason for expecting the two terms to cancel.

Direct measurements of the temperature coefficient of $W(V, T)$ in Eq. (5) have been carried out by Krüger and Stabenow² for molybdenum, tungsten, and tantalum. These workers measured the heat lost by a wire during thermionic emission. The temperature variation is within the experimental error in molybdenum and has the value $0.6 \cdot 10^{-4}$ ev/deg per electron in both tungsten and tantalum. If the integral in the expression (9) is evaluated with the use of this coefficient and with the assumptions that $(\partial W/\partial T)$ approaches zero near absolute zero, so that the contribution from the lower limit of integration may be neglected, and the temperature coefficients at constant volume and constant pressure are practically the same, it is found that $\log \left[\frac{A^*}{A(1-r)} \right] \sim -5.1$. Hence, A^* would be about 0.66 amp/deg-cm² if r were zero. Unfortunately, these workers did not measure A^* on the specimens for which this work was carried out. The values of A^* measured by other workers are about one hundred times larger.

¹ Since the last term is independent of temperature, it is the work function discussed in the preceding sections of this chapter.

² F. KRÜGER and G. STABENOW, *Ann. Physik*, 22, 713 (1935).

CHAPTER XII

THE EXCITED ELECTRONIC STATES OF SOLIDS

93. Introduction.—Both the Bloch and the Heitler-London approximations have been used to treat the excited states of solids. Although these methods have not been rigorously tested in particular cases, qualitative and semiquantitative arguments may be used to show that one approximation is more suitable than another in a given case. For example, we may expect that the Bloch approximation is more suitable¹ to use in a discussion of the excited states of metals because it alone leads² to the low-lying continuous, conducting levels that are characteristic of these solids (*cf.* Sec. 66). Similarly, we shall find that the Heitler-London approximation is more applicable to the lower levels of molecular and ionic crystals.

In this chapter, we shall discuss the general principles upon which computations of excited states are now based and shall also present some simple results. This discussion begins with a survey of the uses of the Bloch method and is followed by a similar survey of the Heitler-London scheme. Problems in which intermediate approximations are applicable will be discussed in later sections.

94. Excited States in the Band Scheme.—The band scheme is based upon a one-electron approximation in which the ψ have the form

$$\psi_{\mathbf{k}} = \chi_{\mathbf{k}} e^{2\pi i \mathbf{k} \cdot \mathbf{r}} \quad (1)$$

and satisfy Fock's equations

$$-\frac{\hbar^2}{2m} \Delta \psi_{\mathbf{k}} + (V + A) \psi_{\mathbf{k}} = \epsilon(\mathbf{k}) \psi_{\mathbf{k}} \quad (2)$$

where V is the coulomb or Hartree potential and A is the exchange operator. The entire wave function of the solid may be constructed from determinants of wave functions of type (1). V and A are not appreciably altered if one of the $\psi_{\mathbf{k}}$ in the set is replaced by another, since the ψ

¹ Direct evidence for the qualitative correctness of the Bloch approximation is also obtained from a study of the soft X-ray emission spectra of metals (*cf.* Sec. 104).

² It should be pointed out that the Heitler-London scheme would also include the Bloch states if we considered atomic wave functions of the type associated with continuous spectra as well as the localized wave functions of the type associated with the discrete atomic levels. We shall explicitly avoid including the first type of wave function, however.

extend over the entire lattice and have small amplitude in any given cell. Hence, V and A may be chosen the same for both normal and excited states. It follows from Koopmans' theorem that $\epsilon(\mathbf{k}) - \epsilon(\mathbf{k}')$ is the energy required to excite the crystal from a given state to the one in which $\psi_{\mathbf{k}}$ is replaced by $\psi_{\mathbf{k}'}$. Thus, the possible excited levels may be obtained from the one-electron energy diagrams of the zone scheme. The highest occupied zone is not completely filled in a metal, whence the lowest states of the solid as a whole have a quasi-continuous system of energy levels. Since the conduction properties of metals require this type of continuum, the band approximation is naturally suited for a semiquantitative description of these solids. As we have seen in Chap. X, the process of improving the zone approximation for a metal does not simply effect a compromise between the Bloch and Heitler-London approximations but consists in treating correlations more accurately. This does not mean that some atomic properties are not retained in passing from the free atoms to the solid, for the functions $\chi_{\mathbf{k}}$ in Eq. (1) preserve many of the features of atomic wave functions.

The band scheme can be applied to insulators as well as to metals. In these cases, the highest occupied zone is completely filled in the normal state, so that the first excited level is a finite distance above the lowest level. We have seen in Sec. 64 that the lowest state of an ionic or molecular crystal is described with equal accuracy by either the Bloch or the Heitler-London scheme. We shall see in the next section, however, that the Heitler-London scheme leads to excited levels that are not contained in the Bloch approximation. For this reason, the zone scheme is not always adequate for a qualitative description of insulators.

95. Excited States in the Heitler-London Scheme.—Let us apply the Heitler-London scheme to sodium chloride, which is a typical insulator. We shall attempt to follow the behavior of the lowest atomic and ionic energy levels as the ions are brought together to form the normal lattice and are kept in crystalline arrangement during the process. The excited ionic levels will be neglected for the moment since they cannot be treated properly without including a discussion of continuous spectra.

At infinite separation, the ionic and atomic levels of sodium and chlorine are as illustrated in the right-hand side of Fig. 1 in which the halogen-ion level is given relative to that of the neutral atom and the level of neutral sodium is given relative to that of Na^+ . Thus, the normal state of Cl is at -3.8 ev, and the lowest level of neutral sodium is at -5.2 ev. Other levels are neglected for the present. The minimum energy required to transfer an electron from a halogen ion to an alkali ion is -1.4 ev at infinite separation. Since this value is negative, the infinitely separated system is more stable as a set of neutral atoms than as a set of ions. This situation is gradually altered as the ions approach

one another. If r_0 is the distance between nearest neighboring ions in the lattice, the electrostatic potential at the negative ions is $1.748e^2/r_0$ and that at the positive ions is the negative of this. Thus, the halogen-ion levels and alkali-ion levels are respectively raised and lowered in accordance with the equations

$$\left. \begin{aligned} \epsilon(\text{Cl}^-) &= -3.8 - \frac{47.3}{(r_0/a_h)} \\ \epsilon(\text{Na}^+) &= -5.2 + \frac{47.3}{(r_0/a_h)} \end{aligned} \right\} \quad (1)$$

in which the unit of energy is the electron volt. These equations are valid only as long as the ions do not overlap appreciably. When they

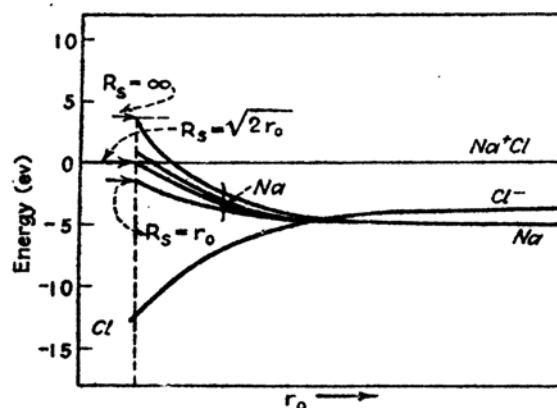


FIG. 1.—Levels of the entire solid in the Heitler-London approximation based on an ionic model. At large separations the state of neutral atoms is most stable because the ionisation potential of metal atoms usually is larger than the electron affinity of the halogen atom. This situation is reversed as the atoms are brought together, because the Madelung energy favors the ionic state. Corresponding to each level of the metal atom there are an infinite number of levels of the entire solid, each of which is related to a particular value of R_s in Eq. (3).

do overlap, additional energy terms should be added to (1) in order to include the effects of exchange and correlation interactions. Since these terms are only about 10 per cent of the electrostatic terms, we shall neglect them temporarily. The energy $\Delta\epsilon$ required to transfer an electron from a halogen ion to a distant alkali metal ion then is the difference between the two terms in (1), namely,

$$\Delta\epsilon = -1.4 + \frac{94.6}{(r_0/a_h)} \text{ (ev).} \quad (2)$$

This is 16.5 ev for the normal interionic distance of $5.29a_h$ in sodium chloride. There are, however, an infinite number of levels lying below this one, for an energy e^2/R less than (2) is required to transfer the excited

electron from a halogen ion to an alkali ion that is at a finite distance R . That this statement is true may be seen from the fact that the Madelung potential at a given alkali ion is decreased by e^2/R if an electron is removed from a chlorine atom at distance R . Thus, the normal and excited levels

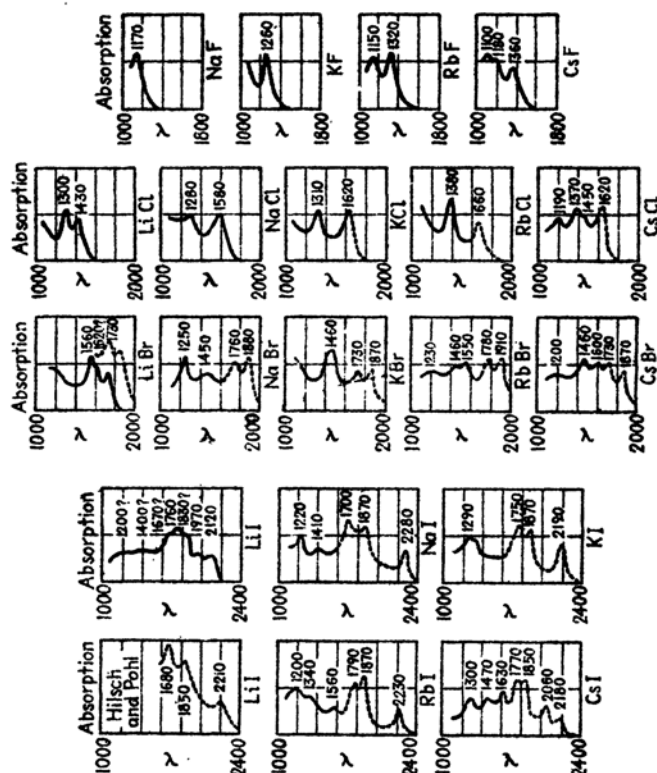


FIG. 2.—The ultraviolet absorption bands of the alkali halide crystals. The dotted portions of the curves represent measurements by Hilsch and Pohl. [After Schneider and O'Bryan, *Phys. Rev.*, **51**, 293 (1937).]

are disposed as in Fig. 1, in which the excited levels are separated from the ground state by an amount

$$\Delta\epsilon_s = -1.4 + \frac{94.6}{(r_0/a_h)} - \frac{27.08}{(R_s/a_h)} \text{ (ev).} \quad (3)$$

At the observed lattice distance, the first excited level, which is 11.3 ev above the ground state, corresponds to $R_s = 5.29a_h$, the distance between nearest neighboring ions, whereas the next level, which is 12.9 ev above the ground state, corresponds to $R_s = \sqrt{2}r_0$. Hilsch and Pohl¹ first pointed out that the difference in the quantum energies of the first two

¹ R. HILSCH and R. W. POHL, *Z. Physik*, **59**, 812 (1930).

ultraviolet absorption bands of sodium chloride is very nearly equal to the difference between these computed excited levels. The first band has its peak at 1580 Å (cf. Fig. 2), and the second has its peak at 1280 Å; these values correspond to energies of 7.8 and 9.6 ev, respectively. The difference between these energies is 1.8 ev, which agrees closely with the difference of 1.6 ev for the calculated absorption energies. von Hippel¹ has attempted to refine this simple calculation of the first excited levels by including corrections for the perturbation of the neutral alkali and halogen atoms. His computed values and the observed values are compared in Table LXXI.

TABLE LXXI.—COMPARISON OF COMPUTED AND OBSERVED ENERGY DIFFERENCES BETWEEN THE GROUND STATE AND THE FIRST EXCITATION STATE OF THE ALKALI HALIDES (AFTER VON HIPPEL)

	Observed, ev	Calculated, ev
NaCl	7.8	8.3
KCl	7.6	8.0
RbCl	7.4	7.7
LiBr	6.7	8.1
NaBr	6.5	7.4
KBr	6.6	7.3
RbBr	6.4	7.1

The excited discrete levels, shown in Fig. 1, actually are highly degenerate, for the excited electron may be removed from any one of the N halogen ions of the crystal and may be carried to any one of its g_s neighboring alkali metal ions that are at distance R_s without altering $\Delta\epsilon_s$. Thus, the levels are Ng_s -fold degenerate. Since $\sum_s g_s$ is equal to N , the total number of alkali metal ions, we see that the total number of excited levels in the system is N^2 .

In the simple model used above, the first excited level is sixfold degenerate, if spin is neglected, for each chlorine ion has six equidistant neighboring alkali ions. This degeneracy is partly accidental, for the six functions do not have the proper symmetry to have the same energy in a cubic crystal. Thus, the degenerate levels would split if interatomic interactions were taken into account. In first approximation, the new functions should be linear combinations of the six functions ψ_i that are localized on the separate alkali ions. The electronic distribution of the new functions should be spread over all six neighboring ions. The lowest state evidently is the symmetrical function that is formed by adding all

¹ A. VON HIPPEL, *Z. Physik*, 101, 680 (1936)

six ψ_i , and is analogous to an atomic s function. Above this, there are a triply degenerate set that is analogous to the three atomic p functions

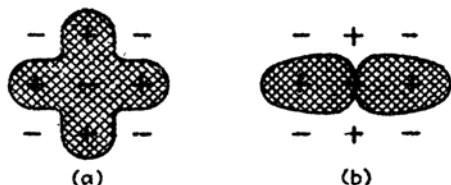


FIG. 3.—Schematic diagram of the wave functions of excited electrons in ionic crystals. Function a is the analogue of an atomic s function and distributes the electronic charge equally about all of the neighboring positive ions. Function b is the analogue of a p function and has opposite signs in the two "wings."

and a doubly degenerate level that has no atomic analogue. Two of the four possibilities for a two-dimensional case are shown in Fig. 3.

Let us now consider the continuum of levels corresponding to ionization of the halogen ion. The continuum remains essentially unchanged as long as the ions are far apart; however, the free electrons begin to be perturbed appreciably when the ions occupy a considerable

fraction of the volume of the crystal. Since the new levels should be computed by determining the wave functions of the free electrons in the field of the entire crystal, they are obviously the same as the excited levels that are computed on the basis of the Bloch approximation. We know, however, that the Bloch bands contract into the levels of the excited states at infinite separation. Hence, we may conclude that the levels of the continuum tend to cross and combine with the excited discrete levels of the Heitler-London scheme, broadening these lines into bands. The extent to which this broadening actually takes place depends upon the interatomic distance. Only the series limit is affected in a case in which the lattice is highly extended, whereas in the opposite extreme of a highly compressed lattice the continuum overlaps even the lowest level, making the solid a metal. In the intermediate case the lowest level is discrete, as shown schematically in Fig. 4. One important case is that corresponding to the point A of Fig. 4 in which the continuum has overlapped

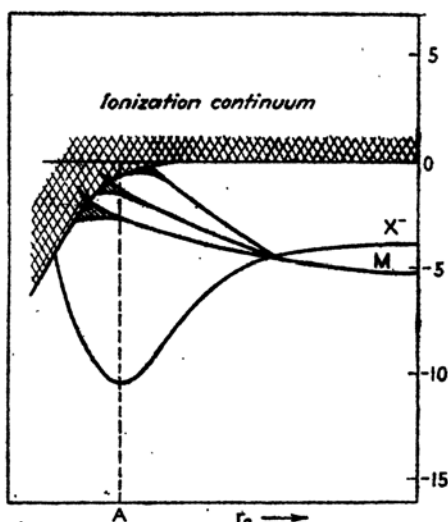


FIG. 4.—Schematic diagram showing the behavior of the levels of an ionic crystal as the atoms are brought together (see also Fig. 1). At large distances the neutral system is stable, whereas at intermediate distances the ionic system is stable. The bands corresponding to the ionization continuum broaden and spread and may actually overlap all of the discrete levels. The broadening of the excited nonconducting levels at A corresponds to the formation of excitation bands, which is discussed in the next section. The energy units are electron volts for NaCl.

all the excited levels except the lower ones. In this case, the first excited levels should be determined approximately by means of the Heitler-London scheme, as von Hippel has done, whereas the higher states should be determined by the Bloch approximation. Experimental evidence, which will be presented in the next chapter, indicates that this corresponds closely to the state of affairs in the simpler ionic crystals. It should be noted that we have removed some of the degeneracy of the lowest excited levels in Fig. 4 before they merge with the continuum. The origin of this effect is discussed in the next section.

Frenkel,¹ who was one of the first to discuss correctly the possible relationships between normal and excited states in insulators, has called the lower excited levels of Fig. 4 that have not mingled with the continuum at A "excitation levels," since they are analogous to the excited states of atoms. Similarly, he has called the higher levels that should be treated by the band approximation "ionization levels," since they are analogous to the ionized states of atoms. The ionization states have already been discussed in Chap. VIII. We know from this presentation that the crystal should become photoconducting when excited to these levels since the excited electron is then free to roam throughout the lattice. On the other hand, the excited electron remains fixed relative to the atom from which it came in the excitation states. Hence, we should not expect photoconductivity to accompany optical excitation to these states. Photoconductivity actually does not seem to occur as a result of absorption in the first fundamental absorption bands of the alkali halides, a fact which indicates that the first excited levels in these solids correspond to excitation states.²

It should be mentioned in passing that Valasek³ has presented good experimental evidence that the excited X-ray levels in salts such as sodium chloride and potassium chloride should be described by the atomic scheme, or by the exciton scheme which is discussed in the next section, rather than by the band scheme.

¹ J. FRENKEL, *Phys. Rev.*, **37**, 17 (1931); **37**, 1276 (1931).

² The cases of the alkaline earth oxides and sulfides, such as zinc oxide and zinc sulfide, are still uncertain, for the structure of the absorption bands of these solids has not been thoroughly investigated. A recent experimental investigation of this problem for the alkali halides has been carried out by L. P. Smith and J. N. Ferguson (see paper 177, program of Washington Meeting, American Physical Society, 1940). These observers find photoconductivity in the long wave length tail of the fundamental bands, but not in the interior. According to the exciton viewpoint (see next section) this conductivity arises either from direct ionization of impurity atoms or lattice-defect atoms (Chap. XIII), or from secondary ionization of these atoms by excitons. In view of results of this kind, however, it must be admitted that there is no conclusive experimental evidence that photoconductivity would not occur at least in the tail of the fundamental band of a pure perfect crystal.

³ J. VALASEK, *Phys. Rev.*, **47**, 896 (1935); **53**, 274 (1938).

