# Electronic Energy Bands in Metals 

J. C. Slater, Massachusetts Institute of Technology

(Received April 5, 1934)


#### Abstract

The method of Wigner and Seitz is extended to the computation of the excited bands of electrons in a metal, with particular applications to sodium. Instead of using simply one $s$ wave function, as Wigner and Seitz do, a combination of eight separate functions is used, one $s$, three $p$, three $d$, and one $f$. Boundary conditions for an arbitrary electron momentum are satisfied at the midpoints of the lines connecting an atom with its eight nearest neighbors. The solution is carried out for an arbitrary direction of propagation in one of the principal planes. Energy levels and wave functions are determined as functions of internuclear distance, leading to the


#### Abstract

following qualitative results: At the observed distance of separation, energy levels are given with remarkable accuracy by the Fermi-Sommerfeld theory, the gaps fall approximately where they should as computed from de Broglie waves, and the wave functions act accurately like plane waves in the region between atoms, but fluctuate violently, like $s, p, \cdots$ functions, near the nuclei. Gaps in energy are precisely filled up, though in each definite direction of propagation there are gaps. As the internuclear distance increases, gaps in energy appear at definite points, the allowed regions shrinking to zero breadth about the atomic energy levels at infinite separation.


## Introduction

WIGNER and Seitz ${ }^{1}$ have proposed an important improvement in the methods of finding wave functions for electrons in periodic force fields, such as one encounters in metals. Previous methods had proceeded by perturbation theory; a good account of these methods is found in the excellent article by Sommerfeld and Bethe ${ }^{2}$ in the Handbuch der Physik. Many of the results quoted in that article will be assumed familiar to the reader, as well as the paper of Wigner and Seitz. The advance of Wigner and Seitz lay in their observation that the potential field acting on an electron in a metal is very approximately spherically symmetrical in the neighborhood of a nucleus, so that the wave equations can be solved by separation of variables, and numerical integration of the equation for the radial function, as in problems of isolated atoms. Only the boundary conditions distinguish the problem from an atomic one. They imagine the crystal to be made up of closepacked cells, one surrounding each nucleus, approximately spheres but bounded by planes so that they exactly fill the space. Then the wave function must satisfy the condition that it be continuous

[^0]with continuous derivative in going from one cell to the next. Further, they consider only the lowest electron level, for which the wave function is periodic, repeating in each cell. It is easy to see that their conditions then demand that the normal derivative of the wave function be zero around the surface of the cell, which they approximate by using a spherical cell of the same volume as the actual one, and making the radial derivative of the wave function zero on the surface of the sphere. By applying this condition they obtain an energy level as function of internuclear distance (or of radius of the sphere), showing a minimum near the observed internuclear distance, and which they could bring into connection with the heat of dissociation of the crystal. They estimate the excited energy levels, corresponding to electrons having linear momentum of translation through the crystal, from the Fermi distribution, adding the kinetic energy of free electrons to the energy derived for the lowest state, and correctly pointing out that the actual band of energy levels will necessarily be narrower than the band so computed. It is the purpose of the present paper to extend the method of Wigner and Seitz to an actual calculation of these excited energy levels.

## 1. Formulation of Boundary Conditions

Consider a wave which is modulated, in passing through the crystal, according to the plane wave function $\exp (i(\mathbf{k} \cdot \mathbf{r}))$, so that $\mathbf{k} / 2 \pi$ is a vector along the wave normal, equal numerically to the wave number, and $\mathbf{k} h / 2 \pi$ would be the mechanical momentum, if the wave represented a free electron. Then according to a theorem of Bloch (for this and similar references, see Sommerfeld and Bethe ${ }^{2}$ ), if we go from one point in the lattice to an equivalent point in another cell, the wave function is multiplied by the factor $\exp \left(i\left(\mathbf{k} \cdot \mathbf{r}_{i j}\right)\right)$, where $\mathbf{r}_{i j}$ is the vector from the nucleus $i$ of the first cell to the nucleus $j$ of the second. In particular, consider one cell and an adjacent cell, so that $\mathrm{r}_{i j}$ is not only the distance between adjacent nuclei, but the perpendicular distance between the opposite faces of the cell, which are cut at right angles by the line joining the nuclei in question. Going from a point of one of these faces to the perpendicularly (not diametrically) opposite point of the other is then such a translation as we have considered. Our boundary condition then is: in going from a point of the surface of the cell to the perpendicularly opposite point, the wave function must be multiplied by $\exp (i(\mathbf{k} \cdot \mathbf{r}))$, where $\mathbf{r}$ represents the perpendicular distance between faces. Similarly, each of the three components of the gradient must be multiplied by the same factor. The latter condition can be restated. The tangential components of gradient will be automatically taken care of, so that we consider only the normal components. We note that the absolute direction in space which is along the outer normal for one face is along the inner normal for the other. Thus we require that the derivative along the outer normal be multiplied by $\left[-\exp \left(i\left(\mathbf{k} \cdot \mathbf{r}_{i j}\right)\right)\right]$ in going from one point to the perpendicularly opposite point. This is the general and exact boundary condition.

