REVIEWS OF

MODERN PHYSICS

VOLUME 6 OCTOBER, 1934 NUMBER 4

The Electronic Structure of Metals

J. C. SLATER, *Massachusetts Institute of Technology*

TABLE OF CONTENTS

I. Introduction

II. Fermi Statistics and Free Electrons

III. Fermi Statistics and Variable Potential Field

IV. The Wave-Mechanical Method: Wave Functions

V. The Wave-Mechanical Method: Conductivity and Energy

Appendix

I. INTRODUCTION

1. The general nature of metallic structure

Lorentz[1] provided us with our first usable electronic theory of metals. He assumed that a metal consisted of empty space containing atoms, acting like hard spheres. In the interstices between the atoms moved free electrons. These electrons were subject to forces of only two sorts: external applied electric fields, and forces of elastic collision with the atoms. By equilibrium with heat radiation, for example, the electrons take up the kinetic energy distribution characteristic of gas molecules at the same temperature. Now if an electric field acts on the metal, it accelerates the electrons, producing an electric current, consisting merely of a drift velocity of the charges. With constant acceleration, this current would increase linearly with time after the application of the field, except for the collisions of the electrons with the nuclei. It can be readily shown that these collisions, when averaged out, act like a resisting force proportional to the velocity, or current. Thus the current rapidly comes to a fixed value, proportional to the applied field. We can see this from a very simple equation. If $-e$ is the charge on an electron, v its velocity in the direction of the applied field, then $-ev$ is the current which it carries. Suppose the external field is E, so that its force on the electron is $-Ee$. Finally the resisting force of collisions is $-kv$. Then, writing force equal to mass times acceleration, we have $m\,dv/dt=-kv-eE$, or $m\,dv/dt+kv=-eE$. The solution of this equation is the sum of two parts: the steady drift, $kv=-eE$; and the transient, $v=\text{constant}\times\exp\,[-(k/m)t]$. The transient damps down so rapidly that it is not appreciable unless times as short as one period of optical light waves are important. We are left, then, with the steady drift, giving the current proportional to the applied field, or Ohm's law. If there are N electrons per unit volume, the current density is $-Nev$, so that we have $-Nev=(Ne^2/k)E$. Thus the quantity Ne^2/k is the specific conductivity of the metal. It is greater the more free electrons there are, but smaller the greater the resisting force on an electron. More detailed calculation of the mechanism of collision processes gave a definite value for k, in terms of the size of the atoms, etc. In Section 4, we shall see how a mechanism of collisions leads to an equivalent result.

The present theories of metals seem enormously complicated, in contrast with the beautiful simplicity of Lorentz's theory. This appearance is partly justified, but partly it is not. For Lorentz's theory, as we can see by a little consideration, is very far from being complete. The modern theories try to go much farther than it does, and therein lies some of their complication. Let us examine a few points in which, quite apart from the quantum theory, Lorentz's treatment needs amplification.

In the first place, it is obvious that with our present knowledge of atomic structure, we should never be content with a picture in which the atoms were treated as hard spheres, the electrons as colliding elastically with them. We must bring the collisions between electrons and atoms in metals into connection with what we know regarding the collisions between electrons and atoms in other places, for instance in gases. Secondly, we cannot be content with simply regarding the conduction electrons as constituting an electron gas, without further examination. The conduction electrons of a metal are altogether too much like the valence electrons of the individual metallic atoms, certainly executing some sort of orbits about the nuclei, as well as passing from atom to atom, and we shall not be satisfied until we see the connection between these two pictures. Then we cannot accept Lorentz's crude assumption about the forces acting on the electrons. Surely the electrons repel each other; and surely the atoms are really positive ions, and attract the electrons. These large electrostatic forces must be considered. Of course, the metal as a whole is electrically neutral, so that the forces from distant electrons and ions cancel each other, but still there are forces from nearby particles. In connection with this, we are led naturally to the question regarding the forces holding the metal together. A set of positive ions, immersed in a gas of negative electrons uniformly distributed, would be a system in which forces of attraction would act between the positive and negative charges, tending to pull them together, while the pressure of the electrons, growing larger as the metal was compressed and the density increased, would tend to keep them apart. This might well result in a stable solid of definite size and compressibility. These questions of metallic cohesion lead us to the idea that a metal is like a large molecule, which is held together by electrical or valence forces of one kind or another, and we should look into the exact relationship between metal and molecule. All these questions are things which occur to us, quite independently of the quantum theory, and the answering of them leads us into a theory much more involved, but at the same time much more satisfactory, than Lorentz's.

We have just spoken of some of the ways in which the general ideas of Lorentz's theory needed amplification. But in addition, that theory was based on classical mechanics, and one feature of quantum theory, namely, the exclusion principle, or the Fermi statistics,[2, 3, 6] has proved to be of great importance in the theory of metals. The fundamental papers of Sommerfeld[17] were devoted to applying this feature of the quantum theory to Lorentz's picture, showing that thereby certain difficulties of his theory, relating mostly to the specific heat, were removed. For the specific heat of the free electrons, according to Lorentz, should be the same as that of a corresponding molecular gas. Experimentally it was very much smaller. According to the Fermi statistics, where the electrons do not change their energy appreciably with temperature until we reach a much higher temperature than we are accustomed to, the specific heat should be small, as it is found to be. But aside from these advances, this first paper of Sommerfeld left the theory of metals much as Lorentz had treated it. Most of the more recent papers on metals have been devoted to filling in the many unsatisfactory gaps in this theory, which we have mentioned in the last paragraph, until at the present time it is possible to give a tolerably complete picture of the nature of metals, and of electrical conduction in them.*

Let us see what general picture of metallic structure we can build up. Metallic atoms consist of nuclei surrounded by a number of shells of electrons, of which the outermost is the shell of valence electrons, readily removed to form ions, and which are active in chemical processes. When the atoms come together to form a crystal, these outer shells of electrons of neighboring atoms overlap, to such an extent that the midpoint between adjacent atoms is at about the distance of maximum density of valence electrons for a single atom. This means that the valence electrons are profoundly affected by the formation of the solid, metallic state, though the inner electronic shells are relatively uninfluenced. There are two qualitatively important results of this interaction: a mobility of the electrons, resulting in electrical conduction; and energy relations, resulting in the binding or cohesion of the metal.