Let us consider another example of a system to which the Heitler-London method may be applied, namely, that of an insulating crystal which contains a neutral impurity atom. As we shall see in the next chapter, this system corresponds to a semi-conductor, such as ZnO, in which there are interstitial zinc atoms. At present, we shall be interested in the case in which the ionization potential of the neutral atom is less than the energy of the first absorption band of the pure crystal and the interstitial atom occupies a position of zero potential at which it is symmetrically surrounded by positive and negative ions. These conditions are approximately fulfilled in zinc oxide. We may

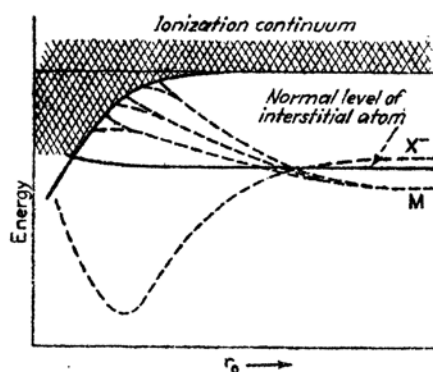


FIG. 5.—Behavior of the levels of an interstitial atom in an ionic crystal. The dotted lines correspond to the energy-level curves of the bulk material. In this case, the excitation and ionization energies of the bulk material are larger than those of the interstitial atom at the equilibrium position, which corresponds to the minimum of the lower curve. At small interatomic distances, the lowest level of the impurity atom may merge with the ionization continuum. This probably does not happen in actual cases.

expect, by analogy with the case of sodium chloride, that the energy required to transfer an electron from a negative to a positive ion is negative at infinite separation, for the electron affinities of negative ions are usually smaller than the ionization potential of metal atoms. This situation is altered as the atoms are brought together, for the positive ions are surrounded principally by negative ones, and vice versa. At infinite separation, the ionization continuum of an interstitial atom lies at the same position as that of the negative ions and must blend into the Bloch bands of the entire solid as the atoms are brought together, because these are the levels of a free electron in the lattice. The normal state of an interstitial atom behaves in the manner shown symbolically in Fig. 5

In this case, the level remains discrete until it merges with the spreading ionization continuum.

96. Excitation Waves.—As we have seen in the preceding section, there is reason to believe that there are nonconducting excited levels beneath the conducting states described by the Bloch approximation in insulators such as sodium chloride. We shall attempt to describe these excitation states more fully in the present section, using a simple model that was considered first by Frenkel¹ and by Peierls² and more recently by Slater and Shockley.³

¹ FRENKEL, *op. cit.*; *Physik. Z. Sowj.*, **9**, 158 (1936).

² R. PEIERLS, *Ann. Physik*, **13**, 905 (1932).

Let us consider a system of identical atoms that possess only one electron and are ordered in a simple crystalline array. For simplicity, we shall neglect electron spin and shall assume that both the lowest and first excited levels are nondegenerate. It will be evident that the important conclusions which may be drawn for this simple system are valid for a similar system of atoms of any type in another lattice.

Let us assume that the atoms are located at the positions

$$\mathbf{r}(n) = n_1\boldsymbol{\tau}_1 + n_2\boldsymbol{\tau}_2 + n_3\boldsymbol{\tau}_3 \quad (1)$$

where the $\boldsymbol{\tau}$ are the primitive translations of the lattice and the n range over all integer values. We shall let ψ_n and ψ'_n be the normal and excited wave functions of the electron at $\mathbf{r}(n)$, and we shall assume¹ that we are dealing with a case in which the wave functions on different atoms overlap so little that the ψ are practically the same as atomic functions. ψ_n and ψ'_n then are orthogonal to one another and to the wave functions of electrons on other atoms. A wave function for the lowest state of the entire system may be constructed by taking a determinant of the form

$$\Psi_0 = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{r}_1) & \dots & \psi_1(\mathbf{r}_N) \\ \vdots & & \vdots \\ \psi_N(\mathbf{r}_1) & \dots & \psi_N(\mathbf{r}_N) \end{vmatrix} \quad (2)$$

where N is the total number of electrons and of atoms. The mean energy of this wave function is

$$E_0 = \int \Psi_0^* H \Psi_0 d\tau \quad (3)$$

where H is the Hamiltonian of the entire system. This integral may be expanded in terms of the eigenvalues of the ψ and the exchange and coulomb integrals between atoms.

Let us consider next wave functions for the case in which one atom is excited. The system of wave functions Ψ_n obtained by replacing ψ_n in (1) by ψ'_n are not the best excited wave functions, for the integrals

$$E_{mn} = \int \Psi_m^* H \Psi_n d\tau \quad m \neq n \quad (4)$$

do not vanish. It is easy to show in fact that, under the orthogonality conditions on the ψ_n and ψ'_n that were assumed above, (4) is equal to

$$e^2 \int \frac{\psi'_m(\mathbf{r}_1)\psi_n(\mathbf{r}_2)\psi_m(\mathbf{r}_1)\psi'_n(\mathbf{r}_2)}{r_{12}} d\tau_{12} - e^2 \int \frac{\psi'_m(\mathbf{r}_1)\psi_n(\mathbf{r}_2)\psi_m(\mathbf{r}_2)\psi'_n(\mathbf{r}_1)}{r_{12}} d\tau_{12} \quad (5)$$

¹ It should be emphasized that the following approximation is accurate only when the atoms are not too close together. There may be no nonconducting excited states if the atoms are pushed together sufficiently (cf. Secs. 66 and 95).

when $m \neq n$. When $m = n$, (4) may be expanded in terms of the energy levels of the normal and excited atoms and the exchange and coulomb integrals between normal and excited atoms. We shall designate the difference between E_{nn} and E_0 , which is the order of magnitude of the resonance energy of an isolated atom, by

$$\epsilon = E_{nn} - E_0. \quad (6)$$

We shall attempt to diagonalize the N -dimensional matrix formed by the E_{mn} . This process is equivalent to finding those linear combinations Ψ' of the Ψ_n which have the form

$$\Psi' = \sum_n a_n \Psi_n, \quad (7)$$

in which the a satisfy the equations

$$\sum_m a_m E_{mn} = E' a_n. \quad (8)$$

It follows from the symmetry of the crystal that E_{mn} depends only upon the difference between the integer sets m and n . This fact suggests that we should reduce the N equations (8) to the same form by making the substitution

$$a_m = a_k e^{2\pi i \mathbf{k} \cdot \mathbf{r}(m)} \quad (9)$$

where \mathbf{k} is a vector in the reciprocal lattice of the crystal. Equation (8) then becomes

$$E'_k = E_{nn} + \sum_l' E_{n,n+l} e^{2\pi i \mathbf{k} \cdot \mathbf{r}(l)} \quad (10)$$

where $E_{n,n+l} = E_{m+r,m+r+l}$, if l and r are arbitrary integer sets. The prime in this summation indicates that the term for $l = 0$, which appears outside the sum, is to be excluded. The normalized wave function associated with the wave number \mathbf{k} may be found by substituting Eq. (9) in (7) and is

$$\Psi_k = \frac{1}{\sqrt{N}} \sum_n e^{2\pi i \mathbf{k} \cdot \mathbf{r}(n)} \Psi_n. \quad (11)$$

The independent values of \mathbf{k} range over a single zone if there is one atom per unit cell and the excited state is nondegenerate, whereas they range over αg zones if there are α atoms per unit cell and the degeneracy of the excited level is g .

In view of the assumption that overlapping is small, it is reasonable to assume that E_{mn} is zero for all except nearest neighbors. If I is the

value of E_{nn} in this case, (9) may be written

$$E'_k = E_{nn} + I \sum_{\tau} e^{2\pi i \mathbf{k} \cdot \tau} \quad (12)$$

where τ is to be summed over nearest neighbors. Equation (12) is simply

$$E'_k = E_{nn} + I(\cos 2\pi k_x a + \cos 2\pi k_y a + \cos 2\pi k_z a) \quad (13)$$

in the case of a simple cubic lattice having lattice constant a . This equation, which is similar to the equations derived in Sec. 65 for the case of narrow conduction bands, shows that the excited levels form a band the width of which is of the order of magnitude I . Equations (10) and (11) are valid only as long as I is appreciably smaller than ϵ . Otherwise, more atomic states must be considered in diagonalizing the Hamiltonian matrix.

The excited atom is not localized in the states described by (11); instead it is distributed throughout the crystal. By constructing wave packets, it is easy to show that the excitation moves with the group velocity

$$\mathbf{v} = \frac{1}{\hbar} \text{grad}_{\mathbf{k}} E_{\mathbf{k}} \quad (14)$$

in the energy state $E_{\mathbf{k}}$ (cf. Sec. 68).

The current associated with a given wave function $\Psi_{\mathbf{k}}$ is the mean value integral

$$\mathbf{I}_{\mathbf{k}} = \frac{e\hbar}{2mi} \int \sum_i (\Psi_{\mathbf{k}} \text{grad}_i \Psi_{\mathbf{k}}^* - \Psi_{\mathbf{k}}^* \text{grad}_i \Psi_{\mathbf{k}}) d\tau_i, \quad (15)$$

which may be reduced to

$$\frac{e\hbar}{2miN} \sum_i \sum_{m,n} [e^{2\pi i \mathbf{k} \cdot (\mathbf{r}(m) - \mathbf{r}(n))} \dots e^{2\pi i \mathbf{k} \cdot (\mathbf{r}(n) - \mathbf{r}(m))}] \int \Psi_m \text{grad}_i \Psi_n d\tau_i. \quad (16)$$

If Ψ_m and Ψ_n are expanded by substituting their determinantal form, it is found that $\int \Psi_m \text{grad}_i \Psi_n d\tau_i$ vanishes for $m \neq n$, for one or more vanishing integrals of the type

$$\int \psi_n \psi'_m d\tau, \quad \int \psi'_n \psi_m d\tau, \quad \int \psi_n \psi_m d\tau, \quad \int \psi'_m \psi'_n d\tau$$

appear in each term.¹ Hence, (15) is zero and the excitation waves carry no current. We may, if we choose, regard the excitation wave as though

¹ If the overlap integrals for immediate neighbors do not vanish, the current will not be strictly zero, but will be very small, corresponding to motion of an electron an atomic distance.

it were an uncharged particle, created by exciting the crystal, that may move about the lattice. This convenient concept was first introduced by Frenkel, who called the imaginary particle an "exciton."

The selection rules for optical transitions from the ground state Ψ_0 to the excitation state Ψ_k are determined by the integral

$$\int \Psi_0^* \left(\sum_i \text{grad}_i e^{-2\pi i \mathbf{r} \cdot \mathbf{r}_i} \right) \Psi_k d\tau(x_1, \dots, x_n) \quad (17)$$

where i is summed over all N electrons and \mathbf{n} is the wave-number vector for the light quantum.¹ Substituting from Equation (11), we find that (17) becomes

$$\sqrt{N} \sum_{\mathbf{n}} e^{2\pi i \mathbf{k} \cdot \mathbf{r}(\mathbf{n})} \int \Psi_0 \text{grad} e^{-2\pi i \mathbf{r} \cdot \mathbf{r}_n} d\tau(x_1, \dots, x_n) \quad (18)$$

where \mathbf{r} is the coordinate vector of any one of the electrons. If the determinantal form of the Ψ is used, the integral in this equation is reduced to

$$\frac{1}{N} \int \psi_n \text{grad} e^{-2\pi i \mathbf{r} \cdot \mathbf{r}_n} \psi'_n d\tau(x, y, z), \quad (19)$$

which determines the selection rules for optical transitions in isolated atoms. Ordinarily, \mathbf{n} is so small that $e^{-2\pi i \mathbf{r} \cdot \mathbf{r}_n}$ does not vary appreciably over a single atom and may be replaced by $e^{-2\pi i \mathbf{r} \cdot \mathbf{r}(\mathbf{n})}$. Thus, (17) becomes

$$\sum_{\mathbf{n}} e^{2\pi i (\mathbf{k} - \mathbf{n}) \cdot \mathbf{r}(\mathbf{n})} \frac{1}{\sqrt{N}} \int \psi_n \text{grad} \psi'_n d\tau = \sqrt{N} \left(\int \psi_n \text{grad} \psi'_n d\tau \right) \delta_{\mathbf{k}\mathbf{n}}. \quad (20)$$

It may be concluded that the transition probability is zero unless the condition

$$\mathbf{k} = \mathbf{n} \quad (21)$$

is fulfilled and unless the excited state ψ'_n is one to which transitions from ψ_n are allowed. Since \mathbf{n} is very small, (21) is equivalent to the condition that $\mathbf{k} = 0$.

If we consider, instead of the system described above, one such as sodium chloride, in which the chlorine ion has excited states when in the crystal, the practical problem of constructing the excited states is complicated by the fact that the constituent ions contain more than one electron. This should not affect the qualitative results of the preceding discussion, such as that the width of the excitation band increases with

¹ The optical properties associated with excitation bands are discussed in Chap. XVII.

increasing interionic interaction and that the excitation wave carries no current.

The lowest states of the electronegative ions in simple ionic crystals are *S*-like, since these ions have closed-shell configurations, whereas the first excited states have *P*-like symmetry, as we have seen in the last section. Since optical transitions between these types of level are allowed, the condition (21) determines the selection rules.

Wannier¹ has proposed a very simple semiquantitative method of looking at the excitation bands in insulators. If we remove an electron from the highest filled band of an insulator, we produce a positive charge that should be able to move about freely in an undistorted lattice. In the Bloch picture, the excited electron is independent of the positive hole and is also able to move about freely. Actually, the hole and the electron should attract one another with a force that is coulomb-like at large distances. Wannier has shown that the excitation bands are analogous to the discrete levels of a hydrogen atom in the sense that in these states the electron and hole revolve about one another in closed orbits. The different levels in a given band correspond to the different translational levels of an excited hydrogen atom. On the basis of this picture, Wannier has derived a set of simple approximate equations from which the wave functions and energy levels of the exciton may be determined.

¹ G. H. WANNIER, *Phys. Rev.*, **52**.

CHAPTER XIII

THE ELECTRONIC STRUCTURE OF THE FIVE SOLID TYPES

97. Introduction.—The present chapter, in which we shall present a survey of the electronic constitution of the normal and excited states of the five solid types, is the central chapter of the book since all the preceding chapters are preparatory for it. A large part of this discussion is necessarily qualitative and probably will remain so until computational technique has been developed much further. Thus, we shall use the one-

electron approximations freely in cases in which they do not lead to qualitatively incorrect results. In other cases, we shall employ the method of description in which the energy levels of the entire solid are used.

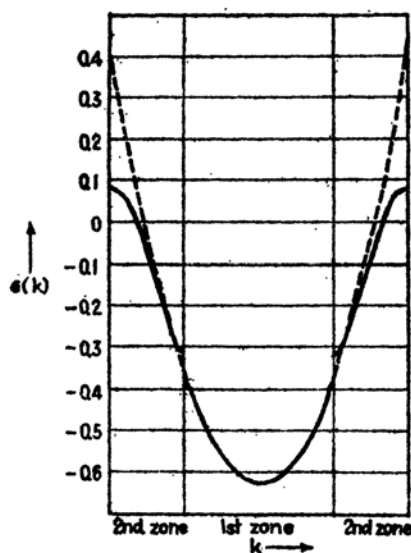


FIG. 1.—The $\epsilon(k)$ curve for the (110) direction of sodium (full line). The dotted line is the free-electron parabola. The energy scale is in Rydberg units. (After Slater.)

A. METALS

98. General Remarks.—Although the correlation terms that were discussed in Chaps. IX and X probably are important in a quantitative determination of any property of a metal, it is unlikely that they often affect the qualitative properties. The possible exceptions occur in connection with those low-temperature effects, such as superconductivity, which are not well understood at present. For this reason, we shall discuss the valence electrons of simple metals on the basis of the band approximation. Many of

the qualitative properties of d -shell electrons can also be treated adequately in this way. The method is not entirely satisfactory, however, for many other properties of d -shell electrons can be explained better with the Heitler-London approximation. This fact shows that neither of the one-electron schemes is very good in this case and that the d shells should be treated as a whole. This more accurate procedure has been used only in a few cases, such as in the

spin-wave theory of ferromagnetism which is developed in Chap. XVI. It is usually assumed at present that the accurate solution would yield the same result as the one-electron schemes in those cases in which the latter appear to give a satisfactory description of affairs.

99. Simple Metals. *a. The Alkali Metals.*—We discussed most of the known facts concerning the electronic levels of the alkali metals in Chap. X. A few additional results follow.

The zone structure of sodium¹ has been investigated by several workers. All the results show that the gaps are very narrow and that the effective electron mass is close to unity. In view of Shockley's investigation of the empty lattice by the cellular method (*cf.* Sec. 73), we may

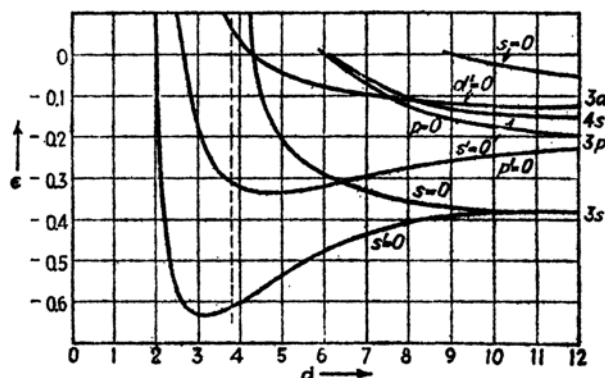


FIG. 2.—The dependence of the energy bands of sodium on interatomic distance d . It should be observed that the s - and p -level bands overlap strongly at the observed value of d . This behavior is characteristic of the simpler metals. The energy scale is in Rydberg units. (*After Slater.*)

say that the electrons in sodium are free, within the accuracy of this method. The zone structure determined by Slater is shown in Fig. 1, and the dependence of energy levels on interatomic distance is shown in Fig. 2. It should be noted that the s - and p -level bands overlap a great deal at the actual interatomic spacing. This overlapping of s and p levels is characteristic of all metals.

Bardeen² has pointed out that exchange terms have a very important effect on the density of electronic levels near the top of the filled region when the electrons are nearly free. In this case, the exchange energy is

$$\epsilon_{ex} = -0.306 \left(2 + \frac{1 - \frac{k^2}{k_0^2}}{k/k_0} \log \frac{|k_0 + k|}{|k_0 - k|} \right) \frac{e^2}{r_s} \quad (1)$$

¹ E. WIGNER and F. SEITZ, *Phys. Rev.*, **43**, 804 (1933); **46**, 509 (1934). J. C. SLATER, *Phys. Rev.*, **45**, 794 (1934); *Rev. Modern Phys.*, **6**, 209 (1934).