Our problem is to solve a spherically symmetrical wave equation, subject to the boundary condition stated above on the surface of a polyhedral cell. The problem is in essence soluble. We may separate variables, and solve the equation in a series of products of spherical harmonics of the angle, multiplied by functions of $r$, the coefficients of the series being arbitrary. Such a
solution can be found for any energy value, and for any energy value we can satisfy the one boundary condition that the function be finite at the origin, though except for the characteristic values of the equation these solutions will become exponentially infinite at infinity. This however does not concern us, since we work in finite cells. It is now in principle possible to choose the infinite number of coefficients so as to satisfy the boundary conditions. The procedure, however, is impracticable, and here we introduce an approximation. We try to get the approximate solution by summing a finite number out of the infinite series of solutions, and consequently we can satisfy the boundary conditions, not at every point of the surface, but only at certain discrete points. In the actual calculations described, we have dealt with a body centered cubic lattice, and we have satisfied the boundary conditions at the eight midpoints of the lines joining an atom to its eight nearest neighbors; that is to say, referring to $W$. and $S$. Fig. 1, at the centers of the eight hexagonal faces of the cell. Corresponding to this, we require only eight wave functions; and using the conventional spectroscopic notation for $l$ values, we find that this demands one $s$ function, three $p$ 's, three $d$ 's, and one $f$. It now proves to be the case that we can satisfy the boundary conditions at the centers of faces particularly easily, for then corresponding points are diametrically opposed (at opposite ends of a line passing through the nucleus) as well as perpendicularly.

Let $u$ be the wave function we are considering, a function of three coordinates. Let it be a sum of an even part $u_{g}$ which is unchanged on going from one point to the diametrically opposite point, and an odd part $u_{u}$ whose sign changes. Then if the function is $u_{g}-u_{u}$ at the midpoint of one face, it will be $u_{g}+u_{u}$ at the opposite point, and similarly if its normal derivative is $u_{g}{ }^{\prime}-u_{u}{ }^{\prime}$ at one midpoint, it will be $u_{g}{ }^{\prime}+u_{u}{ }^{\prime}$ at the opposite point, where primes indicate normal derivatives. Hence our boundary conditions may be stated as

$$
\begin{aligned}
\left(u_{g}+u_{u}\right) & =[\exp (i(\mathbf{k} \cdot \mathbf{r}))]\left(u_{g}-u_{u}\right) \\
\left(u_{g}^{\prime}+u_{u}^{\prime}\right) & =-[\exp (i(\mathbf{k} \cdot \mathbf{r}))]\left(u_{g}^{\prime}-u_{u}^{\prime}\right)
\end{aligned}
$$

where the $u$ 's are to be computed at the mid-


Fig. 1. Energy bands for sodium. Symbols Ia-IIb represent type of state allowed in each band.
points of faces, $r$ is the vector joining midpoints, and there are in our case four such conditions, for the four pairs of opposite faces.

The equations above can be easily transformed to a more usable form. The first one gives at once

$$
\frac{u_{g}}{u_{u}}=-\frac{1+\exp (i(\mathbf{k} \cdot \mathbf{r}))}{1-\exp (i(\mathbf{k} \cdot \mathbf{r}))}=-i \cot (\mathbf{k} \cdot \mathbf{r} / 2),
$$

and the second gives

$$
u_{g}{ }^{\prime} / u_{u}{ }^{\prime}=i \tan (\mathbf{k} \cdot \mathbf{r} / 2)
$$

Since it is allowable to multiply the whole solution by an arbitrary factor, let us assume that $u_{g}$ is real. Then the equations demand that $u_{u}$ be pure imaginary, so long as $k$ is real, which is necessary for real, undamped propagation. If $\left|u_{u}\right|$ is the absolute magnitude of $u_{u}$, we then have $u_{u}=i\left|u_{u}\right|$, so that the equations above may be combined in the form

$$
\left|u_{u}\right| / u_{g}=-u_{g}{ }^{\prime} /\left|u_{u}\right|^{\prime}=\tan (\mathbf{k} \cdot \mathbf{r} / 2)
$$

The four equations of this sort give eight equations in all to be satisfied for the boundary conditions.