The mobility of the electrons is a result of the lowering of the potential barriers between atoms as they approach. The potential energy of one of the electrons of an isolated atom goes from a constant value, say zero, at infinite distance, decreasing according to the law $-e^2/r$, where e is the magnitude of the electronic charge, as the electron approaches the remainder of the atom, then decreasing more rapidly, until very near the nucleus it becomes negatively infinite according to the law $C-Ze^2/r$, where C is a constant, Z the atomic number. Such a potential energy is shown in Fig. 1a. The energies of the various types of electron may be shown as horizontal lines in this diagram, indicating really that the depth of such a line below zero energy measures the energy required to remove the corresponding electron from the atom. It is observed that even the highest of these energy levels is several volts below the asymptotic energy, measuring the ionization potential of the valence electron. If an electron moved with such an energy in such a potential field according to the classical mechanics, it could travel only in the region where the kinetic energy was positive, the region denoted by AB in Fig. 1a. In the wave mechanics, it is well known that the electron can penetrate into the region of negative kinetic energy, but that its wave function there is exponentially damped off to zero, instead of being sinusoidal as in a region of positive kinetic energy. Thus the wave function, and the probability of finding the electron, extend out beyond AB, but not far, with appreciable intensity, and the maximum

---

* For excellent resumes of this recent work, see references 61, 98, 116, 117. In some fields these go much further than the present paper.

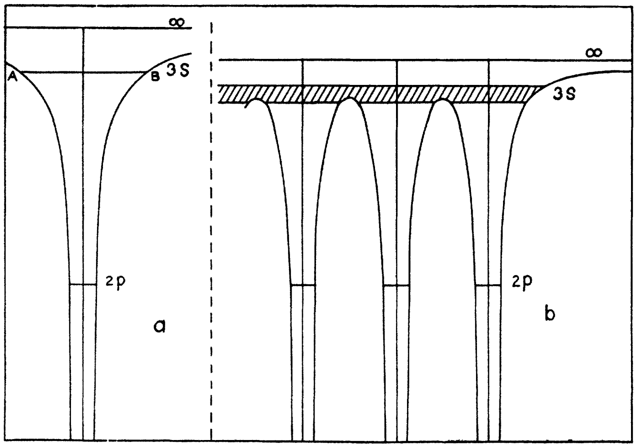

FIG. 1. a, Potential within a sodium atom. b, Potential within a sodium crystal. Potentials are so adjusted that x-ray levels agree, in which case potential at infinite separation is lower for metal than for atom, since image force is only half as great as Coulomb force at same distance. Shaded area in b indicates band of free electrons, derived from 3*s* of separated atoms.

density is necessarily within *AB*. The electron is confined to the atom by the potential wall around it, rising to a height of several volts.

Now if two atoms approach, and if they have incompleted outer shells of electrons, capable of taking in more electrons, as with metallic atoms, then it would be possible, if it were not for the barrier between them, for an electron to go over from the one atom to the other, executing an orbit about either nucleus. This is effectively prevented by the barrier, however, for even though it is possible on the quantum theory for an electron to penetrate a barrier, this would be very unlikely with such a high wall. As the atoms get close enough, however, the barrier begins to get lower, as well as narrower. Qualitatively we might say that this happens when two curves like Fig. 1a, get close enough so that at the midpoint the height of either curve is appreciably lower than the potential at infinity; though we shall see later that this simple view of the situation is not quite accurate. At any rate, it is clear that when the atoms are as close as in Fig. 1b, so that the regions in which the valence electron was likely to be found, in the two atoms, coalesce, the potential barrier between them will have disappeared. As this situation is approached, all obstacles to the free motion of electrons from one atom to the other will be removed, and the whole molecule will become essentially a conductor. In a similar way, as more and more atoms are added, the potential walls between them will disappear, and the whole interior of the metal will become a region without potential barriers. This will allow free motion of electricity from one part to another, leaving only the potential walls around the surface of the metal, now responsible for the work function, or the work required to remove an electron from the metal. This situation is indicated in Fig. 1b, where the potential is shown for the atoms in the neighborhood of the edge of a crystal.

Figs. 1a, b indicate at the same time another feature of the effect of the approach of the atoms on the energy levels: the single level of the valence electron in the isolated atom becomes broadened into a band of levels, of breadth of the order of magnitude of five or ten volts, in the metal. There are several ways to consider this. One of these makes use of the characteristic quantum relation, *Energy* $= h\nu$. As two atoms approach close enough so that electrons jump with appreciable frequency from one to the other, the energy levels of the electrons become split up, and the total splitting in energy, in magnitude, measures roughly the frequency of jumping from one to the other, by this quantum relation. Considered qualitatively, we may say that the stationary states of the separate atoms in this case are no longer exactly stationary; on account of the probability of jumping, there is a finite lifetime for the stationary state, and consequently a broadening or splitting of the energy level, as always with finite lifetimes. When the jumps become as frequent as they are in the metal, when practically every electron which strikes the barrier passes over it, the frequency of jumping becomes of the same order of magnitude as the frequency of motion of the electron in its orbit, and the broadening of the energy level is comparable with the whole electronic energy. With the molecule of a few atoms, the energy levels are split into a few components, while with the metal they are broadened into a continuum. This point of view concerning the broadening, though it is difficult to follow out in a quantitative way, is quite legitimate and very informing.

Quite a different point of view concerning the broadening of levels is obtained from the Fermi statistics. As the atoms are brought closer and closer together, the valence electrons are crowded into a smaller and smaller space. Now according to the Fermi statistics, or to the exclusion principle, the greater the density of an electron gas becomes, the greater must become its mean kinetic energy; in fact, the kinetic energy increases on compression essentially as in adiabatic compression of a gas. The actual density of valence electrons in a metallic crystal represents a considerable concentration, and when we compute the necessary spread of kinetic energy according to the Fermi statistics, we see that it requires electrons of kinetic energy all the way from zero to about ten volts, resulting in the band of energies which we have noted. This is the explanation given in Sommerfeld's quantum form of Lorentz's free electron theory. It is instructive to see that the same bands of electron energy levels which are thus explained as coming from free electrons are those which we have already followed through as being the broadened and distorted energy levels of the valence electrons of the atoms. In fact, there is no distinction which can be drawn between valence and conduction electrons: in the solid metal, the valence electrons acquire mobility and become able to travel from atom to atom, carrying current.