² J. BARDEEN, *Phys. Rev.*, **50**, 1098 (1936).

so that the total dependence of electronic energy on k is

$$\epsilon(k) = \epsilon_0 + \frac{\hbar^2}{2m}k^2 + \epsilon_{ex}. \quad (2)$$

The number of levels having values of k in the range from k to $k + dk$ is

$$dn = 8\pi k^2 V dk = 8\pi k^2 V \frac{1}{d\epsilon/dk} d\epsilon. \quad (3)$$

We may see from Fig. 5 of Chap. IX, that $d\epsilon_{ex}/dk$ is infinite when k is equal to k_0 . Thus, the density of levels is much smaller at $k = k_0$ than it would be if ϵ were simply a parabolic function of k . Since the electronic specific heat should be proportional to the density of levels in this region, it follows that the electronic specific heat of a free-electron gas, in which the exchange interaction is included, should be less than the value

$$k \frac{\pi^2}{2} \frac{kT}{\epsilon_0},$$

which was derived in Sec. 27. In fact, Bardeen has shown that the specific heat should vary as $-(\log T)/T$ at low temperatures when $\epsilon(k)$ has the form of Eq. (2). The low-temperature specific heat of the alkali metals has not been measured accurately enough to check this behavior. It is possible that correlation terms have an effect which may compensate for the effect of exchange.¹

This infinity in the slope of the exchange energy is accompanied by a singularity in curvature. In fact, the mean value of the second derivative of (1) becomes infinite as $\log |k_0 - k|$ when $k \rightarrow k_0$, a fact implying that the electronic mass of the uppermost electrons in the Fermi band approaches zero at the absolute zero of temperature. This singularity is usually ignored in computations of such effects as conductivity and diamagnetism (see Chaps. XV and XVI) because of the undetermined influence of correlations. This procedure seems to be justified at room temperature by the fact that results obtained are usually in good agreement with experiment.

The zone structure of lithium² was investigated by Millman, using the cellular approximation. His results show that the effective electron mass is greater than unity, in this case, as we already have seen in Sec. 78. In all other respects, the zone scheme is like that of sodium.

The matrix elements that determine the transition probabilities for optical absorption are zero for perfectly free electrons and are undoubt

¹ This possibility was pointed out by E. Wigner, *Trans. Faraday Soc.*, **34**, 678 (1938).

² J. MILLMAN, *Phys. Rev.*, **47**, 286 (1935).

edly small for the nearly free electrons in the alkali metals. Were this not so, the alkali metals would probably be colored, since the first allowed transition of the type $k \rightarrow k + K$ should occur at 1.5 ev in sodium and at about 2 ev in lithium, according to the zone diagrams for these metals.

b. The Noble Metals Copper, Silver, and Gold.—The monovalent noble metals differ from the alkali metals in that they have newly completed d shells in the atomic configurations. These d levels lie so close to the s levels that configurations such as $3d^{10}4s$ and $3d^95s^2$ in the free atoms are about $1\frac{1}{2}$ ev apart. When the atoms are brought together, the s and d levels split into overlapping bands. Naturally, the d -electron band is narrower than the s - p band because the d electrons are partly screened by the others.

The first investigation of the d band was made by Krutter,¹ who applied the cellular method to copper:

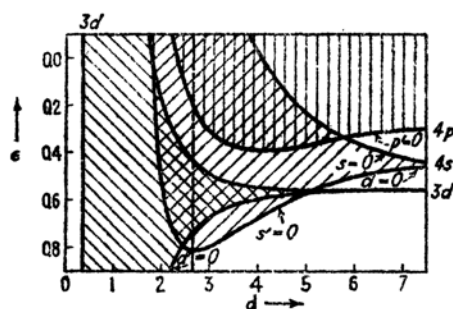


FIG. 4.—Dependence of energy bands of copper on interatomic distance d . It should be observed that the s , p , and d bands overlap appreciably at the actual interatomic distance. The energy scale is in Rydberg units. (After Krutter.)

figure shows that the limit of the filled region is far above the uppermost

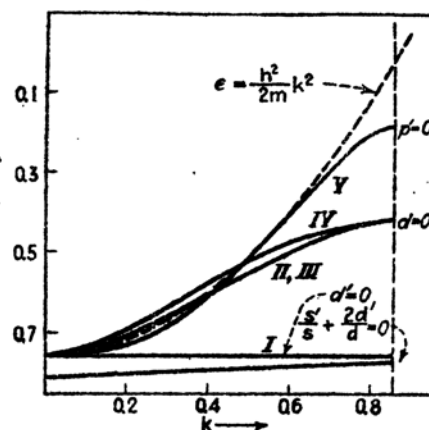


FIG. 3.—The $\epsilon(k)$ curves for the (110) direction of copper. The five branches I to V meet at $k = 0$, giving a five-fold degenerate point. Curve V corresponds to the s - p band in simple metals and is nearly the same as the free-electron curve. Actually all levels are mixtures of s , p , and d states. Curves I to IV and the lowest curve correspond to the d band. The energy scale is in Rydberg units. (After Krutter.)

When computing the s - p bands, he assumed that the field within each cell is that of the free Cu^+ ion; when computing the d band, he used a field obtained from the $3d^94s$ configuration of Hartree's atomic wave functions for copper. In view of the discussion of Chap. X, we may say that these simplifications are reasonable for semi-quantitative work. Figures 3 and 4 show the dependence of the bands upon interatomic distance, and the reduced-zone structure in the (110) direction at the observed lattice distance. The second fig-

¹ H. M. KRUTTER, *Phys. Rev.*, **48**, 664 (1935).

level of the d band, so that this band is completely filled. It is doubtful whether the d band actually is as wide as Krutter's work indicates. Hartree neglected exchange and correlation terms in deriving the field used in computing these wave functions for copper, so that they extend farther from the nucleus and overlap more than they should. The five d -band $\epsilon(\mathbf{k})$ curves meet at $k = 0$ in Krutter's results. His approximate method of applying the cellular scheme is responsible for this degeneracy, for in a more accurate solution the level would split into a twofold and a threefold degenerate level.¹

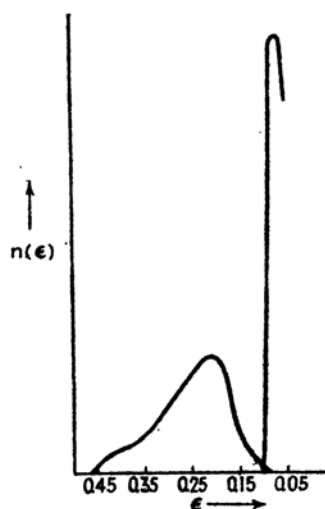


FIG. 5.—The density of energy levels in calcium. The contributions from the first and second zones are indicated separately. The energy scale is in Rydberg units. (After Manning and Krutter.)

metal has a face-centered cubic lattice, and the methods and approximations used in obtaining the energy contours were similar to those used by Krutter for copper.

The alkaline earth metals should be insulators for large interatomic spacing since the atoms have closed-shell configurations in the normal states. The conductivity arises from overlapping of the s , p , and d bands at the observed interatomic distance. Manning and Krutter

The reddish color of copper is attributed to an optically induced electronic transition from the d band to the s - p band. According to Krutter's results, the minimum difference between levels for which a transition is allowed is of the order of 3 ev in the (100) direction, which implies strong absorption in the blue region of the visible spectrum. Since silver is not so strongly colored as copper, we may conclude that the difference between the d and s bands is larger for this metal. The difference presumably decreases again in gold since it is colored.

Tibbs² has carried through similar computations for both copper and silver, including the conduction electrons in more detail.

c. Calcium.—The only extensive calculation on the zone structure of the alkaline earth metals, aside from that for beryllium, which was discussed in Sec. 81, is the work of Manning and Krutter³ for calcium. This

¹ This fact may be derived from a group-theoretical treatment of crystalline wave functions. A cubic crystal cannot have wave functions for $\mathbf{k} = 0$ that are higher than threefold degenerate. Thus, the fivefold degenerate atomic d function splits into a twofold and a threefold degenerate level.

² S. R. TIBBS, *Proc. Camb. Phil. Soc.*, **34**, 89 (1938).

³ M. F. MANNING and H. M. KRUTTER, *Phys. Rev.*, **51**, 761 (1937).

found that this overlapping occurs not in the three principal crystallographic directions, but instead in the (021) direction. Figure 5 shows the density of levels as a function of energy for the first band and part of the second. According to this result, the amount of overlapping actually is very small, a fact which suggests that calcium, like beryllium, is very nearly an insulator.

100. Metals with Irregular Structures.—Mott and Jones¹ have made a detailed investigation of the zone structure of metals such as mercury, white tin, and bismuth that have unusual valencelike structures. In all these metals, it is found that the edge of the filled system of energy levels is very close to a prominent zone boundary, that is, to a boundary that corresponds to a strong X-ray reflecting plane. Since the energy gaps probably are large at a boundary of this type (cf. Sec. 62), there is only a small amount of overlapping of the filled and unfilled zones in these metals. Calcium, which was discussed in part c of the preceding section, is a simple case of this type. Figure 6 shows the prominent zone boundary for the bismuth lattice which contains five electrons per atom.

This observation that irregular metals possess nearly filled bands gives a very satisfactory phenomenological explanation of the fact that their properties lie between those of ideal metals and of valence types. The gaps in an ideal valence crystal are wide enough to keep the occupied and the unoccupied zones apart, for these solids are insulators. On the other hand, the gaps are very narrow in ideal metals. Since the gaps in irregular metals are intermediate between those of these two cases, we may expect that other properties should be intermediate.

The question of why the irregular metals choose those structures which have nearly filled bands rather than others in which the properties are more metallic can be answered accurately only by computing the lattice energy for a nonmetallic and a typically metallic structure, as has been done in the case of hydrogen (cf. Sec. 79). Since such computations have not yet been carried out, we must be satisfied for the present with the chemist's type of answer, namely, that the constituent atoms of irregular metals bear a resemblance to hydrogen in that they prefer to form a structure in which the atoms are coordinated, as in valence crystals.

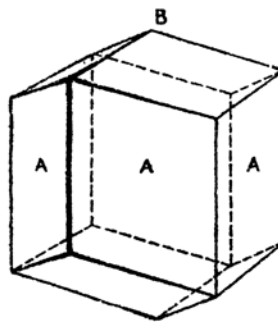


FIG. 6.—The prominent zone boundary for bismuth. This contains $10n$ states so that it is nearly filled. The overlapping of the levels of this zone and the next is believed to occur at points A. (After Jones.)

¹ N. F. MOTT and H. JONES, *Theory of the Properties of Metals and Alloys*, pp. 162 ff.

101. Transition Metals.—Mott¹ and Slater² have used a zone model as the basis for qualitative discussions of the transition metals as a class. Two types of band are employed in this model, namely, a wide low-density valence-electron band, which arises from the s and p atomic states, and a narrow high-density d -electron band (cf. Fig. 7). This scheme of levels occurs in copper, as we have seen in Sec. 99. The essential difference between copper and the transition metals is that the d band is not completely filled in the latter.

An important property of the ferromagnetic transition metals that is readily explained by the zone theory is that their gyromagnetic ratio is nearly equal to 2. A mechanical moment may be induced in a ferromagnetic substance by magnetizing it, and the ratio of the magnetic moment, expressed in units of the Bohr magneton, to the angular momentum,

expressed in units of \hbar , is called the gyromagnetic ratio.³ If the magnetic moment arises purely from orbital motion, the ratio should be unity (cf. Chap. V); if it arises from electronic spin, the ratio should be 2; and if it arises from a combination of spin and orbital motion, it should lie between zero and 2. The fact that the value usually is nearly 2 (for example, the value for iron is 1.93) indicates that most of the orbital angular momentum that the d electrons possess in the

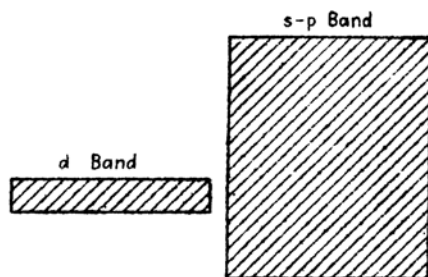


FIG. 7.—Schematic diagram of the relative positions and widths of the d band and s - p band (cf. Fig. 4 for copper). The narrow d band has room for ten electrons per atom, whereas the s - p band has room for only two.

free atom is "quenched" in passing from the gas to the solid and that principally the spin magnetic moment remains. The orbital angular momentum is negligible in the band scheme,⁴ since the zones are symmetrically filled in such a way that the electrons may be paired in groups which move in opposite directions with equal velocities. Accurate measurements such as the one cited above for iron show that the orbital magnetic moment is not entirely quenched, a fact indicating that the band approximation is not entirely accurate.

The cellular method has been applied in detail to only one transition metal, namely, tungsten, which is discussed later, in part *b*. Slater,

¹ N. F. MOTT, *Proc. Phys. Soc.*, **47**, 571 (1935).

² J. C. SLATER, *Phys. Rev.*, **49**, 537 (1936); *Jour. Applied Phys.*, **8**, 385 (1937).

³ See E. C. STONER, *Magnetism and Matter* (Methuen & Company, Ltd., London, 1934); see also S. J. BARNETT, *Rev. Modern Phys.*, **7**, 129 (1935).

⁴ This point has been carefully investigated by H. Brooks (paper to appear in *Phys. Rev.*).

however, has used the level system computed by Krutter for copper to discuss some of the details of the iron-group metal alloys, assuming that the relative positions of the d and the s - p bands do not change very much throughout the iron group (cf. Sec. 103). The density of levels in the s - p and d bands of copper is shown in Fig. 8; the vertical lines indicate the extent to which these levels would be filled in transition metals with different numbers of electrons per atom.

a. The Iron-group Transition Metals.—We shall now discuss several properties of the iron-group transition metals on the basis of the band picture.

1. *Cohesion.*—One of the striking properties of the cohesive energies of the transition metals immediately preceding copper, silver, and gold is

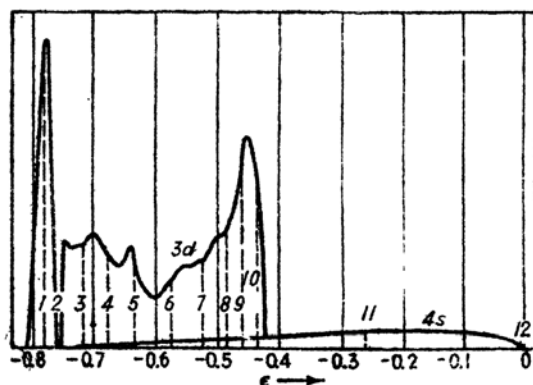


FIG. 8.—The density of levels in the iron-group series. The vertical lines designate the limit to which the levels are filled in the elements having the number of electrons corresponding to the integers given. Thus the d band is completely filled in copper (11 electrons) and is not quite filled in nickel (10 electrons). This figure is based on Krutter's work on copper. The abscissa is expressed in Rydberg units. (After Slater.)

the fact that they usually are larger than the cohesive energies of the monovalent metals. This fact is illustrated by the following sequences:

Ni	85 kg cal/mol	Cu	81 kg cal/mol
Pd	110 kg cal/mol	Ag	68 kg cal/mol
Pt	127 kg cal/mol	Au	92 kg cal/mol

Mott has proposed the following qualitative interpretation of this fact. The electrons in the s - p band are principally responsible for the cohesion of all these metals, since the d shells are nearly filled. Fuchs has estimated the d -shell interaction in copper (cf. Sec. 80) and has found it to be of the order of 0.5 ev per atom. If it is assumed that the electronic levels are very nearly the same in the transition-group metals and in the simple metals that immediately follow them (for example, in the sequence from iron to copper), it should be expected that the differences in binding properties arise from differences in the way in which the levels are filled.

Let us consider iron, cobalt, nickel, and copper. In the first three cases, the d band is not completely filled, and the s - p band presumably is filled to the same height as the d band. Since the density of d levels is very high, this fact means that the s - p band is filled to very nearly the same point in each case, if it is assumed that the relative positions of the two bands remain fixed. On the other hand, the d band of copper is completely filled, so that the additional electrons fill the s - p band to a point far above the top of the d band. Mott points out that the bottom of the s - p band, the top of the filled region, and the mean energy per atom in copper are related in the manner shown in Fig. 9, in which curve I is the bottom of the s - p band, curve II is the top of the filled region, curve III is the mean energy, and the zero of energy is referred to the state in which all atoms are infinitely separated. Since the electrons between curves II and

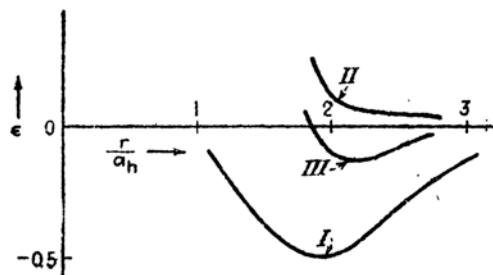


FIG. 9.—Schematic representation of the electron energies in copper. Curve I is the $\epsilon_0(r_s)$ curve; curve II is the top of the filled band; and curve III is the mean energy. Mott points out that curve III would be lowered if there were fewer electrons in the s - p band.

III have more energy than the average, it might be expected that curve III would be lowered if the electrons in the upper part of the filled region were either removed or placed in lower levels. These electrons are essentially removed in the transition metals, according to the zone model. Hence, it may be expected that, in these cases, curve II and curve III are lower than in copper. It is only fair to say that this argument can be used only in a qualitative way, since a shift in the filling of one-electron levels has associated with it changes in exchange and correlation energies that cannot be included in a simple energy diagram.

It has been suggested by other workers¹ that at least part of the binding of the transition metals is related to a lowering of the center of gravity of the occupied d levels in passing from the free atom to the solid. Recent work on tungsten, which is discussed below, indicates that this effect is probably the largest source of cohesive energy in the platinum series of transition metals. Whether or not it is important in the iron group remains to be seen.

¹F. SEITZ and R. P. JOHNSON, *Jour. Applied Phys.*, **8**, 84, 186, 246 (1937).

In addition, Pauling,¹ reasoning on the basis of the empirical knowledge of the properties of transition metals, has suggested that some of the d electron wave functions combine with the s - p functions to form a scheme of levels in which there is a larger number of valence electrons per atom than the one s - p function per atom suggested by Krutter's band scheme for copper. In principle, Pauling's suggestion is equivalent to assuming that there are two types of d -electron band in the transition metals, namely, a wide band (see Fig. 10), which is similar to the s - p band but contains 2.6 of the 5 d functions per atom of given spin, and a narrow band, which contains the remaining 2.4 d functions. The first class of function

(A type), along with the s - p functions, is responsible for the large cohesive energy of the transition metals since its existence implies an increase in the number of binding electrons in the metal, whereas the second class of function (B type) is responsible for the magnetic properties in a manner that will be described under 2. It was

pointed out in Sec. 99, in connection with Krutter's work on copper, that the five d zones should break into two separate systems containing two and three zones, respectively (see footnote 1, page 424), and that Krutter's approximation does not give this splitting.

It is possible that Pauling's scheme of levels would be obtained from Krutter's if the band approximation were applied with a higher degree of accuracy. We shall see in Sec. 104, however, that the experimentally determined levels do not seem to agree with Pauling's assumptions.