To satisfy these conditions, it proves to be necessary to build up $u_{g}$ as a sum of four independent functions, and similarly for $u_{u}$. Now all
the spherical harmonics can be written in such form that they have the same magnitude at each of the eight midpoints in question, differing only in sign. There are then just eight different types of function: (1) a function which has the same value at all eight points; (2) three functions in which there is a change of sign on reflection in one of the three coordinate planes; (3) three functions in which there is a change of sign on reflection in one of the three coordinate lines; (4) one function which changes sign on reflection through the origin. Any one of these types can be set up in an infinite number of ways from the infinite number of wave functions at our disposal. But we assume that, at any rate for states of low energy, we shall get a good approximation by taking in each case only the lowest state of the required symmetry. Thus we take in case (1) an $s$ function, in (2) the three $p$ functions with spherical harmonics $x / r, y / r, z / r$, in (3) the three $d$ functions with spherical harmonics $x y / r^{2}, y z / r^{2}, z x / r^{2}$, and in (4) the one $f$ function $x y z / r^{3}$. It is obvious that in building up our solution from these functions, we cannot hope to reproduce correctly the energy levels originating in any atomic level except an $s$ or $p$; for we omit entirely two of the five $d$, and six of the seven $f$, functions.

Let $R=r / 2$ be the distance from a nucleus to the midpoint where we satisfy boundary conditions. Let $s, p, d, f$ represent the values of the radial parts of the corresponding functions, at this distance $R$ from the nucleus. Then, if we put in the suitable values of $x / R$, etc., for the corresponding midpoints, we have

$$
\begin{aligned}
u_{g} & =A s+3(B x y+C y z+D z x) d / R^{2} \\
\left|u_{u}\right| & =3^{\frac{1}{2}}(E x+F y+G z) p / R+H(3)^{\frac{3}{2}} x y z f / R^{3}
\end{aligned}
$$

In particular, we have four pairs of midpoints. Let us compute the values of the $u$ 's in succession for the four points whose coordinates $x / R$, etc., are $3^{-\frac{1}{2}}$ times (111), $(11-1),(1-11),(1-1-1)$. We have, for instance for the second, $u_{g}=A s$ $+(B-C-D) d, \quad\left|u_{u}\right|=(E+F-G) p-H f$. For setting up the equations, we need also the quantities $\tan (\mathbf{k} \cdot \mathbf{R})$. Remembering that the components of $R$ are each $\pm R / 3^{\frac{1}{2}}$, the second of these quantities for instance is $\tan \left(\left(k_{x}+k_{y}\right.\right.$ $\left.\left.-k_{z}\right) R / 3^{\frac{1}{2}}\right)$. Let us name the four tangents, in the order stated above, $K, L, M, N$. Then our eight equations describing the boundary conditions become

$$
\begin{gathered}
K(A s+(B+C+D) d)-(E+F+G) p-H f=0 \\
L(A s+(B-C-D) d)-(E+F-G) p+H f=0 \\
M(A s+(-B-C+D) d)-(E-F+G) p+H f=0 \\
N(A s+(-B+C-D) d)-(E-F-G) p-H f=0 \\
\left(A s^{\prime}+(B+C+D) d^{\prime}\right) \\
+K\left((E+F+G) p^{\prime}+H f^{\prime}\right)=0 \\
\left(A s^{\prime}+(B-C-D) d^{\prime}\right) \\
+L\left((E+F-G) p^{\prime}-H f^{\prime}\right)=0 \\
\left(A s^{\prime}+(-B-C+D) d^{\prime}\right) \\
+M\left((E-F+G) p^{\prime}-H f^{\prime}\right)=0 \\
\left(A s^{\prime}+(-B+C-D) d^{\prime}\right) \\
+N\left((E-F-G) p^{\prime}+H f^{\prime}\right)=0
\end{gathered}
$$