One difficulty in connection with what we have just said will occur to one at once: why is conductivity confined to metals? Our arguments seem so general that it is hard to see why they do not apply to any solid. The answer to this difficulty can be found by considering the energy levels of the valence or conduction electrons. In Fig. 1, we have followed through a single energy level from the isolated atom to the metal, in which it is broadened into a band. This level was occupied, both in atom and metal. But we could equally well have started with a higher, virtual level, not occupied in the normal state of the atom, but connected with an excited state. This would similarly have broadened into a band, lying higher than the one already considered. Each excited level goes into such a band, and the infinite number of excited levels forms a set of bands which, as the atoms move together, stretch to greater and greater positive energies. These bands are discussed in detail later in Section 14, and are shown in Fig. 9. Now it is plain that for large distances of separation these bands will be distinct, but for small distances they can coalesce into a single continuum of levels. Only a certain number of the levels, counting up from the bottom, will be filled with electrons, and the rest will correspond to virtual orbits of excited electrons. It turns out to be the case that in insulators, the filled levels form a band by themselves, separated by a finite interval from the next higher unfilled levels, while in conductors there are empty levels immediately adjacent to the filled ones. Thus in Fig. 1b, above the part of the *3s* band which is filled with electrons, shaded in the diagram, is an unfilled part of the band, which is not indicated. This has an immediate effect on the conductivity. For if we put the substance in the electric field, and if it is a conductor, the effect of the field will be to accelerate all the electrons. Those whose original direction was opposite to the field will be slowed down and decreased in kinetic energy, while those travelling in the direction of the external force will be speeded up, and their energy will be increased. This means that each electron will change its energy level, and the easiest way to consider this is as a gradual shift of electrons from their original energy levels through a continuous set of energy levels, either with increasing or decreasing energy. But now if the electrons originally filled a band of energy levels, there is no possibility for the electron of maximum energy to increase its energy further, as it would have to in conduction, for the next higher level is not adjacent, but requires a finite jump of energy, which the field cannot give it. Thus conduction is impossible.* The crystals in which the filled bands are separated from the unfilled ones in this way are in general those in which the electrons of the component atoms fill closed shells. The energy bands of the crystal into which these shells merge hold the same number of electrons as the atomic shells, and are likewise filled. On the other hand, if the atomic shells are only partly filled, the crystal bands are also only partly filled, leading to conductivity. It is thus plain why metals, with their partially filled atomic shells, are conductors, while ionic crystals, composed of ions formed of closed shells, are insulators. In other more complicated cases, a detailed study of the energy levels is required to determine the conductivity, and this in general has not been given up to the present.

We have now described the first of the two effects produced by the approach of metallic atoms to form a crystal: the mobility of the electrons, resulting in conductivity. The second effect is the effect on the energy, resulting in binding or cohesion between the atoms, holding them together to form a solid. This binding is similar though not identical to valence binding, but that does not suffice as an explanation, since one can interpret valence binding in more fundamental terms. The question is more difficult than that of conduction, for it depends essentially on the interaction of electrons, whereas conduction involves principally the action of separate electrons. Fundamentally, however, the binding results from electrostatic attractions between the positive ions and the negative electrons. As we mentioned at the very beginning, a set of positive ions immersed in a uniform negative volume charge would tend to be attracted into a compressed state, the attraction being balanced by the pressure of the condensed electron gas. In terms of our potential curves of Fig. 1, the potential energy function for an electron becomes decidedly modified as the atoms approach, the energy at any point in the outer part of the atom decreasing, since the electron finds itself in the metal under the attractive action of neighboring atoms as well as of the atom to which it is attached. This is a result of the penetration of one atom by another, according to which an electron of one atom is really partly inside the neighboring atoms, and therefore is not perfectly shielded from their nuclei by their outer electrons. This lowering of potential energy is connected with a corresponding lowering of the total energy, and consequently a binding. This effect is roughly equivalent to the "Coulomb interaction" which has been frequently discussed in valence binding. But there are also other complicating features. Thus as the atoms approach, the charge of the valence electrons redistributes itself, and the first effect is a tendency to concentrate between the atoms, just where the favorable effect of the change of potential is greatest, resulting therefore in further increased binding. This corresponds roughly to the "exchange interaction" of valence binding, though it is probably much less important in metals than in molecules. The whole situation is not as simple as this would indicate, however, for at the same time there are changes in kinetic as well as potential energy, and there are complicating features in the interaction between electrons. But the essential mechanism of interaction is the one we have described above.

II. FERMI STATISTICS AND FREE ELECTRONS

2. The Fermi statistics for metallic electrons

The simplest way of getting a fair approximation to the electrons in a metal is the use of the Fermi statistics, combined with the classical mechanics.[17, 61, 68] This is most easily formulated in a phase space of six dimensions, in which the three coordinates and three momenta of an electron are plotted, and each electron is represented by a point. We assume that each electron moves, not under the influence of its neighbors,

---

* For the application of these ideas to the conductivity of poor conductors, see references 59, 60.

but in a suitable averaged field of force which depends on position only. The simplest form of statistical mechanics is then applicable to the problem. Liouville's theorem applies to the swarm of points representing the various electrons, and tells us that they move so that the density of points is constant as we follow along the streaming points. Further, the motion of each point takes place on a surface of constant energy in the phase space. Thus by familiar arguments, the whole distribution will be independent of time if, and in the general case only if, the distribution is a function of energy only. That is, if $f(x, y, z, p_x, p_y, p_z)\, dx\, dy\, dz\, dp_x\, dp_y\, dp_z$ represents the number of electrons whose coordinates and momenta are in the element $dx \cdots dp_z$, and if actually $f=f(H)$, where H is the Hamiltonian function, a function of the coordinates and momenta $x \cdots p_z$, then f will be independent of time.