2. *The filling of levels in the ferromagnetic elements.*—This topic was previously introduced in Sec. 27, Chap. IV, in which we used the band theory to explain the low-temperature specific heats of transition metals. It was postulated there that in the ferromagnetic elements the half of the d band associated with one kind of electron spin is completely filled and that the saturation magnetic intensity, expressed in Bohr magnetons per atom, is equal to the number Δn_d of unoccupied levels per atom in the other half of the d band. The number of electrons per atom n_s in the s - p band may be computed from this hypothesis. If m is the total number

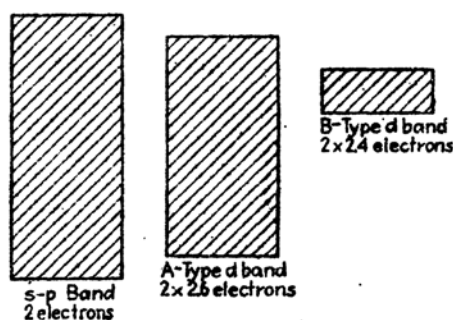


FIG. 10.—The energy bands in the iron-group metals according to Pauling. The s - p functions combine with some of the d functions to form two broad bands, namely, the left-hand band, which is designated as the s - p band, but which contains a mixture of d functions, and the A type of d band, which has room for 1.6 electrons per atom of given spin. The B type of d -electron band is narrow and has room for 2.4 electrons per atom of given spin.

¹I. PAULING, *Phys. Rev.*, 54, 899 (1938).

of d and s - p electrons in the atom, it follows that Δn_d , n_v , and m must satisfy the relation

$$10 - \Delta n_d + n_v = m$$

or

$$n_v = m + \Delta n_d - 10.$$

We find, using the measured saturation moments, that n_v is, respectively, 0.2, 0.7, and 0.6 for iron, cobalt, and nickel. The first value undoubtedly is too small if the cohesive energy of iron is to be explained in terms of the energy of s - p electrons, whereas the other two are reasonable. This low value suggests either that half the d band is not filled in this case or that the d band is lower in iron than in the metals following it and the cohesive energy is related to this lowering. The second possibility is not entirely unreasonable, for, as was mentioned above, the cohesive energy of tungsten (see part *b* of this section) seems to be related entirely to the behavior of the d band. The first possibility is readily explained on the basis of Pauling's suggestion; for in iron the B type band of Fig. 10 would be completely drained of the 2.4 electrons per atom having one type of spin and 0.2 electron per atom having the other type of spin would be removed, whereas in cobalt and nickel only a fraction of the electrons in half this band would be removed. We are not able to decide between the two alternatives on the basis of the present knowledge of energy levels, however, and in subsequent discussions we shall arbitrarily assume that the first is correct.

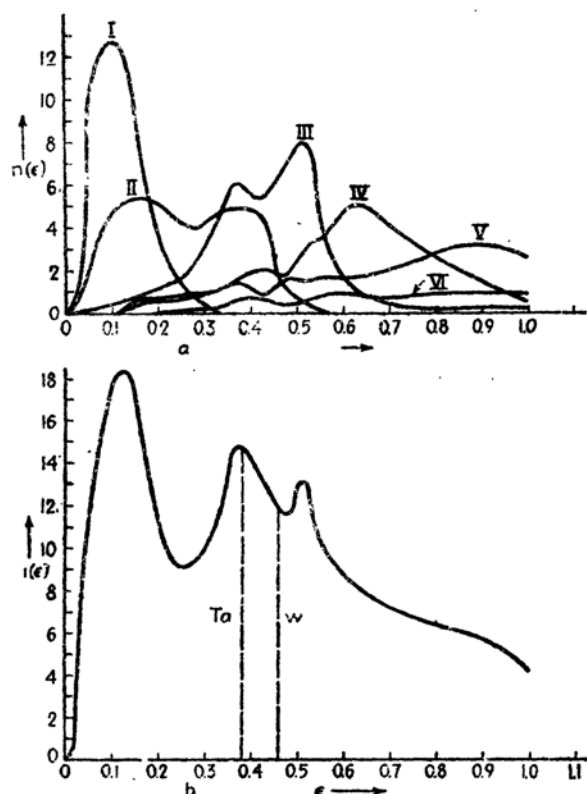
The Heitler-London scheme has been used with considerable success in discussing the spin-aligning forces of ferromagnetism, as will be seen in Chap. XVI. Since this approximation does not provide a satisfactory simple explanation of the low-temperature electronic heat of the transition metals, it cannot be used in place of the band theory for all purposes.

3. *The paramagnetic transition metals.*—We have already seen in Sec. 29 that the band theory cannot explain even semiquantitatively both the specific heat and the magnetic susceptibility of the paramagnetic transition metals. This failure lends additional support to the statements made above concerning the limitations of the band approximation when applied to d -shell electrons.

b. Tungsten.—The band structure of metallic tungsten, which has a body-centered lattice and whose atoms possess six valence electrons and a newly filled f shell outside a rare gas configuration, has been investigated by Manning and Chodorow.¹ These workers assumed that the f shell is unaffected by solid binding and obtained wave functions and energy levels for the remaining six electrons per atom. In first approxi-

¹ M. F. MANNING and M. I. CHODOROW, *Phys. Rev.*, **56**, 787 (1939).

mation, an effective field for these electrons was obtained from a charge distribution derived by renormalizing the parts of the free-atom wave functions lying within the atomic sphere of the lattice. Valence-electron wave functions and $\epsilon(\mathbf{k})$ contours were then computed with the use of this field. In second approximation, the charge density obtained from the results of the first approximation was used to compute a new field. This second approximation was very nearly self-consistent.



11.— $n(\epsilon)$ curves for the five d zones and one s - p zone of tungsten. The energy scale is in Rydberg units. (After Manning and Chodorow.)

The $n(\epsilon)$ curves for the six lowest zones are shown in Fig. 11a. The set labeled with roman numerals I to V correspond to the five d bands, and curve VI corresponds to the s - p band. The total $n(\epsilon)$ curve for all six bands is given in Fig. 11b. The limit of the filled regions of tungsten and of the neighboring element tantalum are marked by vertical lines. According to these results, the number of electrons in the s - p band is of the order of magnitude 0.1 electron per atom in both these metals. Manning and Chodorow estimate that the center of the filled region is

about 8 ev below the mean of the occupied levels of the free atom. These results imply that the large cohesive energy of tungsten, namely, 210 kg cal/mol, arises from the fact that practically all of the six valence electrons may occupy the low-energy portion of the *d* band. In addition, they imply that the contribution to cohesion from the *s-p* electrons, which are responsible for most of the cohesion in simple metals, is negligible in this case.

Using the computed $n(\epsilon)$ curves, Manning and Chodorow estimated the electronic heat of tungsten and tantalum by the use of the equations derived in Chap. IV. The observed and calculated values are given in Table LXXII.

TABLE LXXII.—A COMPARISON OF OBSERVED AND CALCULATED ELECTRONIC HEATS OF TUNGSTEN AND TANTALUM
(In units of 10^{-4} cal/deg-mol)

	Theoretical	Experimental	
		Low temperature	High temperature
W	4.8 <i>T</i>	5.1 <i>T</i>
Ta	6.2 <i>T</i>	27 <i>T</i>	7 <i>T</i>

The low-temperature value for tantalum was obtained indirectly from conductivity measurements near absolute zero, and the high-temperature values were obtained after subtracting the $3R$ lattice vibrational heat and questionable $C_p - C_v$ corrections from the observed molar heats. A discussion of this work may be found in the original paper by Manning and Chodorow. It is difficult to say whether the discrepancy between low-temperature and high-temperature values implies error in the simple theory of electronic heats developed in Chap. IV or in the treatment of experimental results. In any case, the agreement between the theoretical results and the high-temperature values is excellent.

102. Simple Substitutional Alloys.—In the experimental survey of Chap. I, it was seen that the Hume-Rothery electron-atom ratio rule correlates the solid-phase portions of the phase diagrams of different substitutional alloy systems. This rule states that a given phase occurs for a fixed electron-atom ratio in a number of different alloy systems. As a result of an extensive investigation, Jones¹ has found that the edge of the filled region of levels lies close to a prominent zone boundary when the Hume-Rothery rule is satisfied. This observation allows us

¹ H. JONES, *Proc. Roy. Soc.*, **144**, 225 (1934); **147**, 396 (1934). MOTT and JONES, *op. cit.*, Chap. V.

to replace the Hume-Rothery rule by the statement that the stable alloy phases have nearly filled systems of zones.

As we mentioned in Sec. 100, it is not evident why a nearly filled zone system should be more stable than another, and this fact has not yet received a completely satisfactory explanation. Figure 12 shows

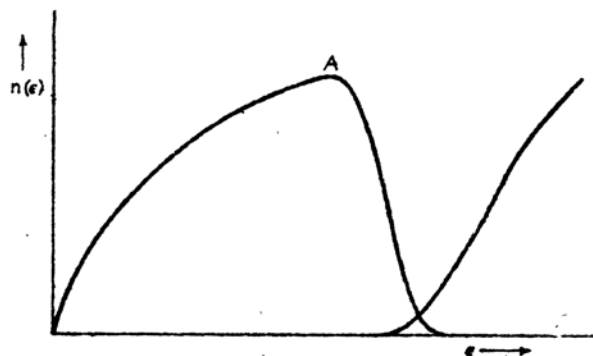


FIG. 12.—Schematic behavior of density of levels near a prominent zone boundary of a metal (cf. Fig. 5 for calcium and Fig. 8 of Chap. X for beryllium).

the behavior of the density of states per unit energy range near a zone boundary at which the gaps are large. The density of levels in the lower zone increases at first as the zone boundary is approached because the $\epsilon(k)$ curves bend over. After this rise, the density falls and approaches the axis sharply. It should rise sharply again in the higher zone in the manner illustrated. We may conclude that the two zones overlap in the substitutional alloys from the fact that these alloys are metallic conductors. Jones assumes that the maximum A of the lower zone occurs at that value of energy for which the contours in wave-number space just touch the zone boundary. This assumption has been justified by a detailed treatment of the $\epsilon(k)$ curves for a number of zones with the use of simplified models (cf. Sec. 65). He then postulates that the stability of a phase increases as the levels are filled to the point A and then decreases rapidly beyond this point, because the average energy of the additional energy is much larger than the mean energy of all electrons. If this assumption is true, the electron-atom ratio associated with A should be the value for which the phase is most stable.

Four phases are commonly met in simple substitutional alloy systems. The α phase is face-centered cubic, the β phase is body-centered cubic,

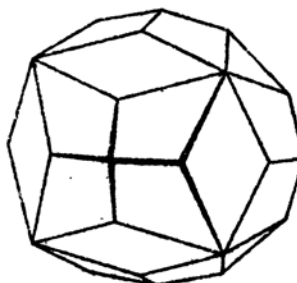


FIG. 13.—The prominent zone boundary for the γ brass phase. This zone contains room for 90 electrons per unit cube of the lattice. (After Jones.)

the γ phase has a more complex cubic structure, and the η phase is close-packed hexagonal. The first zones for the α , β , and η phases were shown in Chap. VIII. The prominent zone for the γ phase is shown in Fig. 13. This zone contains 90 states for a cubic cell of 52 atoms, or 1,731 states per atom. The Hume-Rothery electron-atom ratio is $\frac{2}{3}$, or about 1.615. Jones has computed the ratio corresponding to the point A, Fig. 12, for each of these four phases, assuming that the energy contours are spheres. His values are given in Table LXXIII along with Hume-Rothery's fractional estimates. The two values differ slightly but agree equally well with the experimental values.

TABLE LXXIII

Phase	Hume-Rothery's fractional value	Jones' value
α	1.362
β	$\frac{2}{3} = 1.5$	1.480
γ	$\frac{2}{3} = 1.615$	1.538
η	$\frac{1}{4} = 1.75$	-1.7

108. Alloys Involving Transition Metals.—The properties of transition metals explained most readily by the band theory are (1) the quenching of ferromagnetism by the addition of nontransition metals that form solid solutions and (2) the dependence of the saturation magnetic moment of ferromagnetic alloys on atomic composition. We shall discuss these two topics together.

If similar phases of the iron-group elements and copper and zinc have practically the same zone structure, we may expect that the magnetic properties of their alloys depend principally upon the extent to which zones are filled, that is, upon the electron-atom ratio. We shall use the following two principles in correlating the saturation magnetic moments:

1. The number of valence electrons per atom in the s - p band of all ferromagnetic metals is about 0.7. If, in addition, we were to accept Pauling's postulate, we should also assume that the A type d band of Fig. 10, which contains 2.6 electrons per atom, is filled or nearly filled and that the electrons are removed from or added to the B type band.

2. The saturation magnetization, expressed in Bohr units per atom, is approximately equal to the number of holes per atom in the d band. The word "approximately" is inserted because the saturation magnetization seems to be less in iron, as we have seen in Section 101. We shall try to give additional insight into this point in the following paragraphs.

Before presenting a general survey of results, we shall consider two typical cases. Suppose that some of the nickel atoms in a specimen of

nickel are replaced by copper atoms. Since copper has one electron per atom more than nickel, we should expect that each copper atom which is added has the same effect as if one electron were added to the bands of pure nickel. These additional electrons enter the d band and should decrease the number of holes at the rate of 1 per copper atom. Since there is 0.6 hole per atom in pure nickel, we should expect the saturation magnetization to decrease linearly with the concentration of copper and to be zero when the atom fraction of copper is 0.6. Sadron's measurements show this to be the case (see Fig. 54, Chap. I). We should expect zinc atoms to have twice the effect of copper atoms since zinc has two valence electrons instead of one. This is also found to be true.

Let us consider next the effect of alloying nickel and cobalt. Nickel has 0.6 hole per atom in the d band and cobalt has approximately 1.7. According to the band model, the number of holes per atom in the alloy that contains an atomic fraction of x nickel atoms and $(1 - x)$ cobalt atoms should be

$$n_x = 0.6x + 1.7(1 - x), \quad (1)$$

so that the saturation magnetic moment should be n_x Bohr magnetons per atom. This rule is closely obeyed in the nickel-cobalt system, as we shall see below.

Figure 14 shows¹ the relation between the saturation magnetization per atom and the number of holes per atom for a number of substitutional alloys of the iron-group elements. The number of holes per atom in the d band is computed by the use of equations of the type (1) on the assumption that there are 0.7 $s-p$ electron per atom in all transition metals except nickel, which has 0.6. Nontransition elements are assumed to have a negative value, corresponding to 0.7 minus their valence (that is, -0.3 for copper and -1.3 for zinc). If α_a and α_b are the number of holes per atom in the pure metals A and B , the number per atom in the alloy that contains a fraction f_a of A and f_b of B is

$$n_h = \alpha_a f_a + \alpha_b f_b, \quad (2)$$

analogous to (1). The abscissa of Fig. 14 is the value of n_h computed from (2); the ordinate is the saturation magnetic moment per atom σ .

¹ J. C. SLATER, *Jour. Applied Phys.*, **8**, 385 (1937).

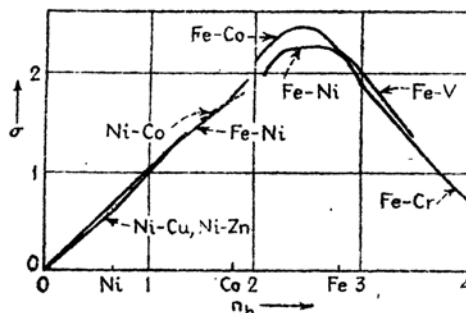


FIG. 14.—Relation between the saturation magnetization, expressed in magnetons per atom, and the number of holes per atom. (After Slater.)

A straight line corresponds to strict proportionality between these two quantities. It should be observed that this relationship is closely obeyed until the point at which n_h is about 2.2. The experimental curve then bends over smoothly and approaches the axis along what appears to be a straight line. The small value of the saturation magnetization of iron that was mentioned in Sec. 101 is in accordance with this bending. Pauling's assumption of the existence of *A* type and *B* type *d* bands is primarily based upon this fact, for in his picture the *B* type band is completely drained of electrons of one type of spin at the composition corresponding to the peak of the curves of Fig. 14, that is, when n_h is about 2.4. As one passes farther to the right, toward iron, manganese,

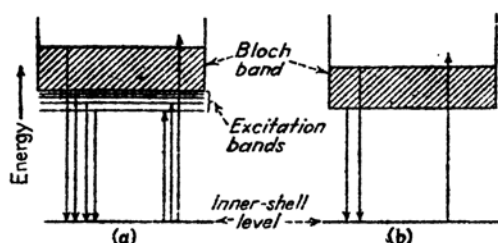


FIG. 15.—In case *a* there are excitation bands for the inner-shell electrons, and both continuous bands and discrete lines may be expected in emission. In case *b* the excitation bands have been absorbed into the continuum. The shaded region indicates the filled part of the bands. It should be observed that emission occurs from occupied valence levels, absorption to unoccupied ones.

and chromium, the remaining electrons are removed from the *B* type band and σ decreases linearly.

Mott and Jones¹ have used essentially the same principles to determine the number of holes in several paramagnetic transition metals. For example, the paramagnetism of palladium decreases when it is alloyed with gold. Since the paramagnetism vanishes when a fraction* of 0.55 palladium atoms has been replaced by gold atoms, these workers conclude that palladium contains 0.55 hole per atom.

104. Level Densities from Soft X-ray Emission Spectra.—Experimental values of the level-density curves of the valence electrons in metals may be obtained from the soft X-ray emission and absorption spectra of metals.² These curves have particular value in deciding whether the excitation-band picture should be applied to metals as well as to insulators or whether the band approximation is accurate for qualitative work.

¹ MOTT and JONES, *op. cit.*, pp. 199–200.

² See the survey article by H. W. B. SKINNER, *Reports on Progress in Physics* V (1938), (Cambridge University Press, 1939). This possibility was first pointed out by W. V. HOUSTON, *Phys. Rev.*, 38, 1797 (1931).

Let us suppose that an electron is missing from an inner shell of an ion in a metal. We may describe these inner-shell levels in the Heitler-London approximation and may designate the wave function of the missing electron, which is localized about a single ion, by $\psi_i(r)$. If the Heitler-London or the excitation-band picture were valid for the excited states of this electron, there would be a set of discrete levels or a set of nonconducting excitation states beneath the ionization limit; the latter corresponds to the beginning of the Bloch levels (cf. Fig. 15). Thus, the emission and absorption spectra would consist of a continuum corresponding to transitions between the Bloch band and the lowest level and of discrete lines corresponding to transitions between the excitation levels and the lowest levels. On the other hand, if there are

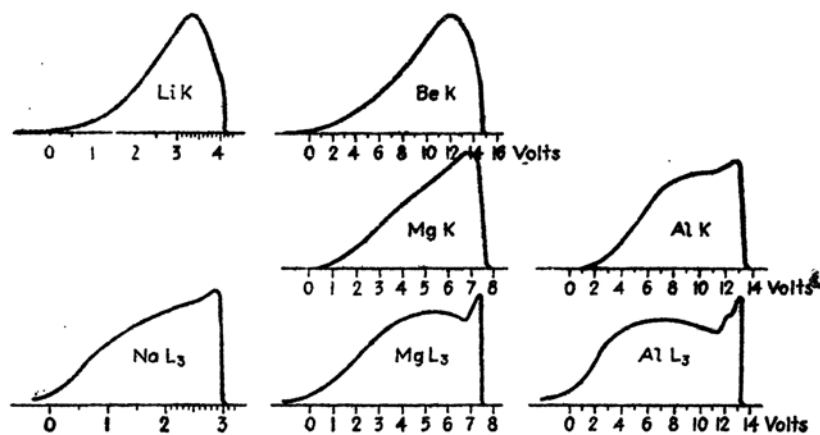


FIG. 16.—The soft X-ray emission spectra of several metals. (After Skinner.)

no excitation bands¹ because the ionization band has broadened enough to absorb them, only the continuum should be present. In the transition from one of these cases to the other, we may expect the intensity of the discrete lines to predominate over that of the continuum at first, then decrease, and finally disappear.