These are eight simultaneous linear homogeneous equations for the eight coefficients $A \cdots H$, and as such they have no non-vanishing solutions unless the determinant of their coefficients is zero. This gives a single relation between the four quantities $K, L, M, N$. A second relation between these quantities arises from the fact that they
are derived from only three independent variables, the three components of $\mathbf{k}$. With these two relations between $K, L, M, N$, only two of them may be taken to be arbitrary. We may for instance assume arbitrarily two variables determining the direction of $\mathbf{k}$, so that for each energy, internuclear distance, and wave normal, the equations determine the magnitude of the electronic momentum or wave-length. The solution corresponds to a real wave, however, only if $K, L, M, N$ are all real. The regions of energy and internuclear distance where this condition is satisfied correspond to the allowed bands of energy, those where some of these quantities are complex correspond to the forbidden bands, in which only damped waves can be propagated.

## 2. Propagation in the $x y$ Plane

The equations above are too difficult to handle in the general case. However, if we limit ourselves to propagation in an arbitrary direction in the $x y$ plane, they simplify enough so that they can be solved, and at the same time the results are of sufficient generality to indicate the essentials of the general solution. In this case $k_{z}=0$, from which $K=L, M=N$. Now we set up eight new equations, of which the first is half the sum of the first two above, the second half the sum of the third and fourth above, and so on, and the fifth to eighth are half the corresponding differences. These equations are

$$
\begin{array}{r}
K(A s+B d)-(E+F) p=0 \\
M(A s-B d)-(E-F) p=0 \\
\left(A s^{\prime}+B d^{\prime}\right)+K(E+F) p^{\prime}=0 \\
\left(A s^{\prime}-B d^{\prime}\right)+M(E-F) p^{\prime}=0 \\
K(C+D) d-G p-H f=0 \\
M(-C+D) d-G p+H f=0 \\
(C+D) d^{\prime}+K\left(G p^{\prime}+H f^{\prime}\right)=0 \\
(-C+D) d^{\prime}+M\left(G p^{\prime}-H f^{\prime}\right)=0
\end{array}
$$

These equations break up into two groups of four each, the first containing the four variables $A, B, E, F$, and the second containing $C, D, G, H$. They can thus be solved separately. For all eight variables to be different from zero demands
that the determinant of coefficients of each of the two sets of equations be zero, giving one more condition than we had before, so that in general we cannot satisfy both conditions. As a result, we must satisfy our equations by letting all variables of one of the two sets be zero, and satisfying the determinantal equation connected with the other set. This divides our solutions into two independent sets: (I) in which $A, B, E, F$ are different from zero, the others equal to zero, and (II) in which $C, D, G, H$ only are different from zero. On looking back at the significance of the constants $A \cdots H$, we see that solutions of the first class are unchanged on reflection in the $x y$ plane, while those of the second class change sign on reflection. We now consider separately the two classes.

## Class I

Eliminating $A, B, E, F$, between the four equations, the determinantal equation by elementary algebraic manipulation can be put in the form

$$
\begin{aligned}
& {\left[K^{2}+\frac{1}{2}\left(p / p^{\prime}\right)\left(s^{\prime} / s+d^{\prime} / d\right)\right]} \\
& \quad \times\left[M^{2}+\frac{1}{2}\left(p / p^{\prime}\right)\left(s^{\prime} / s+d^{\prime} / d\right)\right] \\
& \quad=\left[\frac{1}{2}\left(p / p^{\prime}\right)\left(s^{\prime} / s-d^{\prime} / d\right)\right]^{2}
\end{aligned}
$$

This determines a single relation between $K^{2}$ and $M^{2}$, assuming that $s, s^{\prime}$, etc., are regarded as known, which they are if we first fix the energy and internuclear distance. Plotting $M^{2}$ against $K^{2}$, it is the equation of a rectangular hyperbola, with the 45 degree diagonal passing through the foci, and the asymptotes at $K^{2}, M^{2}=-\frac{1}{2}\left(p / p^{\prime}\right)$ $\times\left(s^{\prime} / s+d^{\prime} / d\right)$, the two intersections with the 45 degree line coming at $K^{2}=M^{2}=-\left(p / p^{\prime}\right)\left(s^{\prime} / s\right)$, $-\left(p / p^{\prime}\right)\left(d^{\prime} / d\right)$.