The different types of statistics correspond to different choices of the function f. The choice is not arbitrary, but is dictated by the physical principles underlying the problem. Actually in any physical case the various particles (as electrons in this case), will not be independent, but will rather be able to interact with each other. We neglect this interaction when we replace the mutual effects by an average field of force, and this is approximately justified, but not exactly. Actually, individual pairs of electrons sometimes collide sharply with each other, producing perturbations of the simple situation we have considered, and in particular allowing the colliding electrons to change their individual energies, though conserving the total energy. In such a collision, the representative points of the two electrons move from the energy surfaces they were on to new surfaces, a type of motion of representative points not contemplated in Liouville's theorem. And for real statistical equilibrium, the distribution must remain constant even though processes of this type occur. For particles colliding according to the classical mechanics, f must be given by the Maxwell-Boltzmann distribution, $f=c \exp(-H/kT)$, where k is Boltmann's constant, T the absolute temperature, and where the constant c is determined by the normalization condition $\int f\, dx \cdots dp_z = N$, N being the total number of particles. This is shown by the well-known method of the Boltzmann H-theorem. Particles obeying the exclusion principle, however, will give a distribution independent of time, only if f is given by the Fermi distribution, where, if we consider particles of one spin only,

$$f=\frac{1/h^3}{\exp\left[(H-H_0)/kT\right]+1},$$

where H_0 is determined by the normalization condition $\int f\, dx \cdots dp_z = N$. The power of the statistical method comes through the fact that the interactions of the electrons are largely taken into account merely by the choice of the statistical distribution function f. Once that is settled, we can go back to the picture of electrons moving in a potential field without influencing each other, and obtain results of a considerable degree of accuracy.

Before going further, it will be well to anticipate by pointing out the principal features in which the statistical method is inaccurate. First, even in classical mechanics, no problem in which there are mutual interactions of particles can be treated with real accuracy by a six-dimensional phase space. Such a space does not take into account the mutual relations, and in computing any such thing as mutual energy of interaction, it fails, for it does not give correctly such things as the probability that one particle will be at a definite distance from a neighbor. This point is discussed in more detail in Section 5. The space we have used is the μ space; to express our information completely, we need a $6N$ dimensional space, in which a whole system of N particles is represented by a single point, and we treat ensembles of such systems. In this Γ space, the interactions are correctly indicated. Secondly, in wave mechanics, a phase space cannot properly be used anyway. The principle of uncertainty states that the uncertainty in a coordinate, multiplied by the uncertainty in the corresponding momentum, is necessarily at least as great as h. If we multiply together the uncertainties $\Delta x \cdots \Delta p_z$ in all coordinates and momenta, we then have a minimum of h^3. That is, the uncertainty principle forbids us to look at the phase space in a way in which we can distinguish smaller detail than a volume h^3. It is not, then, a picture which can be carried through in great detail. Instead, we must use separate coordinate and momentum spaces as in Section 12.

The application of the uncertainty principle has important relations to the question of quantum states, and of individual electrons. The Sommerfeld quantum condition tells us that if the system is quantized, there will be one stationary state per volume h^3 of phase space. We see then that if we look at the problem from the standpoint of the phase space, we cannot resolve finely enough to distinguish individual stationary states from each other. No information depending on particular quantum numbers, electron shells, and the like, can be properly represented in the phase space; if we tried to do it, the information would become lost on transforming to the more correct wave mechanics. It is for this reason that the Fermi statistics makes no effort to consider separate stationary states, and that the Thomas-Fermi model of the atom contains no information about the x-ray shells of electrons. Similarly, the uncertainty principle prevents us from considering separate electrons, and the corresponding discontinuous density in μ space. The maximum possible value of the Fermi distribution function f equals $1/h^3$, corresponding to one electron per volume h^3 of phase space. If we looked at such a distribution function with a fine resolving power, the density would be highly discontinuous, being infinite where an electron was, zero in between. By virtue of the uncertainty principle, however, we are directed not to look at the density too closely. In fact, we are required to smooth it out by just the right amount to convert a discontinuous distribution into a continuous density function. The Fermi statistics thus does not consider individual electrons, but rather a smoothed-out distribution of charge.

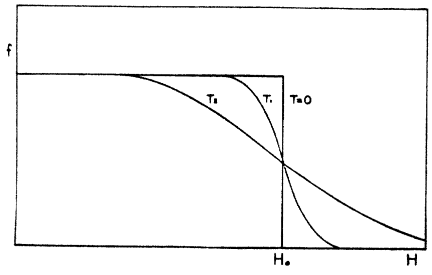

FIG. 2. Density distribution in Fermi statistics. f as function of energy H, for absolute zero, and two higher temperatures.

With these explanations, we may now proceed to use the Fermi statistics. In the Fermi function $f=(1/h^3)\{\exp[(H-H_0)/kT]+1\}^{-1}$, H_0 proves to be approximately independent of temperature, for low temperatures. Then when H is less than H_0, and $H-H_0$ is large numerically compared with kT, the exponential is very small, and f is practically equal to $1/h^3$, while when H is greater than H_0 by a corresponding amount, f becomes practically equal to $(1/h^3)\exp[-(H-H_0)/kT]$, much smaller than $1/h^3$, and following a law formally similar to the Maxwell-Boltzmann law. Thus in an energy range of the order of magnitude of kT the distribution falls from its maximum value practically to zero. The behavior of the function is shown in Fig. 2. As the temperature becomes lower and lower, the range within which the fall takes place becomes smaller, until finally at the absolute zero the function f is equal to $1/h^3$ when H is less than H_0, zero when H is greater than H_0.

3. The free electron model with Fermi statistics

The simplest model which can be made of a metal, and one which has been much used, is that in which the potential energy of an electron is assumed to be constant, so that H equals simply the kinetic energy, $(p_x^2+p_y^2+p_z^2)/2m$. In this case, the distribution function f is independent of the coordinates, depending on the momenta. We may then profitably consider the distribution of momentum, a function obtained by integrating f over coordinates: $\int f\,dx\,dy\,dz$ is a function giving the probability of finding an electron in unit volume of momentum space, independent of coordinates. Since f is in this case independent of coordinates, within the volume V of the container, though of course it is zero outside, this function simply is fV. It approximately equals V/h^3 if $p^2/2m$ is less than H_0, zero if it is greater than H_0. That is, in momentum space, all that region with the sphere $p_x^2+p_y^2+p_z^2=2mII_0$, of radius $(2mH_0)^{1/2}$, is uniformly filled to a density of V/h^3 with electrons, while the region outside the sphere is

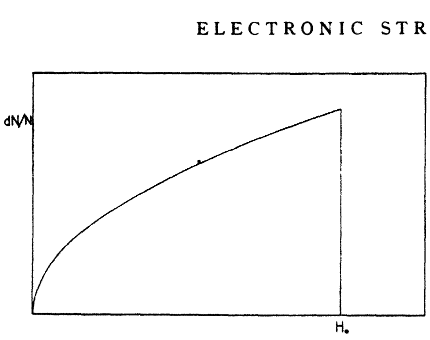

FIG. 3. Energy distribution for free electrons.

empty. At a temperature above absolute zero, the transition is not perfectly sharp, but is rounded off, the transition region occupying an energy band of the order of kT. If electrons of both spins are present, each sort can fill the space independently of the other, so that now the maximum density is $2V/h^3$.