The observed soft X-ray emission spectra of lithium, sodium, beryllium, magnesium, and aluminum, as determined by O'Bryan and Skinner² and Farineau,³ are given in Fig. 16. The lithium and beryllium bands arise from transitions to the $1s$ level (K band), and the two magnesium

¹ The absence of excitation levels would imply that the free electrons so completely screen the hole in the ion core that there is not enough potential for a discrete level. The quantitative description of this effect would require a very accurate treatment of the many-electron problem.

² H. M. O'BRYAN and H. W. B. SKINNER, *Phys. Rev.*, **45**, 370 (1934).

³ J. FARINEAU, *Compt. rend.*, **203**, 540 (1936), **204**, 1108 (1937), **204**, 1242 (1937), **205**, 365 (1937); *Nature*, **140**, 508 (1937). See footnote 1, p. 440, also.

and aluminum bands arise, respectively, from transitions to the 1s level and 2p level (L_{III} band). In contrast with the absorption spectra of the alkali halides (cf. Sec. 95), these spectra show no indication of strong discrete lines, which implies that the excitation levels are not prominent. In order to determine the extent to which the low-energy tails of these bands are vestiges of the transitions from excitation states, we must examine the theory of emission more closely.

We shall assume¹ that the periodic wave functions may be expressed in the form

$$\psi_{\mathbf{k}} = \chi_0(\mathbf{r})e^{2\pi i\mathbf{k}\cdot\mathbf{r}}$$

where χ_0 is an s function in the vicinity of the nucleus. The intensity of the line emitted in the jump from $\psi_{\mathbf{k}}$ to ψ_f then is proportional to

$$|\int \psi_{\mathbf{k}}^* \text{grad } \psi_f d\tau|^2 \quad (1)$$

(cf. Sec. 43). For small values of \mathbf{k} or \mathbf{r} , we may expand the exponent in powers of $\mathbf{k} \cdot \mathbf{r}$ and keep only the first two terms. Thus, (1) becomes

$$|\int \chi_0(1 - 2\pi i\mathbf{k} \cdot \mathbf{r}) \text{grad } \psi_f d\tau|^2. \quad (2)$$

If ψ_f is a p function, the integral of the term in (2) that involves \mathbf{k} vanishes, leaving

$$|\int \chi_0 \text{grad } \psi_f d\tau|^2, \quad (3)$$

which is independent of \mathbf{k} . This result should be valid for fairly large values of \mathbf{k} since ψ_f is localized exceedingly close to the nucleus. Thus, the intensity of the emission band as a function of energy should depend only upon the density of levels and, as a result, should vary as $\sqrt{\epsilon}$ near the low-energy side. On the other hand, if ψ_f is an s function, the term corresponding to (3) vanishes and the remaining term can be reduced to

$$\mathbf{k}^2 \left| \int \chi_0 x \frac{\partial \psi_f}{\partial x} \right|^2, \quad (4)$$

which varies as \mathbf{k}^2 or as ϵ near the low-energy side of the band. Since the density of levels varies as $\sqrt{\epsilon}$ in this energy region, it follows that the intensity of the band should vary as $\epsilon^{\frac{3}{2}}$.

Now, the L_{III} emission curves for sodium, aluminum, and magnesium correspond to the first of these two cases. It is clear from Fig. 16 that the intensity starts out much more slowly than $\sqrt{\epsilon}$. A comparison of

¹ HOUSTON, *op. cit.*; H. JONES, N. F. MOTT and H. W. B. SKINNER, *Phys. Rev.*, **45**, 379 (1934).

the actual curve and the $\sqrt{\epsilon}$ curve that would fit the sodium curve¹ most closely is given in Fig. 17a. This suggests that there is a residue of the excitation bands; however, it is also possible, although not probable, that the measured tail is due to a background emission of impurity atoms. Since the *K* emission curves of Fig. 16 correspond to the second of the two cases, they should rise as $\epsilon^{\frac{1}{2}}$. Figure 17b gives a comparison of an $\epsilon^{\frac{1}{2}}$ curve and the observed curve for lithium. The high-energy cutoff of the theoretical curve is chosen so that the width is the same as that computed in Sec. 78. The discrepancy on the low-energy side may have the explanation suggested in Fig. 17a for sodium. The discrepancy on the high-energy side indicates that the actual density of levels does not

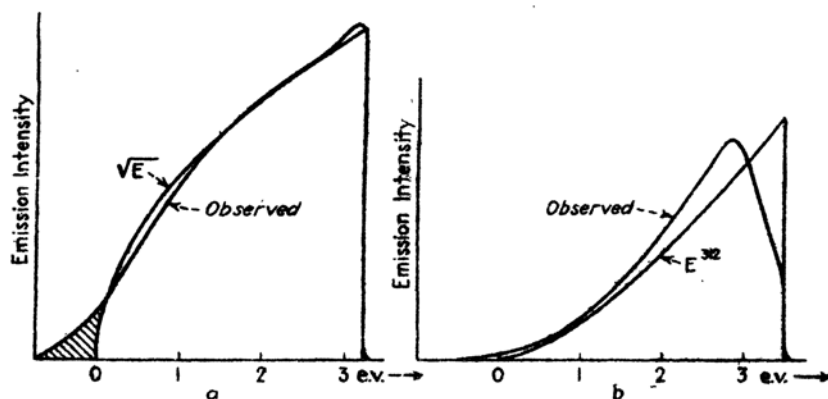


FIG. 17.—(a) Comparison of the actual emission band of sodium and that expected from the simple band theory. It is suggested that the shaded region represents the contribution from vestiges of the excitation bands. The energy scale is in electron volts. (b) Comparison of the actual emission band of lithium and that predicted by the band theory.

vary as $\sqrt{\epsilon}$ in lithium. The premature peak in the curve for lithium has not received a satisfactory explanation. It is possible that it is related to a rapid variation in exchange and correlation energy at the top of the filled region; however, if this is the case, it is not easy to see why sodium does not have a similar peak. If the difference between the curves for lithium and sodium is real, we may expect the electronic specific heat of sodium to be more normal than that of lithium, for its level density is more nearly like that for perfectly free electrons.

It is interesting to note that the beryllium curve of Fig. 16 behaves as though the levels of a single zone were almost completely occupied, a fact indicating that this metal is very nearly an insulator. In mag-

¹ In these comparisons, the effect of both exchange and correlation on the calculated band widths is neglected because the second, which tends to compensate for the broadening effect of the first (see footnote 1, p. 422) is not precisely known.

nesium and aluminum, however, the curves appear as though two zones overlapped extensively.

TABLE LXXIV.—OBSERVED AND CALCULATED WIDTHS OF THE SOFT X-RAY EMISSION BANDS (AFTER SKINNER)

(The theoretical values are those computed in Chap. X. Values in parentheses are free-electron values)

	Observed, ev	Calculated, ev
Li	4.1 ± 0.3	3.4
Na	3.4 ± 0.2	3.2
Be	14.8 ± 0.5	(13.8)
Mg	7.6 ± 0.3	(7.2)
Al	13.2 ± 0.5	(12.0)

The widths of a number of the emission bands determined by O'Bryan and Skinner, and others, are given in Table LXXIV and are compared with the theoretical values in cases in which the latter have been computed. The values

$$\frac{h^2}{2m} \left(\frac{3n_0}{8\pi} \right)^{\frac{1}{3}}$$

for perfectly free electrons are given in parentheses in the other cases.

The emission bands of metals containing filled or partly filled *d* bands have been investigated by a number of workers,¹ among whom are Bearden, Shaw, Beeman, Friedman, Saur, Gwinner, and Farineau. Farineau's curves for nickel, copper, and zinc are shown in Fig. 18. Since the density of levels in the *d* band presumably is much higher than that in the *s-p* band, we may conclude that practically all of this structure arises from the *d* band. One of the important features of these curves is the fact that there is a single peak in the cases of copper and nickel and not two, as we might expect from Krutter's work. Krutter's curve is

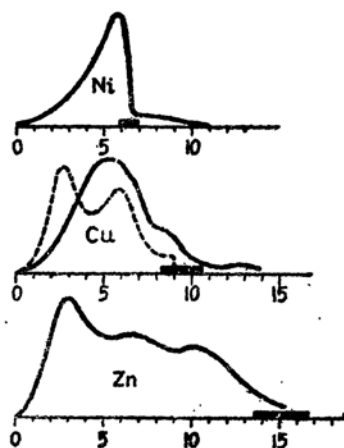


FIG. 18.—The emission curves for nickel, copper, and zinc. Presumably this emission arises mainly from the *d* band. The difference between copper and nickel is mainly due to the fact that copper has one more valence electron than nickel. Zinc, however, has an entirely different structure so that the band structure should be different.

shown in the diagram for copper. If this peak were to be associated with

¹ J. A. BEARDEN and C. H. SHAW, *Phys. Rev.*, **48**, 18 (1935); W. W. BEEMAN and H. FRIEDMAN, *Phys. Rev.*, **56**, 392 (1939); E. SAUR, *Z. Physik*, **103**, 421 (1936); E. GWINNER, *Z. Physik*, **108**, 523 (1938).

Pauling's *B* type of *d* band (see Fig. 10), we should expect that it would occur nearer the high-energy cutoff of the curves, at least in the case of nickel. Assuming that these density curves are trustworthy, we must conclude both that the *d* bands shift to some extent in going from copper to nickel and that the one-electron approximations discussed in previous sections are not very accurate when applied to *d* electrons. The curves for copper and zinc are considerably different; however, this fact is not surprising, for the crystalline symmetry is different in the two cases.

B. IONIC CRYSTALS

105. Plan of Treatment.—In the preceding chapter, we saw: (a) that the lowest state of ideal ionic crystals may be treated approximately by either the Heitler-London or the band scheme; (b) that the lower, non-conducting, excited states may be treated by the method of excitation waves, at least in the case of the alkali halides; and (c) that the higher excited states may be treated by the band scheme. In the following sections, we shall apply these approximations to several alkali halide crystals and alkaline earth oxide and sulfide crystals the experimental properties of which have been investigated with some degree of completeness. The first two sections apply to crystals having ideal, undistorted lattice structures, and the following section applies to ideal crystals having lattice distortions of a type that will be described in more detail in that section.

106. The Alkali Halides.—Zone-structure calculations have been made for two alkali halides, namely, lithium fluoride and sodium chloride. In addition, computations have been made for lithium hydride, which resembles the alkali halides closely since negative hydrogen ions behave like ions of a halogen. In an atomic picture, the eight valence electrons per unit cell of the alkali halides completely occupy the outer *s* and *p* shells of the negative ions. In the band scheme, the same electrons occupy four zones, one of which connects adiabatically with the ionic *s* level and the three others of which connect with the ionic *p* level. Figure 19 illustrates the manner in which the ionic levels broaden into the bands of the zone theory as the ions are brought together. The levels of the negative ions are depressed and those of the positive ions are raised because of the Madelung field. In addition, the levels break into bands when the ions begin to overlap. At the observed lattice distance, the *s* and *p* bands are separated from one another and from the higher unfilled band which connects with the lowest level of the metal ion. The ionic levels do not split into bands in the Heitler-London approximation but remain discrete, roughly following the center of gravity of the bands (cf. the discussion in Sec. 64 concerning the connec-

tion between the Heitler-London and Bloch energy parameters). If the squares of the $2s$ wave function of atomic lithium and of the radial part of the p wave function of ionic fluorine are plotted in such a way that each distribution is centered about points which are separated by the distance between the lithium and fluorine ions in the crystal, it is found that the peaks of the distributions overlap and that the center of gravity of the charge of an undistorted lithium atom would lie in the shells of the surrounding halogen ions (cf. Fig. 1, Chap. II, for lithium hydride).

In treating sodium chloride, Shockley¹ chose an effective chlorine charge distribution by normalizing Hartree's chlorine ion wave functions

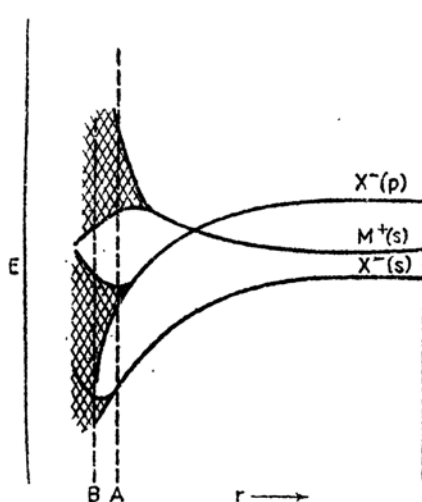


FIG. 19. --Schematic representation of the manner in which the ionic levels of the constituents of ionic crystals break into bands in the band approximation. At A the s and p bands of the negative ion are separate, whereas at B they overlap.

within a sphere the volume of which is equal to that of the unit cell of sodium chloride. He assumed that only eight electrons are in this cell at any one time, because of correlation effects, and computed the effective field for a given electron in the sphere by taking the charge of the remaining seven into account. The effective field for an electron near a sodium ion was taken as the ion-core field used in the computations on metallic sodium. To these effective ion fields, Shockley added the Madelung field of the surrounding ions. In addition, he subtracted from the sodium ion-core field the field arising from a single electron that is spread uniformly over the six surrounding halogen ions. The missing electron is assumed to be at the

sodium ion. Boundary conditions were applied to the wave functions in three different ways: (1) by neglecting the sodium wave functions and treating the lattice as though it were composed of chlorine ions, (2) by satisfying only sodium-chlorine boundary conditions, and (3) by satisfying both sodium-chlorine and chlorine-chlorine boundary conditions. The equations employed in the first case were those derived by Krutter for copper, and Shockley derived similar equations for the two other cases. The $\epsilon(k)$ curves that were obtained from these equations are shown in Figs. 20 *a, b*.

¹ W. SHOCKLEY, *Phys. Rev.*, **50**, 754 (1936).

More recently, Tibbs¹ has investigated the conduction band of sodium chloride by a similar method and has shown that the effective mass of the conduction electrons is close to unity.

Attempts to construct self-consistent fields for the lattice were made in the treatments of lithium fluoride and hydride.² This task is much more difficult in the case of diatomic crystals than in the case of monatomic ones, because the total charge in each polyhedron of the unit cell, as well as the relative distribution within a given polyhedron, must be the same in the initial and final solutions. In these ionic crystals, the lattice was divided into cubes of equal volume centered about each of the ions, and the cubes were replaced by equivalent spheres. The Li^+

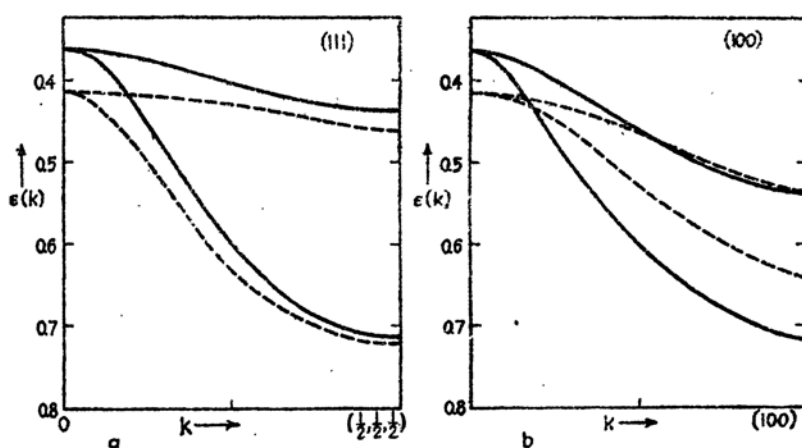


FIG. 20.—(a) $\epsilon(k)$ curve for the (111) direction of sodium chloride. Only the p -band curves are given. The full curve corresponds to the results obtained by neglecting the sodium wave functions; the dotted lines correspond to the most accurate procedure described in the text. (b) Same for the (100) direction. (After Shockley.)

ion-core field was taken from the work on metallic lithium, and the field of the $(1s)^2$ ion core of F^- was taken from Hartree's work. The remaining eight electrons per unit cell were treated by a self-consistent method. The charge distribution of the valence electrons was not determined by computing wave functions for all values of k and taking an average. Instead, it was assumed that the average of the four wave functions associated with $k = 0$ is the same as the average distribution of all electrons. This approximation is justified by the fact that the mean charge distribution in the alkali metals is practically the same as the distribution for $k = 0$. Boundary conditions were satisfied at different points of the polyhedra for several different values of k , group theoretical

¹ S. R. TIBBS, *Trans. Faraday Soc.*, **35**, 1471 (1939).

² D. H. EWING and F. SEITZ, *Phys. Rev.*, **50**, 760 (1936).

methods being used to choose the appropriate combination of zonal harmonics in each case.

Figure 21 shows the agreement between initial and final charge distributions in the final trials for lithium fluoride, and Figs. 22 and 23 give plots of the s - and p -band functions for $\mathbf{k} = 0$ for both lithium fluoride and lithium hydride.

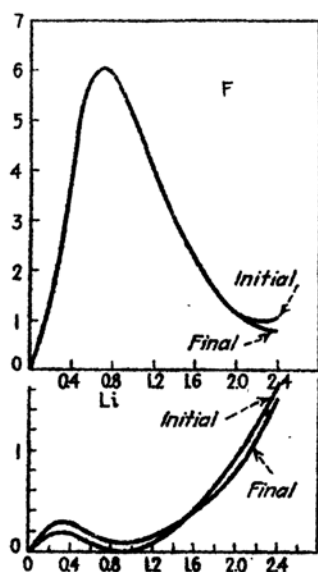


FIG. 21.—The initial and final charge distributions in the Li and F polyhedra in the last computation for the self-consistent field in lithium fluoride. The abscissae are in Bohr units. (After Ewing and Seitz.)

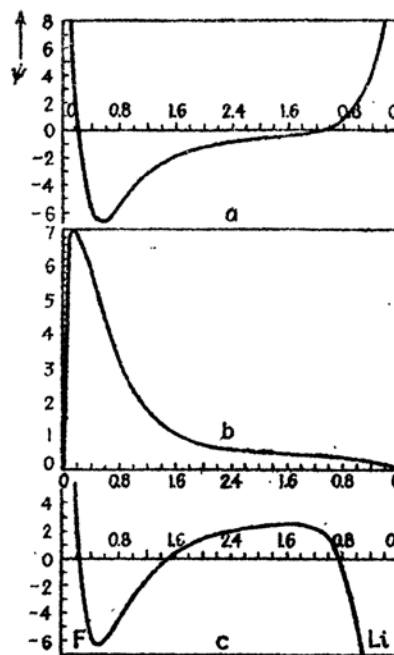


FIG. 22.—Wave functions for $\mathbf{k} = 0$ in lithium fluoride. a is the s -band function; b is the p -band function and c is the function for the first unoccupied band. It should be noted that a and b are distributed principally about the F^- ion, whereas c is distributed about both ions. The abscissae are in Bohr units.