Let us next consider the other relation between $K$ and $M$. In the present case, by assuming propagation to take place in the $x y$ plane, we have already fixed one of the two variables determining the direction of the wave normal. Let the other be $\theta$, the angle between this direction and the $x$ axis. Then $k_{x}=k \cos \theta$, $k_{y}=k \sin \theta$, where $k$ is the magnitude of the vector, so that

$$
K=\tan \left(k_{x}+k_{y}\right) R / 3^{\frac{1}{2}}, \quad M=\tan \left(k_{x}-k_{y}\right) R / 3^{\frac{1}{2}},
$$

from which, eliminating $k$ and $R$ by easy trigonometrical manipulation, $\left(\tan ^{-1} M\right) /\left(\tan ^{-1} K\right)$ $=\tan \left(45^{\circ}-\theta\right)$. This equation between $K$ and $M$ must be solved simultaneously with the hyperbolic equation above, determining $K$ and $M$ as functions of $\theta$, and hence determining $k$, the magnitude of the momentum, as a function of direction of the wave normal. The solution is most easily carried out graphically, plotting both equations in the $K^{2}-M^{2}$ plane. The first equation, as we have seen, is that of a hyperbola. The second gives curves, proceeding out from the origin in a more or less radial manner. On account of the symmetry, we can get all possible types of propagation by merely considering angles $\theta$ between 0 and 45 degrees. For $\theta=0$, $M=K$, so that the curve is the 45 degree diagonal in the $K^{2}-M^{2}$ space. For $\theta=45$ degrees, $M=0$, so that the curve is the axis of abscissas. For intermediate angles, the curve lies between these values. Further, in order to have real propagation, $K$ and $M$ must be real, so that $K^{2}$ and $M^{2}$ are positive. Hence we consider only the intersections of the two curves in the first quadrant. With these descriptions of the curves, the following facts become evident. The equations can have no, one, or two real solutions. If both intersections of the hyperbola with the 45 degree diagonal come for negative $K^{2}$ and $M^{2}$, there are no real solutions. This is the case if $-\left(p / p^{\prime}\right)\left(s^{\prime} / s\right)$ and $-\left(p / p^{\prime}\right)\left(d^{\prime} / d\right)$ are both negative, or if $s^{\prime} / s$, $p^{\prime} / p$, and $d^{\prime} / d$ all have the same sign. If one of these conditions is satisfied, but not the other, there will be one real solution. In this case one branch of the hyperbola cuts the first quadrant. If it cuts both 45 degree line and axis of abscissa, there will be real propagation for all directions; this is true if the asymptotes come for negative $K^{2}$ and $M^{2}$, or if $-\frac{1}{2}\left(p / p^{\prime}\right)\left(s^{\prime} / s+d^{\prime} / d\right)$ is negative. In the contrary case, however, the hyperbola will cut the 45 degree line but not the axis of abscissas, so that there will be real propagation for a certain range of angles about $\theta=0$, but no real propagation in the neighborhood of 45 degrees. Finally if both intersections of the hyperbola with the diagonal are positive, there will be two real solutions. One of these, coming from the positive branch of the hyperbola, will necessarily have real propagation for a limited range of angles only, while the other, coming
from the negative branch, which must cut both axes, will have propagation for all directions.

## Class II

The discussion is throughout similar to that of class I. The determinantal equation is

$$
\begin{aligned}
& {\left[K^{2}+\frac{1}{2}\left(d^{\prime} / d\right)\left(p / p^{\prime}+f / f^{\prime}\right)\right]} \\
& \times\left[M^{2}+\frac{1}{2}\left(d^{\prime} / d\right)\left(p / p^{\prime}+f / f^{\prime}\right)\right] \\
& \quad=\left[\frac{1}{2}\left(d^{\prime} / d\right)\left(p / p^{\prime}-f / f^{\prime}\right)\right]^{2}
\end{aligned}
$$

The asymptotes come at $K^{2}, \quad M^{2}=-\frac{1}{2}\left(d^{\prime} / d\right)$ $\times\left(p / p^{\prime}+f / f^{\prime}\right)$, the intersections with the 45 degree line at $K^{2}=M^{2}=-\left(d^{\prime} / d\right)\left(f / f^{\prime}\right)$, $-\left(d^{\prime} / d\right)\left(p / p^{\prime}\right)$. If these values are substituted for the corresponding ones of class $I$, the same discussion applies in the present case.

For both cases, once $K$ and $M$ are determined as functions of $\theta$, the linear equations can be solved for the coefficients, $A, B, E, F$, or $C, D$, $G, H$, and the whole wave functions can be set up. Of course, only the ratios of coefficients are determined, the absolute values coming from the normalization condition. The solution is perfectly straightforward, and it is unnecessary to reproduce it here.