It is easy to find the relation between the maximum energy H_0 and the number of electrons N, at the absolute zero of temperature. This relation is to be found from the equation $\int f V dp_x dp_y dp_z = N$. Now for electrons of one spin, fV is equal to V/h^3 within a sphere of volume $(4\pi/3)(2mH_0)^{3/2}$, and is zero outside. Hence we have $N=(4\pi V/3h^3)(2mH_0)^{3/2}$. This fundamental relation tells us that the maximum allowable density increases with the 3/2 power of the maximum energy H_0, or that the energy increases with the 2/3 power of the density. We can also find easily the distribution function in energy. For energies below H_0, the number of electrons whose energy is between H and $H+dH$ is equal to V/h^3 times the volume of momentum space between these limits, which is $(8\pi mH)dp = (8\pi mH)m\,dH/(2mH)^{1/2}$. Substituting, this becomes $dN/N=(3/2)(H/H_0)^{1/2}d(H/H_0)$, which is plotted in Fig. 3. From this distribution function we can immediately find the average value of functions of the energy; for instance, the average energy is

$$H_0\left\{\int_0^1 x^{3/2}dx\right\}\Big/\left\{\int_0^1 x^{1/2}dx\right\}=(3/5)H_0.$$

4. Conductivity in the free electron model

Having found our distribution, the two most interesting things to consider are the conductivity, and the total energy. For the conductivity, we can divide the discussion into two steps. First we ask how the electrons would be accelerated in the presence of an external electric field. Next we inquire what the effect of damping or collisions would be, and how these would lead to a constant drift velocity. The first question is easy to answer. The electrons are free, acted on by no forces except the external field. In the absence of a field, the coordinates of the electron increase linearly with the time, the momentum stays constant. Thus if we consider only a momentum space, plotting each electron by a point in this space irrespective of its coordinates, the representative point of an electron stays at a fixed position. The representative points of all the electrons, at absolute zero of temperature, will fill a sphere surrounding the origin of this momentum space, with uniform density, as we have just seen. Now if an external field is imposed, by Newton's second law, the external force will equal the time rate of change of momentum. Let us set up the vector $-eE$, where $-e$ is the charge on the electron, E the external electric field, assumed to be constant over the volume of the metal. This vector will then represent the time rate of change of momentum, or the velocity vector for the representative points in momentum space. All points will flow with uniform velocity in the direction of the external force, so that the spherical swarm of points will move bodily, without change of the relative distribution, in a given direction in momentum space. This will result in a continuously increasing current. For the total current is simply N times the mean current of a single electron, if N is the total number, or is $-Ne\mathbf{p}/m$, where $\mathbf{p}$ is the mean momentum of an electron, a vector. But this is simply the center of gravity of the swarm of points, so that its velocity is $-eE$, and its value, if the external field is applied at $t=0$, is $-eEt$. Hence the total current is $(Ne^2/m)Et$, a current proportional to the external field, increasing uniformly with time. This is the same current that would be found from any free electron picture, whatever sort of statistics were used.

The next problem in discussing conductivity is that of collisions. With really free electrons, there would be no collisions; they must be introduced as an arbitrary, extra influence. For the moment, let us adopt a very simple model, that in which an electron travels on an average for a time τ, then has a collision. At the collision, the momentum of the electron will change suddenly and discontinuously, so that its representative point in momentum space will jump to a quite different point of space. If the collision is completely elastic, and the center of force with which the electron collides is fixed or infinitely heavy, energy will be conserved, so that the point will jump from one point to another on a surface of constant energy, or a sphere whose center is the origin of momentum space. For example, if the collision is like that of two elastic spheres, one moving (the electron) and the other rigidly fixed (the atom), the distribution of scattered particles in angle is uniform, so that after collision the representative point is equally likely to be found on any part of the sphere; the mean momentum after collision is precisely zero. Hence on an average the electrons collect momentum for a time τ, then lose it, so that the average electron will have the momentum characteristic of the average time which has elapsed since the last collision, or $\tau/2$. We thus have for the average current the quantity $(Ne^2\tau/2m)E$, a constant times E, from which the conductivity is $Ne^2\tau/2m$. And the distribution in momentum space is again a sphere, but shifted along to have its center at $-eE\tau/2$.

The formula which we have just found is exactly the same that we should have found from classical theory. There are several features of it, however, which demand further consideration. In the first place, if the collisions were elastic, the electrons would continually gain energy. We can see this by a simple example of one-dimensional motion. If an electron at $t=0$ has a momentum p, and after time τ has a momentum $p+(-eE)\tau$, its kinetic energy will have increased by the amount $(p-eE\tau)^2/2m-p^2/2m=-peE\tau/m+e^2E^2\tau^2/2m$ in the interval. A corresponding electron with initial momentum $-p$ would have had its energy increased by $peE\tau/m+e^2E^2\tau^2/2m$. If the initial momentum p is large compared with the increment of momentum, as it is with the Fermi distribution, the first terms in the two cases are large numerically compared with the second, and we observe that they cancel, so that to a first approximation the changes of electronic energy add to zero. But the second order terms are all positive, and they will not cancel, but will result in a continuous absorption of energy. This absorption, for N electrons, would be of the amount $Ne^2E^2\tau^2/2m$, or just the product of the mean current $Ne^2E\tau/2m$ and the electric field E, per unit time. Thus this absorption of energy, with consequent speeding up of the electrons, is simply the Joulean heat, and this simple picture would indicate that all this heat would be lodged in kinetic energy of the electrons. Actually, however, the collisions are not elastic. The atoms which are hit by the electrons are of finite, not infinite, mass, and even a so-called elastic collision is one in which the electron loses a little energy, the atom gaining an equal amount. As a matter of fact, the electron loses almost precisely the extra, second-order, energy which it has picked up, on the average, so that the electrons gain practically none of the Joulean energy, the atoms receiving practically all of it. This follows from the known fact that the electrons have a specific heat which is very small compared with that of the atoms, so that if there is a certain amount of energy to distribute between atoms and electrons, it will practically all go to the atoms, if the situation is anything like thermal equilibrium. And normally in a metal there is approximate thermal equilibrium between electrons and atoms, though in certain exceptional cases this is not necessarily true.