The integral of the charge distribution inside the lithium sphere of lithium fluoride is $0.95e$ in the Hartree approximation. This result is unquestionably too large since neither exchange nor correlation effects were taken into account. A somewhat better value might have been obtained by including these terms in the way Shockley did, namely, by excluding one unit of electronic charge in determining the field inside the fluorine sphere. A rough estimate shows that this procedure would probably reduce the charge in the lithium sphere to about $0.5e$, which, in turn, would leave about 0.05 valence electron in the sphere the radius of which is equal to the classical lithium-ion radius. This estimate

furnishes a good justification for the Born-Mayer approximation in which the charge in this sphere is assumed to be zero, for the correction to the Madelung energy would be only about 5 per cent, which is less than the importance of repulsive terms.

Figures 24 and 25 show the $\epsilon(k)$ curves for several important crystallographic directions.¹ The values corresponding to the dots are the actual computed cases, whereas the full lines were obtained by interpolation. The upper curve of the second band for lithium fluoride is doubly degenerate in the (100) and (111) directions. This degeneracy also appears in Shockley's results.

One important feature of the zone scheme is the fact that the uppermost filled bands are several volts wide. Although these values probably are too large, because Hartree fields were used in obtaining them, their order of magnitude seems unquestionably to be correct.

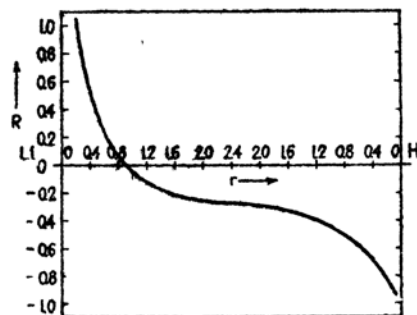


FIG. 23.—The wave function for $k = 0$ in the filled band of lithium hydride. The abscissae are in Bohr units. (After Ewing and Seitz.)

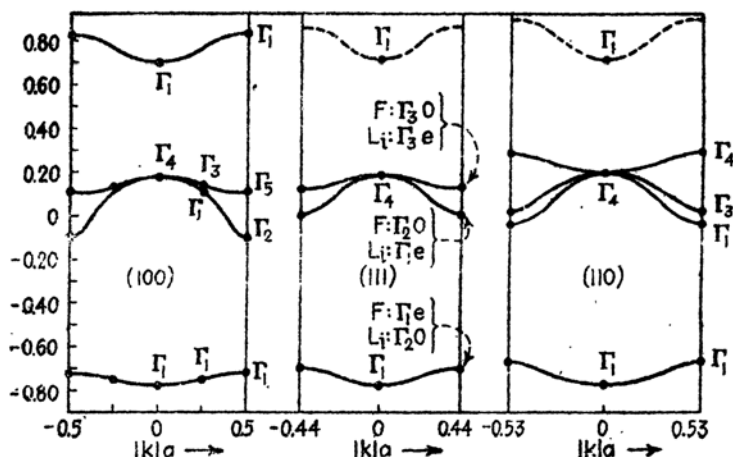


FIG. 24.— $\epsilon(k)$ curves of lithium fluoride for three prominent crystallographic directions. The Greek letters are crystallographic term symbols; a is the interionic distance. The lower two band systems are filled; the upper band is empty. The energy is expressed in Rydberg units. (After Ewing and Seitz.)

In summarizing this discussion of the normal states of the alkali halides, we may say that the charge distribution in the lattice is very

¹ It is interesting to note that the s curves increase with increasing $|k|$, whereas the p curves tend to decrease. This fact was pointed out in Sec. 65 on the basis of the narrow-band approximation.

nearly the same as if the crystal were composed of free positive and negative ions. The lowest s band of the zone picture is very narrow so that the halogen s shell is not appreciably perturbed. The width of the p band is of the order of 1 ev, which indicates that neighboring ions overlap appreciably and that the exchange interaction is of the order of 1 ev. The magnitude of the band width also indicates that the effective mass of a free hole in the p band is comparable with the mass of an electron.

The wave functions for $\mathbf{k} = 0$ in the first unoccupied zone of lithium

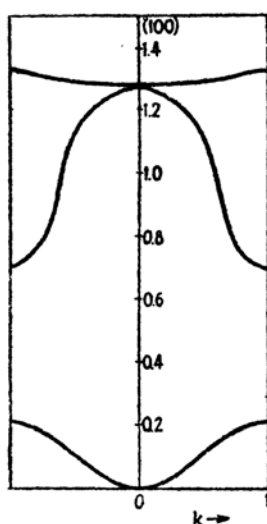


FIG. 25.—Same as Fig. 24 for the (100) direction in lithium hydride. There is only one filled band in this case so that the upper two bands are empty. These bands correspond to the occupied p bands of lithium fluoride (Fig. 24). The energy is expressed in Rydberg units.

fluoride are shown in Fig. 22. Here, the electronic charge is distributed more or less uniformly between the positive and negative ions. The closest vertical distance between the filled and unfilled bands of Fig. 24 is 7.5 ev for the end point of the zone in the (110) direction. This difference should be the energy required to induce photoconductivity in the pure crystal, and it should be greater than the first absorption energy. Actually it is less, since the fundamental peak of lithium fluoride occurs below 1000 Å, which makes the absorption peak greater than 12 ev. The Hartree approximation, can be blamed for this discrepancy, for the exchange terms would lower the filled band much more than the empty one. This tendency for the exchange energy to be smaller in magnitude for an excited electron than for a normal one is shown for perfectly free electrons in Fig. 5, Chap. IX.

We have seen in Sec. 95, Chap. XII, that the position of the nonconducting excited levels of the alkali halides can be estimated fairly closely by the use of an atomic model. According to this work, the first excitation band should lie about 12 ev above the lowest state in lithium fluoride, a value that agrees closely with the threshold absorption frequency for lithium fluoride. The complete structure of the first ultraviolet absorption bands of lithium fluoride and sodium chloride has not been measured. In a typical case such as that of sodium bromide, illustrated in Fig. 26, it seems natural to assume, in analogy with the absorption spectra for atoms, that the peaks A , B , C , and D correspond to transitions to excitation levels and that the ionization edge falls in the short wave length foot of the band, that is, at about 1200 Å. The

corresponding points for lithium fluoride and sodium chloride (cf. Fig. 2 of Chap. XII) undoubtedly lie at energies greater than 14 ev and 12 ev, respectively.

107. Alkaline Earth Oxides and Sulfides.—Although the cellular method has not been applied to any of the alkaline earth salts, it is possible to draw several plausible conclusions about their zone structures by indirect reasoning. The interatomic exchange terms, which give rise to most of the repulsive forces in ionic crystals and arise mainly from the metal-ion negative-ion interactions, are about four times larger in the oxides and sulfides than in the halides. Since these exchange energies are closely related to the widths of the occupied bands (cf. Sec. 64), we may expect that the bands are much broader in the oxides and sulfides. Referring to Fig. 19, we may expect that the point *B* where

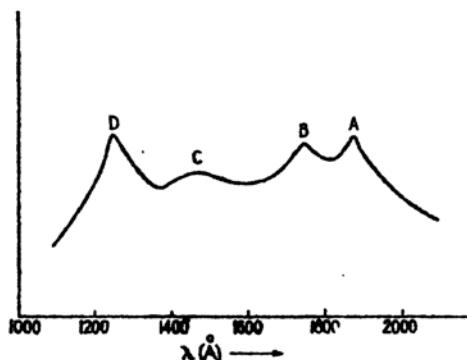


FIG. 26.—The structure of the first ultraviolet absorption band of sodium bromide.

the *s* and *p* bands are very close corresponds to the alkaline earth oxides and sulfides, if the point *A* corresponds to the alkali halides.¹

It is possible to construct reasonable energy-level diagrams for some of the alkaline earth salts by use of energy-level data derived from the Born cycle and from spectroscopic measurements on the free ions. As examples, we shall take zinc oxide and zinc sulfide, which have properties that are typical of other members of this group of salts. The crystal structure of zinc oxide is the wurtzite lattice, which is also the high-temperature structure of zinc sulfide. The low-temperature form of zinc sulfide has the zincblende lattice which is similar to diamond.

We shall begin by considering the energy necessary to remove an electron from a free negative ion and place it on a free zinc ion. The

¹ This overlapping of *s* and *p* bands occurs in diamond, as is shown in Sec. 109. Hence we may expect it to be associated with some valence characteristics. H. M. James and V. A. Johnson (*Phys. Rev.*, **56**, 119 (1939)) have shown, in fact, that the charge distribution in zinc oxide is not perfectly ionic.

total electron affinities of O^- and S^- have been determined approximately by the Born cycle and are about -7 and -4 ev, respectively. Lozier¹ determined experimentally the affinity of neutral oxygen for one electron and found that it is 2.2 ± 0.2 ev, which shows that the negative affinity of oxygen for two electrons is due entirely to the second electron and that the energy necessary to remove one electron from O^- is about -9 ev. The energy level of O^- is plotted² relative to the normal state of O^- in the right-hand column of Fig. 27. Since O^- and S^- have about

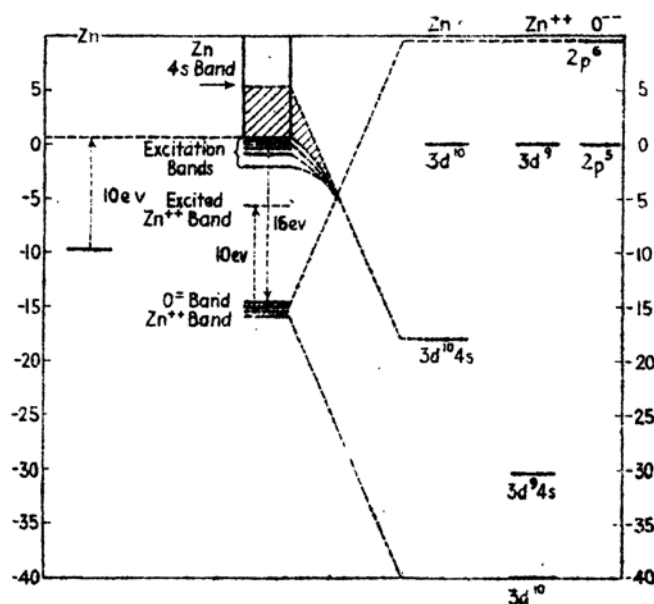


FIG. 27.—The ionic energy-level relations in zinc oxide. The levels of Zn^+ , Zn^{++} , and O^- are given on the right. The behavior of these when a lattice is formed is shown by the dotted lines. The relative position of the levels in a neutral zinc atom is given on the left. (See footnotes 9 on page 450).

the same classical radii as F^- and Cl^- , respectively, we may conclude that the electronic structure of the first pair of ions is very similar to that of the second pair.

Now, the affinities of the halogens decrease as one passes down the periodic chart. For this reason, we shall assume that the affinity of a sulfur atom for a single electron is about 1 ev less than that of an oxygen ion, which makes the energy of S^- about 5 ev relative to the energy of S^- . This is indicated in the right-hand column of Fig. 28. The energy levels of zinc ions have been measured spectroscopically and are given in the second and third columns from the right in Figs. 27 and 28. Since the

¹ W. W. LOZIER, *Phys. Rev.*, **46**, 268 (1934).

² F. SETTZ, *Jour. Chem. Phys.*, **6**, 454 (1938).

energy of Zn^+ relative to Zn^{++} is 17.9 ev, the energy required to remove an electron from a free oxygen ion and place it on a free zinc ion is -26.9 ev. The same quantity is -22.9 for S^- and Zn^{++} . Let us now arrange the ions in the lattices of ZnO and ZnS and gradually decrease the lattice constants from infinity. The lattice potential at the position of the positive ions is negative, and the potential at the negative ions is positive, so that the levels of the negative ions are depressed and those of the positive ions are raised during this process. This change is indicated by the dotted lines of Figs. 27 and 28, which show the total shift as computed from the Madelung potentials of the zinc oxide and zinc sulfide lattices.

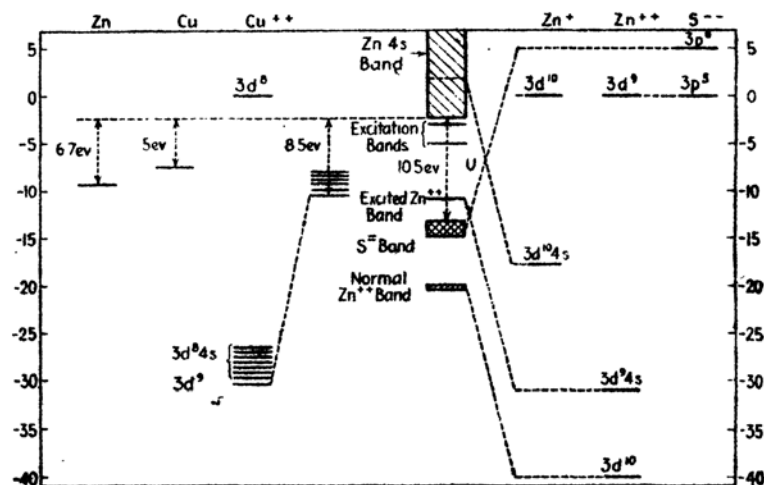
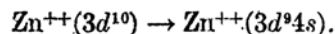


FIG. 28.—Same as Fig. 27 for zinc sulfide. The positions of levels of neutral zinc and copper, and of Cu^{++} are given on the left. (See footnote 1 on page 450.)

At the actual interatomic distances, the discrete levels of the free ions broaden into bands characteristic of band structure, and excitation bands appear¹ below the first unoccupied zone, which is adiabatically connected with the lowest level of free Zn^+ . The energy required to remove an electron from a negative ion and place it on a zinc ion at infinity, as found from these diagrams, is about 16 ev for zinc oxide and 10.5 ev for zinc sulfide. This transition evidently corresponds to ionization. The energy required to carry the electron from a negative ion to a near-lying zinc ion is less, of course, as is indicated in the figures. The first absorption band of zinc sulfide has been measured roughly, and the experimental work seems to show that the first excitation band should be about 6 ev above the ground state. The value of the absorption energy obtained from Fig. 28 is somewhat larger than this. Figure

¹ It is possible, however, that the excitation levels may merge with the ionization levels.

27 indicates that the first absorption band of zinc oxide may correspond to internal excitation of the zinc ion, that is, to the transition



If this is true, the first absorption peak of zinc oxide should lie near 1200 Å. There does not seem to be any work on the ultraviolet absorption properties of this salt.

Diagrams¹ such as Figs. 27 and 28 may be constructed for any ionic crystal for which the necessary data are known. Some additional uses for these diagrams will be found in Sec. 113.

108. Equilibrium Atomic Arrangements for Excited States.—The orderly lattice arrangement of atoms in crystals that is determined by X-ray analysis is the equilibrium distribution for the lowest electronic state. There is no reason for expecting the same distribution to be stable for excited electronic states of insulators. In fact, there is reason to expect the opposite, for each excited electronic state of diatomic or polyatomic molecules has its own equilibrium atomic arrangement. Since the dependence of excited electronic levels on atomic arrangement has not been investigated in a detailed and quantitative way, we shall have to be contented with qualitative pictures.²

Let us consider an ideal ionic crystal, such as one of the alkali halides or one of the alkaline earth salts. If we neglect thermal effects, the atoms occupy lattice sites in the normal electronic state. Suppose that we now use electrons or light quanta to excite the crystal to a higher electronic level. An excited electron and a hole are then produced, and the two should move together if the excited state is not conducting. The crystal still has the equilibrium atomic arrangement of the lowest level immediately after electronic excitation because of the Franck-Condon principle. Now, as we saw in the last chapter, the excited levels occur in systems of quasi-continuous bands, each level of which corresponds to an exciton moving with a definite velocity. If the exciton is produced by optical absorption, it usually is moving slowly, because the selection rules forbid transitions in which the wave-number vector of the exciton lies very far from the center of the zone and because the group velocity $\text{grad}_k \epsilon(\mathbf{k})/\hbar$ is zero when the wave number is zero. This selection rule is not valid, of course, if the excitation is induced by means of cathode rays or alpha particles, which have appreciable momentum; hence, the exciton may move more swiftly in these cases. If the exciton is regarded as an

¹ In both Figs. 27 and 28, the positions of levels of neutral atoms are represented on the left. The ionization energies of these atoms should be altered because of polarization effects such as those discussed in Sec. 112.

² Cf. A. VON HIPPEL, *Z. Physik*, **101**, 680 (1936). F. SEITZ, *op. cit.*, p. 150; *Trans. Faraday Soc.*, **35**, 74 (1939).

excited ion, it is easy to see that the lattice near it is under stress for the normal atomic arrangement, for an excited ion and a normal ion usually interact differently with their neighbors. These stresses would set the excited atom into oscillation about a new equilibrium position if the exciton were permanently at rest. If it is moving even slowly, however, the atoms near the exciton may not have time to move very far during the short time that the exciton is near. For example, the time required for an exciton that is moving at 10^6 cm/sec to traverse a distance of 10^{-8} cm is 10^{-14} sec, and the time required for an atom to make one oscillation is about ten times this. We may expect, however, that some lattice vibrations about the normal atomic positions are stimulated and that the exciton is slowed because it dissipates this vibrational energy. Thus, the exciton should eventually drop to the lowest excited energy state, and it should ordinarily be at rest in this state because $\text{grad}_{\mathbf{k}} \epsilon(\mathbf{k})$ is zero at the lowest point of the exciton band in simple crystals (cf. Sec. 96 of the previous chapter). The atom on which the exciton comes to rest should be set into violent vibrational motion because of the stresses mentioned above. This kind of motion is strongly damped since the atoms are strongly coupled. Thus, the localized vibrational energy should be dissipated during a time of the order of 10^{-12} sec, which is the atomic oscillation period, by the production of elastic waves that radiate from the vibrating atom.

The possible disposition of the normal and excited levels of the crystal is shown symbolically in Fig. 29. The abscissa represents the configurational coordinates of the lattice, that is, the interatomic distances, and the ordinate is the energy of the crystal. The lowest level corresponds to the normal electronic state so that its minimum *A* corresponds to the normal atomic arrangement of the lattice. The second curve represents the lowest excited state, in which the exciton is at rest. The minimum *B* corresponds to the values of the configurational coordinates for which the excited atom and its neighbors are at equilibrium. The quasi-

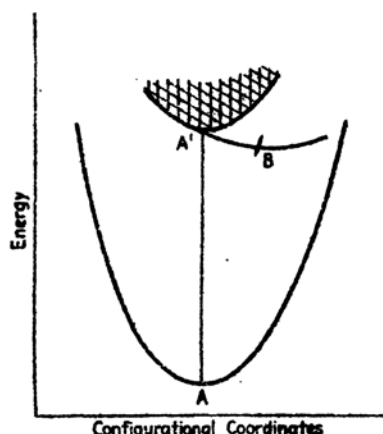


FIG. 29.—Schematic representation of the possible positions of normal and excited states of an insulator as functions of atomic coordinates. The lowest curve corresponds to the normal state and its minimum determines the normal atomic arrangement. The cross-hatched region corresponds to the levels of moving excitons. Their minima are at position *A* because the lattice cannot come to complete equilibrium when the excitons move. The discrete level that separates from this band represents the state of an exciton at rest. Its minimum is not at *A*, because the interactions between normal and excited atoms are different.

continuous band of levels represents the excited states in which the exciton is in motion. These levels have their minima at the same point as the normal state because the stresses are not localized when the exciton is in motion. During excitation, the system jumps from A to A' ; as soon as the exciton comes to rest, the state of the system "slides" toward B with the emission of elastic waves.