## 3. Results

With the same potential curve for sodium employed by Wigner and Seitz, calculations have been made for $s, p, d, f$ wave functions connected with a variety of energies. From these, we get $s^{\prime} / s, p^{\prime} / p$, etc., as functions of internuclear distance for each energy used. We can then construct Fig. 1, in which energies are plotted as functions of internuclear distance, and in which the curves are those on which $s^{\prime}=0, p^{\prime}=0$, etc., and $s=0, p=0$, etc. The lowest curve $s^{\prime}=0$ is that of Wigner and Seitz. These curves divide the space into different regions, within each of which the equations have a definite number of real solutions; for it is obvious that such a quantity as $-\left(p / p^{\prime}\right)\left(s^{\prime} / s\right)$, whose sign determines the number of solutions, will change sign when either $s^{\prime}, p^{\prime}, s$, or $p$ changes sign by going through zero. We see that for large distances, the allowed regions are narrow bands surrounding each energy level of the isolated atom, the bands broadening as the atoms come together, until they touch or overlap. As indicated in Fig. 1, it
is convenient to divide the solutions into their four types, two of class I, two of class II, in which $-\left(p / p^{\prime}\right)\left(s^{\prime} / s\right)$ is positive for class Ia, $-\left(p / p^{\prime}\right)\left(d^{\prime} / d\right)$ is positive for classes Ib and IIa, and $-\left(f / f^{\prime}\right)\left(d^{\prime} / d\right)$ is positive for class IIb, so that Ib and IIa always occur together in the same region. Then we note the following fact from Fig. 1: as the energy bands broaden, and attempt to overlap, two solutions of different type can overlap without interfering with each other; but two solutions of the same type not only do not overlap, but are forced entirely out of the overlapping part of the band. An example is seen in the overlapping of the $3 s$ and $3 p$ bands, in which among the three solutions of $3 p$, there is one, of type Ia, which has the same symmetry as the $3 s$ solution. This type is absent in the overlapping region, only the other two $3 p$ functions occurring in this region. Several other examples are evident in the figure. This effect is similar to that found in second order perturbation theory, in which terms of different symmetry do not perturb each other, but terms of the same symmetry repel. Another fact follows from Fig. 1, if it is extended to larger internuclear distances: while one state originates from each atomic $s$ level, three from each $p$ level, as we should suppose, only three originate from each $d$ level, instead of five, and only one from each $f$ level, instead of seven. This was to be expected, as we mentioned earlier, on account of using but eight functions to set up our approximate solution. It means, however, that above the lowest line $d^{\prime}=0$ we must expect to have missing states, not given by our approximation, so that we can place no great credence in any results of the method for higher energies. At the actual distance of separation, this means that but three states, the $3 s$ and two of the $3 p$, may be expected to be reliable. A next step in the improvement of the method would consist in using, not eight, but fourteen functions, satisfying conditions at the midpoints of lines to the eight nearest and six next nearest neighbors, or at the center of each face of the cell. This would be quite feasible, and should extend the correct results to decidedly higher energies.

The well-known gaps in energy appear at large distances in Fig. 1, as the gaps between energy levels arising from different atomic


Fig. 2. Two inner Brillouin zones, showing energy contours, approximately normal distance of separation.
states. For smaller distances they are not evident in the figure, the allowed states filling all energy levels without gaps in energy. The gaps come about if we consider only a particular direction of propagation; the allowed bands in this case would not fill up the whole regions in Fig. 1, and in general there would be gaps separating the regions. This situation is indicated in Fig. 2, in which we use a space with $k_{x}$ as abscissa, $k_{y}$ as ordinate, essentially a momentum space, and draw contours of constant energy. In this space, as Brillouin has shown (see Sommerfeld and Bethe), certain zones are formed, bounded by straight lines, and the gaps in energy come at the boundaries of the zones. In our case these are three dimensional zones, and in Fig. 2 we give the sections of them made by the plane $k_{z}=0$. They resemble, but are not identical with, the zones of two-dimensional lattices discussed by Brillouin, Sommerfeld and Bethe. Only two out of our three lowest zones, the $3 s \mathrm{Ia}$ and $3 p \mathrm{Ib}$, prove to be cut by the plane $k_{z}=0$, when plotted in this way. We shall not go into the method of identifying the polygonal parts of zones with the central zone, since this question is treated in the references mentioned.