The second special feature of the model of conduction which we must consider is that resulting from the exclusion principle. The collisions with atoms cannot obey exactly the classical laws, for according to those laws an electron after collision might jump to a point of momentum space which was already filled, contradicting the exclusion principle. The modification of the classical law which appears correct is to assume that an electron can have a collision according to classical laws if it jumps to an unoccupied point of momentum space, but that it has no collision at all if it would jump to an occupied point. More precisely, if f is the distribution function at the point of phase space to which the electron wishes to jump, $f_{\max}$ the maximum allowable value, we assume that the probability of jumping to this point is only $(1-f/f_{\max})$ as great as the classical theory would indicate. This, of course, is rather crude when examined in detail, for in some cases it can lead to electrons passing through spherical obstacles without interference, but our method of classical mechanics plus the Fermi statistics is crude anyway. This law of collision has a rather profound effect on the actual motions of the electrons, though not on the resulting conductivity. For consider an electron whose representative point before collision lies well inside the occupied sphere in momentum space. Then all points of the same energy are also occupied, so that no collision whatever is possible. The only electrons which can collide are those whose kinetic energy is very near the maximum allowable energy. This has rather odd results. An electron whose initial velocity is opposite to the accelerating force, but quite large, of the order of magnitude of the maximum velocity, will be slowed down by the external field, so that its representative point will move toward the center of the sphere, taking it out of the region where collision is possible. This will continue until the electron is brought to rest, speeded up in the other direction, and given practically the maximum allowable velocity in the other direction, before collision is possible. The collision will perhaps reverse the velocity without changing its magnitude, and the whole cycle begins again. Now if we look at the coordinates of the electron while its velocity is undergoing these changes, we see that the problem is similar to that of a ball being thrown up in the air, being brought to rest by gravity, and falling down to earth again with its initial velocity but opposite direction. The actual distance through the metal which the electron must travel in the process corresponds to the height at which the ball starts down, and this is the height at which its kinetic energy is wholly transformed into potential energy. In the electron's case, it moves against the potential gradient of the conductor for a distance in which the voltage changes by its initial energy, so that the retarding potential within the conductor stops it, then falls back again to its initial position, and finally has a collision. In the process, it may move several centimeters, since its energy is a number of volts, and in all this time it has no collision. But the essential thing is that as soon as it gets a large enough energy so that collisions are possible, the classical law begins to act, and it goes only a time τ further before colliding. The mean free time, or mean free path, of an electron, for computing conductivity, must be understood to mean the mean time or path for an electron which has already acquired enough energy so that the exclusion principle permits a collision. In spite of the peculiar type of motion of the individual electrons according to this theory, the calculation we have given of the current is still approximately correct. For the spherical distribution of points will still be shifted along by the current by an amount $-eE\,\tau/2$, approximately. Individual points will move uniformly through the sphere, in the direction of the external acceleration, until they strike the boundary of the original sphere. Then they will continue on the average about the time τ more, when they will have a collision which on the whole will shift them to the opposite side of the sphere, from which they will commence the same process over again. The result will be a displacement of the sphere.

For a more exact discussion of conductivity by this method, it is obvious that we must go into all the points we have mentioned in a careful analytical way. We shall not do this, for this report is concerned primarily with the electronic structure rather than the conductivity. But the general method is the following. Evidently in the presence of the external field, and with the collisions, a new stationary state will be set up, different from the original Fermi distribution. It will not be really stationary unless the Joulean heat is somehow removed from the metal, so that it does not continually heat up, but we assume this done, so that the atoms have a velocity distribution characteristic of a fixed temperature. Then the problem is to find a new distribution function f, similar to the Fermi distribution, but different from it. We cannot use thermodynamic methods for finding this distribution, as we could for the Fermi distribution, for the problem is not one of thermodynamic equilibrium, since energy is being transferred from the electric field to the electrons, from them to the atoms, and from them removed by some external process, in an irreversible manner. We must instead use a kinetic method, just as in Boltzmann's H-theorem, or as in Lorentz's discussion of the problem of conduction, first setting up the rate of change of any arbitrary distribution function f with time, then equating the time rate of change to zero for a steady state. As usual in such problems, we assume that the final distribution differs by a small quantity of the first order in E from the Fermi distribution, by using an expansion in powers of small quantities to make the solution practicable. And when we do this, using our modified classical law of collisions, we find a distribution which in fact is very similar to the Fermi distribution, but shifted along by an appropriate amount in the direction of the acceleration. The problem is decidedly complicated by the fact that there are three separate features to consider, all affecting the detailed distribution function in the region where it falls from its maximum density to zero: the temperature distribution, resulting in a gradual falling off of density in an energy range of the order of kT rather than a sudden edge; the energy gained by an electron during a free path; and the energy lost to an atom on a collision. These three energies are independent of each other, and result in decidedly complicated calculations, but still calculations which have been approximately carried through. The reader is referred to the original papers for the details, and is merely reminded that our simple discussion is in essence correct.[*]

5. Energy in the free electron model

The second point which we wish to consider for our simple model of the metal is the total energy. First we find the kinetic energy. This is very simple. Assuming the electrons to be at absolute zero, and all to have the same spin, the mean kinetic energy has been found to be $(3/5)H_0$, where $H_0=(3h^3/4\pi)^{2/3}(N/V)^{2/3}/2m$. This is of interest particularly for finding the pressure, which would be computed as for a perfect gas if we neglected potential energy, which for the moment we do. For a perfect gas, if P is the pressure, we have $PV=2/3\times$average kinetic energy. Hence for a given number of electrons, the pressure is inversely proportional to the 5/3 power of the volume, or $PV^{5/3}=$constant. This is the adiabatic law of compression for a perfect gas, as we should expect; for the type of compression which we give the electrons of the metal when we compress it is adiabatic, no heat being communicated to it, since we assume the process to occur at the absolute zero. The pressure of the gas is that which has been already mentioned in the first section, and it must be balanced by an attraction to produce equilibrium of the metal.