There are two conceivable arrangements of the normal and excited levels. In the first (Fig. 30a), the excited state has its minimum B within the minimum of the lower curve. In the second case (Fig. 30b), the minimum E is outside the lower curve. The system slides to B in the first instance and may then jump to the point C on the lower curve by emitting a light quantum. It slides to D in the second case

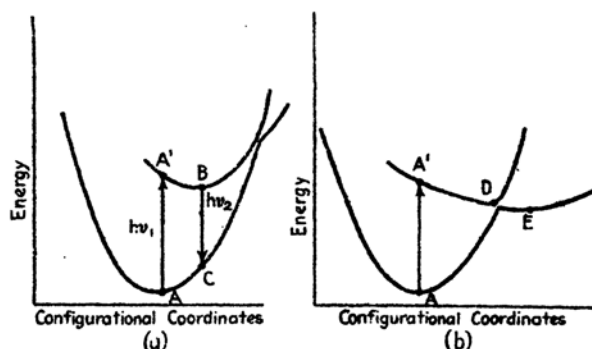


FIG. 30.—Possible arrangements of the normal level and the level of an exciton at rest. In a the minimum of the excited level is inside the lower curve so that fluorescent emission of frequency ν_2 may follow absorption of ν_1 . In case b the minimum of the excited level is outside that of the normal level. Hence, fluorescence cannot occur.

and may then slide either to A or to E . All the energy is dissipated immediately in the form of lattice waves in the first case. The system may rest at E indefinitely in the second, storing a part of the excitation energy. At temperatures above absolute zero, the system oscillates about E and should eventually pass over D and down to A . The crystal may be fluorescent in the case of Fig. 30a, since some of the absorbed energy may be radiated as light; however, it is not fluorescent in the second case. Fluorescence has never been observed¹ unambiguously to accompany absorption in the fundamental bands of ideal crystals, which indicates that the second case occurs commonly.

C. VALENCE CRYSTALS

109. The Zone Structure of Diamond.—In spite of the importance of the subject, practically no quantitative work has been done on valence

¹ See *ibid.*

binding in crystals. The reason for this lack, as we mentioned in Sec. 97, is that the simplest atoms entering into valence crystals have so many valence electrons which are appreciably affected by the crystalline binding that the computations are more complicated than for simple metals or salts. The existing work consists of a semiquantitative investigation of the zone structure of diamond and a qualitative discussion of the most appropriate form of the Heitler-London functions for the atoms in valence crystals with a tetrahedral arrangement of nearest neighbors. Carbon and silicon are the principal atoms to which these considerations apply.

The fact that the lowest state of atomic carbon is degenerate and that diamond is an insulator shows that the energy levels of the entire solid vary very much as the atoms of the lattice are brought together. The lowest level is highly degenerate at infinite separation, and it must broaden into a dense band just as the lowest levels of metal atoms do. Since the crystal actually is an insulator, we must conclude that a single level separates from this dense array and is the lowest level at the observed interatomic distance (cf. Fig. 31). A similar situation occurs in ionic crystals, for the highly degenerate state corresponding to free neutral atoms is more stable than the state of free ions at infinite separation; in this case, however, the two states cross before appreciable splitting occurs, since the Madelung energy favors an ionic state. For this reason, the situation is much easier to understand in ionic crystals. A simple picture of the same type has not been developed for diamond, although the separation of the singlet level can be shown to occur in the band approximation, as will be seen below. In fact, the separation occurs for such a simple carbon field that we may expect the effect is determined primarily by the crystal structure, that is, is connected with the way in which wave functions are diffracted by the diamond lattice.

a. *The Band Approximation.*—Kimball,¹ who is responsible for a semiquantitative zone treatment of diamond, satisfied boundary conditions at the center points of the four faces of the atomic polyhedron, which is shown in Fig. 3, Chap. IX. There are two polyhedra of this type in the unit cell. The boundary conditions, which were taken as the continuity of ψ and its normal gradient at these points, require that ψ be

¹ G. E. KIMBALL, *Jour. Chem. Phys.*, **3**, 560 (1935).

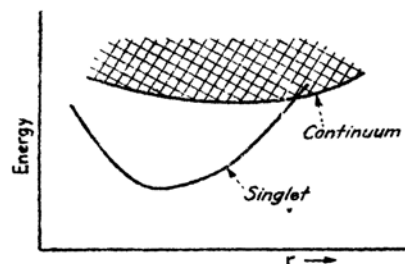


FIG. 31.—The dependence of the levels of diamond on interatomic distance (schematic). At large distances the lowest state is highly degenerate; however, a nondegenerate discrete level separates at smaller distances.

expanded in terms of four surface harmonics. Kimball took these to be one s function and three p functions, the field that was used to obtain the radial parts being the one derived by Torrance for the $2s$ and $2p$ functions of atomic carbon. The energy bands are shown¹ in Fig. 32. The atomic $2s$ level splits into two zones since there are two atoms per unit cell, whereas the atomic $2p$ function splits into six zones. These two systems of zones overlap at $r = 2.7a_h$ and then split into two new systems which contain four zones each. Two of the lower set of zones have zero width in Kimball's approximation, but they probably would

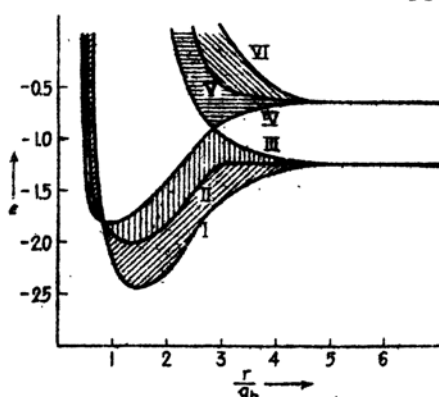


FIG. 32.—The band structure of diamond. It should be noted that the s - p bands overlap and break into two separate systems. From the standpoint of the entire solid this corresponds to the behavior of Fig. 31. (After Kimball.)

have finite width if more boundary points and surface harmonics had been used. Since the lower zone system is just exactly filled by the eight electrons per unit cell, the crystal is an insulator at the observed lattice distance in Kimball's approximation. There is no isolated low-lying group of four zones at large distances, however, so that the crystal should be a metal when r is greater than $2.7a_h$. From the standpoint of the entire crystal, this means that a single level separates from a continuum when r is $2.7a_h$.

Kimball found that his initial and final charge distributions were not the same, which shows that his starting field was inaccurate. That he still obtained the separation of bands which is needed to make diamond an insulator suggests, as we mentioned above, that this separation is determined primarily by the crystal structure. Thus, it is likely that carbon would be a metal if the atoms were placed together in one of the simple close-packed lattices.

The difference between the filled and unfilled bands of Fig. 32 is 7 eV, which corresponds to an absorption peak near 1700 \AA . It is probable, however, that excitation levels lie between these bands. The energy-level splitting is so large that the atomic-perturbation method of Sec. 96 cannot be used as an argument in favor of these levels. Instead, they should be treated by a more general method, such as that suggested by Wannier.

b. The Heitler-London Approximation.—The properties of tetravalent carbon in saturated hydrocarbon compounds suggest to the chemist that

¹ A similar figure has been derived by F. Hund, *Physik. Z.*, **36**, 888 (1935), on the basis of simpler reasoning.

carbon has tetrahedral directional properties. For this reason, he assigns a tetrahedral bond structure to carbon and assumes that this atom prefers to join with other carbon atoms or with hydrogen atoms along these bonds. The structure of diamond supports this viewpoint since each carbon atom in it is surrounded tetrahedrally by four other carbon atoms.

Pauling¹ and Slater² have independently established a set of principles that may be used to understand this tetrahedral character. Let us consider molecular hydrogen for a moment. As we have seen in Sec. 56, the stability of this molecule arises from the following two facts: (1) The field between two protons is stronger than the field of one. (2) Two electrons may share this region and minimize their repulsive energy by correlating their motion so that they are not there at the same time. When the binding is largest, the atomic distributions are distorted in such a way that the wave functions extend along the line of centers, where the field is largest. On the basis of results such as this, Pauling and Slater suggest that the observed atomic arrangements in valence compounds are those for which the Heitler-London functions have the largest peaks along the line of centers. In applying this principle to carbon, they assume that the carbon bonds are so strong that the 2s and the 2p electrons should be treated on an equal footing. Kimball's results support this assumption, for in his model the s levels split into bands that are as wide as the p bands and the two types of state become thoroughly mixed. Pauling found by direct computation that the four orthogonal functions which are linear combinations of one s function and three p functions and which have maximum directional localization extend toward the four corners of a tetrahedron. The equation of any one of these may be placed in the form

$$f(r)\left(\frac{1}{4} + \frac{3}{4} \cos \theta\right) \quad (1)$$

where $f(r)$ is a radial function and θ is the polar angle measured from the directional axis. The angular part of (1) is shown in Fig. 33.

It is evident that the Slater-Pauling principle cannot be rigorous, for carbon also forms a stable crystal in which the coordination is not tetrahedral, namely, graphite.

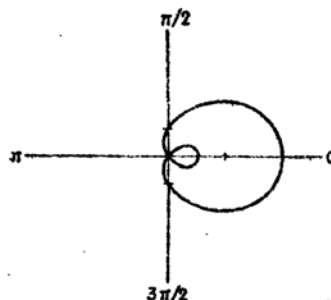


FIG. 33.—Polar plot of the Pauling bond function (1) for carbon.

¹ L. PAULING, *Jour. Am. Chem. Soc.*, **53**, 1367 ff. (1931).

² J. C. SLATER, *Phys. Rev.*, **37**, 481 (1931).

The electronic distribution that can be derived from Kimball's model undoubtedly shows directional properties similar to those of Pauling's tetrahedral functions. Since neither scheme has been used to make a quantitative computation of the cohesive energy of diamond, it is not possible to say which would give a better binding energy.

The same principle may be applied in discussing other simple valence bonds such as those that occur between silicon atoms in solid silicon or between silicon and oxygen in silica. The second case, in which each silicon atom is surrounded by four oxygen atoms and each oxygen atom by two silicon atoms, is complicated somewhat by the fact that oxygen, in place of having two electrons available for binding, actually lacks two electrons from a complete *p* shell. As we have seen in the previous discussion on solids, these holes may be treated like positively charged electrons. Hence, we may say that the silicon-oxygen bond in silica occurs between a directed valence electron of silicon and a directed hole of oxygen. We should expect these bonds to be polar because holes behave like positive charges.

D. SEMI-CONDUCTORS

110. General Principles.—There are two types of semi-conductor, namely, those which contain impurities and those which do not. Many polar salts, such as zinc oxide, belong to both classes. We shall be principally interested in pure semi-conductors, since their experimental properties have been studied more systematically than those of impure semi-conductors.

The principles determining the electronic conductivity of pure salts that have been given appropriate heat-treatment were first recognized and developed by Schottky and Wagner.¹ They pointed out, for example, that the electronic conductivity of pure zinc oxide can be associated with interstitial zinc atoms the presence of which may be understood in terms of ordinary principles of statistical mechanics. We shall present their work in a form that is modified in keeping with the treatment of statistical mechanics used in this book.

There are two main divisions of pure semi-conductors,² namely, those which conduct by free electrons and those which conduct by holes. These two types may be distinguished experimentally by the sign of the Hall coefficient; for the first class has the normal sign, that is, the same sign as the alkali metals and bismuth, and the second has opposite sign. It is easy to construct a system of electronic energy levels that explains

¹ W. SCHOTTKY and C. WAGNER, *Z. physik. Chem.*, **11B**, 163 (1930). C. WAGNER, *Z. physik. Chem.*, **22B**, 181 (1933).

² See, for example, the review by B. Gudden, *Ergebnisse exakt. Natur.*, **13**, 223 (1934); also Chap. I of this book.

qualitatively the difference between these two types. Consider, for example, a typical insulator having the system of filled and unfilled levels shown in Fig. 34a. If we add foreign atoms or distort the atoms in another way, we may expect to introduce new electronic levels in the forbidden regions. The electronic charge associated with these states is localized about the distortion or impurity atom.¹ The detailed properties of the additional levels vary from case to case and should be discussed separately in each one. If all the discrete levels are occupied, as in Fig. 34b, the electrons near the conducting bands may be thermally excited to this band, thereby making the crystal an electronic conductor. We discussed the properties of this type of semi-conductor in Chap. IV

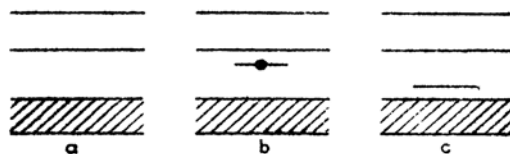


FIG. 34.—(a) The filled and unfilled levels in an insulator. (b) The discrete impurity level is occupied by an electron. This substance may be an electronic semi-conductor if the electron may be thermally excited to the empty band. (c) The level is unoccupied. This substance may be a hole semi-conductor with an 'anomalous' Hall coefficient, if electrons may be thermally excited from the filled band to this level leaving free holes in the lower band.

and found that the low-temperature conductivity σ should vary according to the equation

$$\sigma = n_b \frac{4\sqrt{2}}{3} \frac{c^2 l_0}{h^3} (2\pi m^* kT)^{1/2} e^{-\frac{\Delta\epsilon}{2kT}} \quad (1)$$

where n_b is the number of bound states per unit volume, l_0 is the mean free path, m^* is the effective mass of the free electrons, and $\Delta\epsilon$ is the activation energy for freeing the electrons (cf. Fig. 35). The sign of the Hall coefficient is normal in this case. If the discrete electronic levels are unoccupied (Fig. 34c), the electrons in the filled region may be thermally excited to the lowest unoccupied level, leaving free holes in the band. These holes should behave like free positively charged electrons and should give the solid a Hall coefficient whose sign is opposite to that of the preceding case. Equation (1) should also be valid in this case if the constants n_b , l_0 , m^* , and $\Delta\epsilon$ are reinterpreted in terms of free and bound holes. We shall discuss the electronic levels of several solids in more detail below.

¹ The behavior of models of this type has been discussed by A. H. Wilson, *Proc. Roy. Soc.*, **133**, 458 (1931), **134**, 277 (1931); R. H. Fowler, *Proc. Roy. Soc.*, **140**, 505 (1933); **141**, 56 (1933).

At temperature T , the equilibrium state of an insulator is determined by the condition that its free energy

$$A = E - TS$$

be a minimum. This quantity is simply E at zero temperature so that

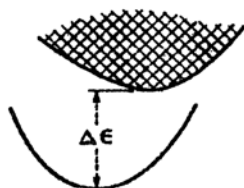


FIG. 35.—Schematic representation of the excited and normal levels in a semi-conductor containing an impurity atom. The minimum of the lower curve corresponds to the equilibrium atomic arrangement when the discrete level of the foreign atom is occupied. The minimum of the upper continuum corresponds to the equilibrium arrangement when the electron is ionized. The two minima are different because the ionized foreign atom does not interact with the lattice in the same way as the un-ionized one. For this reason the energy $\Delta\epsilon$ for thermal ionization is usually less than the energy for optical ionization, for the second process must obey the Franck-Condon principle and corresponds to a vertical jump in this diagram.

the equilibrium condition presumably demands that the ideal crystalline arrangement be most stable at this temperature, since E probably is then a minimum. This arrangement need not be the most stable above the absolute zero; for if the entropy associated with a distorted arrangement, such as that caused by placing some atoms in interstitial positions or by removing normal atoms, is large enough, the distorted state is more stable than the normal one. Let us consider a simplified system consisting of a monatomic lattice of N atoms. We shall let ϵ be the energy that is necessary to remove one atom from a typical lattice site and to place it at the surface in a normal position. It will be assumed that this energy is independent of the number n of atoms removed for small values of n . If we neglect any changes in the vibrational energy that may occur as a result of this transposition, the total entropy change is determined by the number of possible ways in which the n vacancies may be distributed among the N sites. This number evidently is $N!/n!(N-n)!$ so that the entropy S is

$$S = k \log \frac{N!}{n!(N-n)!} \simeq kn \log \frac{n}{N}. \quad (2)$$

Thus, the free energy

$$A = n\epsilon - kTn \log \frac{n}{N} \quad (3)$$

is a minimum when

$$\frac{n}{N} = e^{-\frac{\epsilon}{kT}}.$$

This result shows that we should expect some deviations from the ideal crystalline state at any finite temperature. If ϵ is 1 ev and T is 1000°K, which are values that can reasonably occur, we find

$$\frac{n}{N} \sim 10^{-4.5}.$$

The principles used in this computation may be applied to other cases, most important of which for the theory of semi-conductors are the polar crystals having the composition M_mX_n , where M is a metal atom and X is an electronegative atom. The four independent types of deviation from ideal arrangement that may occur in these solids are as follows: There may be vacancies (a) in the metal lattice or (b) in the electronegative lattice; and there may be (c) interstitial metal atoms, or (d) interstitial electronegative atoms. These four types of *lattice defect* may occur in any one of the various possible combinations. We shall discuss a few actual cases in the following sections.

111. The Alkali Halide Semi-conductors.—Hilsch, Pohl,¹ and their numerous collaborators have made extensive investigations of semi-conducting alkali halide crystals that are produced by heating these solids in alkali metal vapor until they become colored. Figure 36 shows

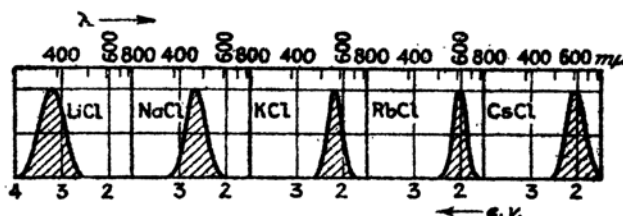


Fig. 36.—*F*-center absorption bands at room temperature in various alkali halides. The wave-length scale is in units of 10^{-7} cm. (After Pohl.)

the spectral dependence of the new absorption band in several cases. Identical discoloration may be produced by X-ray or cathode-ray bombardment. The discoloration is not so durable in these cases, however, for it may be removed by relatively mild heating that does not affect the discoloration produced by heating the crystals in alkali vapor. From the intensity of the absorption bands, it is possible to determine the number of absorption centers that are responsible for the discoloration. This number depends upon the method used to prepare the colored crystal and varies from 10^{18} to 10^{19} per cubic centimeter in the specimens ordinarily used in experiments. Pohl has named these absorbing centers "*F* centers," (*Farbzentren*). We shall use the same term.