In Fig. 2 the gaps of energy are plainly evident: along the 45 degree diagonal, there is no contour for energy -0.35 , the energy jumping from about -0.37 to -0.32 . If we look at the figure in a broader way, however, we see that the lines of constant energy are roughly circles. If the electrons were free, the lines would be exactly


Fig. 3. Energy as function of wave number, for wave in 110 direction, showing first two zones.
circular, the energy depending only on the magnitude of the momentum, not on its direction. From the resemblance of the curves to circles, we see that the free electron picture is not entirely incorrect. The comparison is better shown in Fig. 3, in which we plot energy as a function of the magnitude of $k$, for the 110 direction, or the 45 degree direction of Fig. 2. The free electron distribution would correspond to a parabolic curve, the kinetic energy being proportional to the square of the momentum. As we see, the actual curve agrees rather closely with the free electron parabola, which is drawn with the correct constants. In fact, for the lower half of the bottom zone, which alone is filled with electrons in the normal state of the metal, the agreement is practically perfect. This is an unexpected and significant result of the present calculations. It has been expected that the true curve would be represented by a parabola in its lower part, but it was generally supposed that the curvature of the parabola would be less than for free electrons. As a matter of fact, if we draw
similar curves for a distance of separation decidedly greater than the normal distance, say twice as great, we do find a decidedly smaller curvature, the gaps becoming much larger in proportion, and the curves for the occupied regions much flatter, so that these regions are narrower than for the case of the free electron distribution, as pointed out by Wigner and Seitz. But at, or even in the general neighborhood of, the equilibrium distance, the free electron energy is a good approximation, except in the immediate neighborhood of the gaps. Another point may be noticed from Fig. 3. The forbidden energy bands are closely connected with electron diffraction by the crystal; if electrons are thrown at the crystal in such a direction and with such energy that they would lie in a forbidden band within the metal, they are totally reflected, and form the Bragg reflected beams. The Bragg law of reflection, which is experimentally verified for the electrons, is correct only if the forbidden bands come just at the places predicted by the free electron curves. These energies and momenta lie at the intersection of the parbola with the vertical lines at the edges of the zones, in Fig. 3, and it is significant that they lie within the gaps as actually found. Thus our energy distribution would lead to Bragg reflections, not just at one definite angle, but through a range of angle, resulting in a broadening, but the angle predicted by the elementary law would lie within the actual broadened beam of directions.
The wave functions themselves show interesting form. In Fig. 4, we plot the wave function from the lowest zone, normal distance of separation, energy -0.60 , corresponding to a long wave-length. The wave is travelling in the 110 direction, and a line is drawn through a set of atoms in the 111 direction (the closest direction to the wave normal in which we can find a closely packed line of atoms). The real part of the wave function is plotted as a function of distance along that line, the nuclei being indicated. In the neighborhood of the nuclei, the function has the character of an atomic function, being like a $3 s$ function at the left side of the diagram, a $3 p$ at


Fig. 4. Electronic wave function, real part, for points on a line in 111 direction.
the right. In the imaginary part of the wave function, the rôle of $3 s$ and $3 p$ would be interchanged. But now we note that in between nuclei the wave function follows close to a sinusoidal curve, which is drawn for comparison. The accuracy of the agreement is another unexpected result of the present calculation. It has been checked by cutting through the same wave in other directions, and also by considering waves of other wave-lengths. It means that in all that part of the metal when the valence electron is not actually penetrating into the core of an atom, its wave function is much like that of a free electron, just as its energy is approximately that of a free electron. Here again, for distances of separation much larger than the normal distance, we no longer have agreement with the free electron picture. It is worth pointing out that the essentially new results of the present paper, namely the indication of the accuracy of the free electron picture both in energy and in the part of the wave function outside the atom (but not inside), and the detailed investigation of the nature of the energy gaps and zones (which can be carried much further than we have indicated) both depend on the possibility of actually solving the wave equation for an electron in a periodic field without important approximations. We could hardly expect them to follow with anything like the same certainty from a perturbation method which would be inaccurate at the actual internuclear separation.


[^0]:    ${ }^{1}$ Wigner and Seitz, Phys. Rev. 43, 804 (1933).
    ${ }^{2}$ Sommerfeld and Bethe, Handbuch der Physik, Vol. XXIV, 2nd Ed.