Before we can compute the potential energy, we must make a little more precise the nature of our model. Of course, electrons are charged, and a dense gas of electrons, with a high negative charge, would be very different from our picture of electrons in a force free space. We must rather assume that there is a fixed uniform distribution of positive electricity, the net amount per unit volume just cancelling the negative charge of the electrons. This positive charge would represent the nuclei of the atoms, spread out uniformly over the volume. We should say at first sight that the potential energy for such a model was zero. The distribution is exactly electrically neutral, so that each element of positive or negative charge would seem to be in no field. In accordance with electrostatics, the total potential energy equals the double integral $\frac{1}{2}\int\int\rho(x_1)\rho(x_2)/r_{12}\times dv_1\,dv_2$, where $\rho(x_1)$ is the charge density at x_1, $\rho(x_2)$ at x_2, r_{12} is the distance between these points, and the factor 1/2 takes care of the fact that each pair of volume elements, dv_1 and dv_2, is counted twice in the double integral. Surely if ρ is everywhere zero, the integral vanishes.

There is, however, an additional refinement which we can introduce, which seems perhaps rather artificial in this simple model, but which becomes very natural in the real metal. This refinement comes in when we remember that we are making a statistical calculation, by means of the Fermi statistics, rather than an electrostatic calculation of the ordinary sort, and that we are dealing with discrete electrons rather than infinitely subdivided ones, so that each electron is acted on by only the $N-1$ other electrons, since it exerts no forces on itself. The statistical

---

[*] Many of the references in the Bibliography deal with conduction, though mostly with some application of wave mechanics. References 17 and 25 are fundamental, but many of the others are important. The problems are thoroughly discussed in reference 98, and reference 117 gives an excellent and readable shorter discussion.

calculation applies only to the electrons, not to the nuclei, which we assume to be uniformly distributed, as we have just mentioned, so that we must treat the charge of the electrons and of the nuclei differently. Let us set $\rho=\rho_e+\rho_n$, the sum of the electronic and nuclear charge densities, so that the energy is

$$\tfrac{1}{2}\int\int\{\rho_e(x_1)\rho_e(x_2)+[\rho_e(x_1)\rho_n(x_2)+\rho_n(x_1)\rho_e(x_2)]+\rho_n(x_1)\rho_n(x_2)\}/r_{12}dv_1dv_2.$$

This consists of three types of term, the first coming from the interaction of electrons with each other, the second from interactions of electrons with nuclei, the third from interactions of nuclei with each other. The first is quadratic in the electronic charge density, the second linear, the third independent of it. When we regard this electronic charge density statistically, ρ_e becomes an average charge density. Then according to electrostatics, in the first term, the quantity $\rho_e(x_1)\rho_e(x_2)$ would be the product of the average charge densities at the two points x_1 and x_2, while according to statistics it will be the average of the product. As usual in statistics, the average of a product is not equal to the product of the averages, so that the first term is not the same as in electrostatics. The other terms are not affected by the statistical treatment, however, since they involve only linear expressions, and using the fact that the total potential energy, as computed by electrostatics, would be zero, and indicating averages by bars over the quantities, we can rewrite the potential energy as

$$\tfrac{1}{2}\int\int(\overline{\rho_e(x_1)\rho_e(x_2)}-\overline{\rho_e(x_1)}\,\overline{\rho_e(x_2)})/r_{12}dv_1dv_2,$$

or as simply the correction to the energy of interaction of pairs of electrons on account of the statistical treatment.

The average $\overline{\rho_e(x_1)\rho_e(x_2)}$ is something involving the statistical correlation between different electrons, and as indicated in Section 2, it can be computed properly only from the probability distribution in the Γ space. Thus let the coordinates of the N electrons be $x_1\cdots x_N$, and their momenta $p_1\cdots p_N$, and let the probability that simultaneously x_1 be in dv_1, x_2 in dv_2, etc., and p_1 in dp_1, etc., be $f(x_1\cdots p_N)\,dv_1\cdots dp_N$. The space in which $x_1\cdots p_N$ are plotted as variables is the Γ space, and f is the distribution function in that space. We are interested only in coordinates, not in momenta; thus let us integrate over momenta, obtaining a function which we can call $F_N(x_1\cdots x_N)$, whose product with the volume element $dv_1\cdots dv_N$ gives the probability that simultaneously electron number 1 be in dv_1, electron 2 in dv_2, etc. Now the potential energy of interaction between electrons is $\frac{1}{2}\sum_{ij}e^2/r_{ij}$, summed over all electrons i and j, for which $i\neq j$. This is to be averaged over the distribution function F_N. On account of the equivalence of electrons, the double sum will be just $N(N-1)$ times one term of it, or the energy equals $\frac{1}{2}N(N-1)\overline{e^2/r_{12}}$, the interaction energy of the pair 1,2, times the number $N(N-1)/2$ of pairs. The average of e^2/r_{12}, however, depends only on the first and second electrons, independent of the others. That is, to compute it, we need only the integral of F_N over the coordinates of all electrons but the first and second, which we may call $F_2(x_1x_2)=\int\cdots\int F_N(x_1\cdots x_N)dv_3\cdots dv_N$. Averaging over this distribution function, the total potential energy of interaction between pairs of electrons is $\frac{1}{2}N(N-1)\int\int F_2(x_1x_2)e^2/r_{12}\,dv_1\,dv_2$. This is what we previously denoted as $\frac{1}{2}\int\int\overline{\rho_e(x_1)\rho_e(x_2)}/r_{12}\,dv_1\,dv_2$. On the other hand, in finding the quantity $\frac{1}{2}\int\int\overline{\rho_e(x_1)}\,\overline{\rho_e(x_2)}/r_{12}\times dv_1\,dv_2$, we must use the average charge density, a function of one electron only. Again on account of the equivalence of electrons, this is $-Ne$ times the probability of finding electron 1 in dv, independent of all other electrons. This requires the probability function to be integrated still another time, giving $F_1(x_1)=\int F_2(x_1x_2)dv_2$. In terms of F_1, our second quantity is $\frac{1}{2}N^2\int\int F_1(x_1)F_1(x_2)e^2/r_{12}\,dv_1\,dv_2$.