The electronic conductivity of the alkali halides containing F centers becomes appreciable above 200°C and is superimposed upon the ionic conductivity, which is also appreciable. The two may be separated by measuring either the conductivities of separate clear and colored specimens or the conductivity before and after the F centers have been removed. This removal can be accomplished by placing the colored

¹ See the survey article by R. W. Pohl, *Physik. Z.*, **39**, 36 (1938).

crystal in a constant field at the temperatures at which the conductivities are measured. The F centers then migrate to the anode and disappear. During this procedure, the conductivity drops, as is shown in Fig. 37. Electronic conductivity may be induced at low temperatures by illuminating the crystal with light in the absorption bands. We shall discuss this photoconductivity in detail in Chap. XV.

The ionic conductivity of the alkali halides usually is caused by the migration of both positive and negative ions, as may be determined by transport measurements, such as those discussed in Chap. I. Frenkel¹ first pointed out that the migrating positive and negative ions probably do not move by squeezing past one another, as they would in an ideal

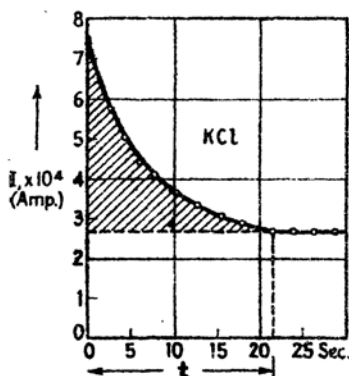


FIG. 37.—Decrease in the current of colored potassium chloride as the F centers are removed by conduction. The initial conductivity is due to ions and electrons, the final conductivity to ions alone. The temperature was 580°C ; the field intensity was 300 volts per cm. (After Pohl.)

lattice, since the activation energy² that would be required for this process is much larger than the activation energies determined by measurements of the temperature dependence of conductivity. He estimated that the activation energy in an ideal lattice would be about the same as the cohesive energy. This is about 7 ev in sodium chloride, whereas the observed value³ of the activation energy is only 1.9 ev. For this reason, Frenkel postulated that alkali halide crystals normally contain vacancies in both the positive- and negative-ion sites. We may conclude that these vacancies are present in equal numbers in uncolored crystals, for otherwise these crystals would be charged. They presumably have the same thermodynamical origin as the vacancies in the simple monatomic lattice discussed at the end of the previous

section and can be discussed in terms of the theory used there.

It is possible to give two reasonable pictures of the discoloration of alkali halides by X rays. In both pictures, it is assumed that the primary action of the X rays is to free an electron from one of the inner shells of an atom of the lattice and that the discoloration center is associated with the absorption properties of this electron when it subsequently becomes trapped in the lattice. The most apparent trapping positions are the vacancies in the negative-ion lattice and the positive ions. The

¹ J. FRENKEL, *Z. Physik*, **35**, 652 (1926).

² The activation energy is the least energy required to interchange two atoms. This is discussed in detail in the next chapter.

³ W. LEHFELDT, *Z. Physik*, **85**, 717 (1933).

vacancies should have an affinity because the Madelung potential is positive at these positions, whereas the positive ions should have an affinity because an electron should be able to polarize the surrounding lattice and produce a stable discrete level lying below the conduction bands discussed in Sec. 106 of this chapter. Evidence obtained from investigations of photoconductivity seems to support the first interpretation of the trapping position and rule out the second. If any alkali atom could trap an electron and produce an F center, the mean free path for trapping of free electrons should be independent of the density of F centers and of vacancies in the negative-ion lattice, since the number of these is far less than the number of alkali metal ions. The experimental work on the photoconductivity of crystals containing F centers shows that the density of trapping points is far less than the density of alkali metal ions and depends upon the density of F centers. Hence, it may be concluded that F centers are electrons trapped in vacant halogen-ion sites.¹ There has been no experimental evidence to show that electrons are ever trapped by the positive ions.

When an alkali halide crystal is heated in alkali metal vapor, some of the atoms of the vapor presumably become absorbed on the surface and lose their electrons. These electrons then wander into the crystal and occupy vacant halogen positions, producing F centers. The ions left behind may then diffuse into the lattice, decreasing the number of positive-ion vacancies and keeping the volume of the crystal unchanged. Let us suppose that the crystal is at temperature T and that it normally contains n vacancies in the positive- and negative-ion lattice. In addition, let us suppose that it is placed in a container of which the volume V is much larger than that of the crystal and which contains N_A neutral alkali metal atoms in vapor form. If n_F atoms are absorbed into the crystal, the number of vacancies in the positive- and negative-atom sites is decreased from n to $n - n_F$. The mixing entropy associated with the vacancies is

$$-2k(n - n_F) \log \frac{n - n_F}{N} \quad (1)$$

where N is the number of ions in the lattice. In addition, the n_F electrons that occupy the halogen sites have the entropy

$$-kn_F \log \frac{n_F}{N}, \quad (2)$$

since they may occupy any of the N sites. The vapor has the entropy

¹ This interpretation of F centers is due to J. H. de Boer, *Rec. trav. chim. Pays-Bas*, **56**, 301 (1937), and has been developed by R. W. Gurney and N. F. Mott, *Trans. Faraday Soc.*, **34**, 506 (1938).

$$-k(N_A - n_F) \log \frac{(N_A - n_F)}{C} \quad (3)$$

where

$$C \cong \frac{V}{h^3} (2\pi\mu kT)^{3/2}$$

in which μ is the atomic mass. Thus, the total free energy as a function of n_F and n is

$$A(n_F, n) = n_F \epsilon + (n - n_F) \epsilon' + kT \left[2(n - n_F) \log \frac{n - n_F}{N} + n_F \log \frac{n_F}{N} + (N_A - n_F) \log \frac{(N_A - n_F)}{C} \right] \quad (4)$$

where ϵ is the energy required to dissociate an atom into an F center and

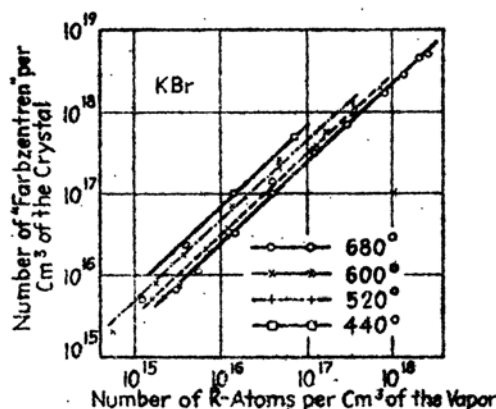


FIG. 38.—Relation between the density of F centers in a potassium bromide crystal and the density of alkali metal atoms in the vapor for different temperatures. (After POHL.)

an ion and ϵ' is the energy required to form normal lattice defects. Then A is a minimum when

$$\frac{n_F}{(N_A/C)} = Ne^{\frac{\epsilon}{kT}}. \quad (5)$$

According to this result, the ratio of the concentration of F centers to the concentration of atoms in the vapor should be constant. Figure 38 shows that this relation is obeyed in KBr over a wide range of concentrations.¹ In addition, the temperature dependence is the same as that predicted by (5). Rögner finds experimentally that $\epsilon = -0.25$ ev for KBr and -0.10 ev for KCl.

¹ POHL, *op. cit.*

If the equilibrium density of F centers is established at one temperature and alkali metal pressure and the crystal is then cooled, the excess F centers should coagulate into colloidal globules of alkali metal unless the cooling takes place so rapidly that the higher density is frozen in.

Following a plan of Gurney and Mott,¹ we may obtain a rough idea of the energy levels of an F center by using classical methods. The Madelung potential at a vacant negative-ion site is $-Ae^2/r_0$ as long as the surrounding ions have perfect crystalline order. Here, A is the Madelung constant and r_0 is the nearest-neighbor distance. This poten-

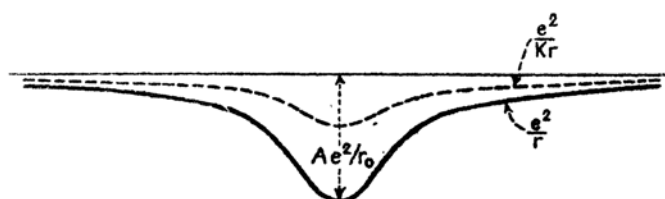


FIG. 39.—The potential trough for an electron near a halogen-ion vacancy. The full curve represents the potential when polarization is neglected, the dotted curve the potential when it is not.

tial is of the order of 9 ev for sodium chloride. At large distances from the vacant site, the total potential is

$$-\frac{e^2}{r} + V_M \quad (6)$$

where V_M is the periodic Madelung potential of a normal lattice and e^2/r is the potential arising from the vacancy. The mean value of V_M is very close to zero for an electron placed in an alkali halide crystal (cf. Sec. 91), and so we are justified in dropping it from (6) in a good approximation. The remaining potential then varies smoothly between the value $-Ae^2/r_0$ at the vacant site and the value $-e^2/r$ at large distances, in the manner shown in Fig. 39. If there were no electron in the vacant site, the surrounding ions would be displaced from their equilibrium positions for the normal lattice. We shall assume, however, that they are nearly in their normal positions when an electron is present, for this electron should have nearly the same electrostatic effect on the neighboring ions as a halogen ion. When the electron is at large distances from the site, it is partly screened from the excess positive charge by the polarization charge that it induces in the crystal. Hence, the potential at large distances should vary roughly as e^2/n^2r instead of as e^2/r , where n is the refractive index. Thus, the potential well in which the electron is trapped should have the form of the second curve, Fig. 39. This trough has an infinite number of discrete levels, which end in a series

¹ R. W. GURNEY and N. F. MOTT, *Proc. Phys. Soc. (sup.)*, **49**, 32 (1937).

limit at zero energy, since the field is coulomb-like at large distances. The lowest of the discrete states is an s state, and the next two are s and p states. The optical transition from the lowest s state to the lowest p state should have the greatest probability and should correspond to the absorption band of the colored crystal. This transition should not lead directly to photoconductivity, for the excited electron is bound to the vacant site just as the normal electron is. The excited state lies nearer the continuum, however, so that the probability for a thermal transition to the continuum is higher from it than from the lowest level. In order to account for the observed photoconductivity, we must assume that the electron actually becomes free by thermal excitation. The probability for this excitation should decrease with decreasing temperature and vanish at absolute zero. This disappearance of photoconductivity near the absolute zero is actually observed.¹ We shall discuss the effect further in Sec. 134.

F centers are not the only color centers that may be introduced into alkali halide crystals.² Thus, Pohl and his coworkers have found that a stoichiometric excess of halogen atoms may be produced by heating the crystal in halogen vapor. Most of the coloration lies in the near ultraviolet region of the spectrum in this case. Although these centers have not been investigated so fully as F centers, it seems probable that they are neutral halogen atoms in halogen sites. The F center absorption band may be destroyed and a new band may be introduced in the far ultraviolet by heating a crystal containing F centers in hydrogen vapor. In this case, it is believed that the hydrogen atoms diffuse to the F centers and form H^- ions at these positions. The far ultraviolet absorption band of these " U centers" presumably corresponds to the first excitation frequency of the internally absorbed hydrogen ions.

112. Zinc Oxide.—Zinc oxide that has been formed at low temperatures is a pure white substance having no appreciable electronic conductivity. After being heated to a high temperature, it develops a brownish hue and is a good electronic conductor at room temperature. A preliminary discussion of these properties was given in Sec. 37, which deals with the free-electron theory of semi-conductors. We saw there that the room-temperature conductivity obeys Eq. (1), Sec. 110 of this chapter. For the purposes of this discussion, this equation may be placed in the form

$$\sigma = Ae^{-\frac{e}{kT}} \quad (1)$$

¹ *Ibid.*

² Recently H. Pick has investigated the optical and electrical properties of colored halides containing divalent strontium halides, *Ann. Physik*, 35, 73 (1939).

where ϵ' and A are not strongly temperature-dependent in the range over which conductivity is measured. It was also seen in Sec. 37 that these constants are influenced by the pressure of oxygen in which the specimen is heated. Since A is related to the mean free path l_0 and the density of centers n_b by the equation

$$A = 0.024l_0n_b^{1/2}T^{1/2} \text{ ohm}^{-1} \text{ cm}^{-1}$$

and since Hall-effect measurements show that l_0 is practically constant for a given specimen, we may conclude that the variation of A with oxygen pressure implies a variation in n_b . This variation has been investigated by Wagner and Baumbach;¹ we shall now discuss their results.

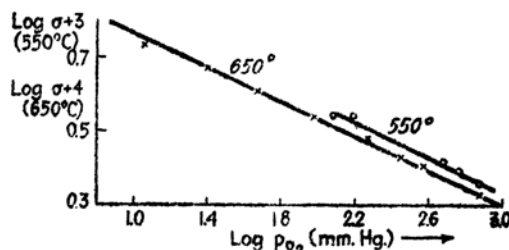


FIG. 40.—Dependence of the conductivity of zinc oxide on vapor pressure of oxygen (After Baumbach and Wagner.)

Figure 40 shows the variation of high-temperature conductivity with oxygen vapor pressure for a specimen that has been heated at two temperatures. Practically all the n_b bound electrons are free at the temperatures at which these measurements are made, for $\Delta\epsilon$ then is appreciably less than kT . Hence, the curves of Fig. 40 give directly the dependence of n_b on oxygen pressure p_o . They show that

$$n_b = cp_o^{-\frac{1}{n}} \quad (2)$$

where $n \simeq 4.2$.

All the experimental results may be explained satisfactorily if we assume that the heated zinc oxide loses oxygen atoms from the surface and leaves excess zinc atoms, which become ionized and diffuse into the interstices of the lattice. Wagner has ruled out the alternative possibility that vacancies are produced in the oxygen lattice and that the conduction electrons are those which might normally occupy these vacancies by showing that the negative-ion transport number is very small compared with the positive-ion transport number. The small observed positive ionic current is carried either by ionized interstitial

¹ H. H. v. BAUMBACH and C. WAGNER, *Z. Physik Chem.*, **22B**, 199 (1933).

zinc atoms or by the normal zinc ions. There is ample room for interstitial zinc atoms in the zinc oxide lattice because it has the porous wurtzite structure.

The displacement of equilibrium with changing oxygen pressure may be treated in the following way. We shall let ϵ'' be the energy required to produce a singly charged interstitial zinc ion, a free electron in the lattice, and one atom of gaseous oxygen, which is attached to another oxygen atom to form an O_2 molecule. We shall let n be the number of interstitial Zn^+ ions and N_{O_2} the total number of O_2 molecules in the gas, where N_{O_2} is much larger than n . The total free energy then is

$$A(n, N_{O_2}) = n\epsilon'' + kT \left[n \log \frac{n}{N} + n \log \frac{n}{C} + \left(N_{O_2} + \frac{n}{2} \right) \log \frac{\left(N_{O_2} + \frac{n}{2} \right)}{C} \right] \quad (3)$$

where the first entropy term is that of the interstitial ions, N being the total number of interstitial sites, the second term is the entropy of the free-electron gas, and the third term is the entropy of the O_2 molecules. The equilibrium value of n is

$$nN \delta_i = B e^{-\frac{\epsilon''}{kT}} \quad (4)$$

where B is a constant. Hence, n should vary as $p_{O_2}^{-1}$ according to this simple theory. The result is in reasonable agreement with the observed variation of $p_{O_2}^{-\frac{1}{4.3}}$.

In the first approximation, we might treat the energy levels of the interstitial zinc atoms as though they were free atoms in a homogeneous, polarizable medium. The principal effect of the polarizability¹ is to decrease the distance between the ground state and the continuum, as we have seen in the previous section. Suppose that we had a hydrogen atom in a medium of refractive index n . Then, the potential between the electron and proton would be $-e^2/n^2r$, where r is the radial distance between the two particles. The presence of n in the potential energy changes the Rydberg constant to R/n^4 where R is the normal value for a free atom. The refractive index of zinc oxide is about 2, so that we should expect the ionization energy to be lowered by a large factor, of the order of magnitude 10. The same qualitative result should apply to zinc, which has an ionization potential of 9.36 ev, and should lower the ionization energy of the interstitial atom to about 1 ev. The observed values of ϵ' in Eq. (1), however, are even lower than this value. For example, ϵ' is commonly less than 0.01 ev in a specimen that has been

¹ N. F. MOTT and R. W. GURNEY, *Proc. Phys. Soc. (sup.)*, **49**, 32 (1937).

heated for a long while in a vacuum. Moreover, Fritsch (*cf.* Sec. 37) has found that ϵ' in Eq. (1) varies with the pressure of oxygen and has shown that ϵ' increases as the density of interstitial atoms decreases. This effect indicates that the interstitial zinc atoms interact with one another and in some way decrease the distance between the bound and free levels. The density of interstitial atoms is of the order of 10^{18} , according to Hall-effect measurements, so that this interaction is conceivable only if the radius of the interstitial atom is ten times larger than the radius of a normal zinc atom. Now, the radius of a hydrogen atom in a medium of index n would be n^2 times larger than the radius of a normal atom. Thus, it is possible that the electrons in the interstitial atoms move in very large orbits because the surrounding medium is highly polarizable.

113. Cuprous Oxide and Other Substances That Involve Transition Metals.—Cuprous oxide is a very useful semi-conductor, but its highly intricate properties are only partly understood. The most reliable evidence seems to show that it has hole conductivity, which indicates that it has either a deficiency of metal atoms or an excess of oxygen atoms. Wagner and his coworkers¹ have evidence to show that the copper ion is much more mobile than oxygen, and they conclude from this that the conducting oxide probably contains vacancies in the copper-ion lattice. These are shown schematically in Fig. 41. Since the deficient copper ion carries away an electron, the lattice should contain one electron hole for each vacancy. This hole may normally reside either on a copper ion, turning a Cu^+ ion into a $\text{Cu}^{+ -}$ ion, or on an oxygen ion, turning an O^- ion into an O^{--} ion. Wagner suggests that the first picture is more probable since copper is commonly bivalent. According to this view, the conductivity of cuprous oxide should result from the motion of the hole from one copper ion to another. The most stable position for the hole should be near the vacancy since there is an excess negative charge at that position.

De Boer and Verwey² have attempted to make a systematic classification of other semi-conductors containing metals with partly filled $3d$ shells. They computed the energy of electrons on atoms and ions near vacant sites, using atomic-model methods similar to those that we have

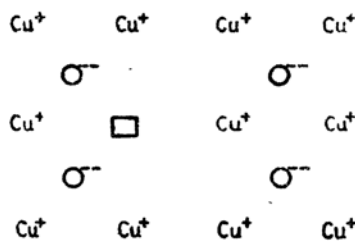


FIG. 41.—Schematic representation of copper vacancies in cuprous oxide. The vacancy leaves the lattice with a deficiency of one electron.

¹ C. WAGNER, *Physik. Z.*, **36**, 721 (1935).

² J. H. DE BOER and E. J. W. VERWEY, *Proc. Phys. Soc. (sup.)*, **49**, 59 (1937).