In considering the total potential energy, it is convenient to rewrite it as

$$\tfrac{1}{2}\int\{-NeF_1(x_1)\}dv_1\int\{-(N-1)eF_2(x_1x_2)/F_1(x_1)+NeF_1(x_2)\}/r_{12}dv_2$$
$$=\tfrac{1}{2}\int\overline{\rho_e(x_1)}dv_1\int(\rho_e'(x_2)-\overline{\rho_e(x_2)})/r_{12}dv_2,$$

where $\rho_e'(x_2)=-(N-1)eF_2(x_1x_2)/F_1(x_1)$. The quantity $\rho_e'(x_2)$, really a function of x_1 as well as x_2, is the average charge density of electrons which would be found at x_2 if an electron were known to be at x_1, whereas $\overline{\rho_e(x_2)}$ is the average density at x_2, independent of whether there may be an electron in dv_1 or not. Since dv_1 is infinitesimal, the chance of finding an electron in it is also infinitesimal, so that $\overline{\rho_e(x_2)}$ is a charge density which integrates over all space excluding dv_1 to N electrons, while $\rho_e'(x_2)$ integrates only to $(N-1)$ electrons, since one of the electrons is definitely located at x_1. Thus, as we stated above, the true interaction energy between electrons involves the interaction of each electron with only $N-1$ others, and not with itself. The total potential energy may be regarded as the interaction of the electronic charge with that single electronic charge included in $\overline{\rho_e(x_2)}$ and not in $\rho_e'(x_2)$, with opposite sign, since it is to be subtracted from zero to give the final result. In a manner of speaking, it is the interaction of the electron with itself, which is incorrectly added into the total in electrostatics, but which we subtract again to give the correct energy.

The particular point of the present calculation is that the density ρ_e' is not so simple as it would seem. If we know that an electron is in dv_1, then we cannot assume that the other electrons are distributed over other volume elements dv_2 in an entirely haphazard way, to give a uniform volume distribution. On account of the exclusion principle, other electrons will be kept a certain distance away from the one in question. Thus, while ρ_e' will be independent of position so long as dv_2 is a considerable distance from dv_1, it will decrease as dv_2 approaches dv_1, presumably becoming zero as they come together. The problem is much like that met in x-ray crystal structure, where we are interested in the probability that, if one atom is in dv_1, another will be found in dv_2. If the two elements are far apart, the probability will be constant, but as they come together it fluctuates, there being some distances, as twice the atomic radius, where there are an abnormally large number of pairs, while with smaller distances the probability goes down to zero, since no pairs approach closer than a certain minimum. In our present case, we have not the material for making a really correct calculation, of ρ', though later, in Section 28, we shall use the wave mechanics to get an accurate treatment. We know two things about ρ', however: first, it is constant, equal to the mean density of electrons in the metal, $-Ne/V$, if r is large, but falls to zero if r approaches zero; secondly, its integral throughout the volume is $-(N-1)e$, corresponding to the $N-1$ other electrons. Putting these two facts together, we can make a very simple hypothesis which proves to be a fair first approximation. We assume that ρ' is zero for r_{12} less than a certain R, and is constant for r_{12} greater than R, where R is chosen so as to make the integral come out right. At large r_{12}, ρ' must equal $-Ne/V$, the average density. The integral of this over the volume V would be $-Ne$, one electron greater than the required integral $-(N-1)e$. Hence the sphere of radius R must be just large enough to include one electron: $4\pi R^3/3=V/N$. This determines ρ' completely. In terms of it, we see that $F_2(x_1x_2)=N/(N-1)\,V^2$ if r_{12}, the distance between $x_1y_1z_1$ and $x_2y_2z_2$, is greater than R, zero if r_{12} is less than R. On the other hand, $F_1(x_1)=1/V$.

In terms of the assumed ρ's, we can now compute the potential energy. We have $\overline{\rho_e(x_2)}=-Ne/V$, and $\rho_e'(x_2)=-Ne/V$ if r_{12} is greater than R, zero otherwise. Thus our integral

$$\tfrac{1}{2}\int\overline{\rho_e(x_1)}dv_1\int(\rho_e'(x_2)-\overline{\rho_e(x_2)})/r_{12}dv_2$$
$$=\tfrac{1}{2}\int(-Ne/V)dv_1\int(Ne/V)/r_{12}dv_2,$$

where the first integration is to be taken over the volume V, and the second over a sphere of radius r_{12} equal to R. Hence the integral is

$$-\tfrac{1}{2}(N^2e^2/V)\int_0^R 4\pi r_{12}^2/r_{12}dr_{12}=-\pi R^2N^2e^2/V.$$

Substituting for R, this is equal to

$$\tfrac{3}{4}(-Ne)e/R,\quad\text{or}\quad -\tfrac{3}{4}(4\pi/3V)^{1/3}e^2N^{4/3}.$$

We have now found the total potential energy, as well as the total kinetic energy, for electrons of one spin. If there are N_1 electrons of +spin, N_2 of −spin, where ordinarily $N_1=N_2$, the total energy will then be

$$\frac{3}{5}\left(\frac{3h^3}{4\pi V}\right)^{2/3}\frac{1}{2m}(N_1^{5/3}+N_2^{5/3})$$
$$-\frac{3}{4}\left(\frac{4\pi}{3V}\right)^{1/3}e^2(N_1^{4/3}+N_2^{4/3}).$$

This differs from the more correct expression derived from the wave mechanics* only in the numerical value of the constant multiplying the second term. The formula becomes more interesting if we take the case in which $N_1 = N_2$, and introduce the radius R as above by the equation $4\pi R^3/3 = V/N_1 = V/N_2$, so that R is the radius of a sphere containing two electrons, one of each spin. Then we can eliminate V. If further we express energies in terms of the Rydberg energy $2\pi^2 me^4/h^2$, and distances in terms of the atomic unit $a_0 = h^2/4\pi^2 me^2$, the energy becomes

$$\frac{9}{5}\pi\left(\frac{3}{4\pi}\right)^{1/3}\frac{(N_1+N_2)}{R^2}-\frac{3}{2}\frac{(N_1+N_2)}{R}$$
$$=3.50N/R^2-1.50N/R,$$

if $N = N_1 + N_2 =$ total number of electrons. In other words, since R will change proportionally to the linear dimensions of the metal if it is compressed, the kinetic energy increases inversely proportionally to the square of the linear dimension, the potential energy inversely as the distance, as we should expect from an electrostatic potential energy.

The model we have used is of course very crude. Nevertheless, it qualitatively agrees with the truth in the matter of the energy. The expression above shows a minimum of energy for a certain distance R, given by setting the derivative equal to zero: $0 = -7.00N/R^3 + 1.50N/R^2$, $R = 7.00/1.50 = 4.67$ atomic units. The resulting density of electrons is, in order of magnitude, in agreement with what we actually have in the alkali metals, if we consider only the free electrons; and it is for these metals and these electrons that we should expect to get the best agreement. The energy of binding per electron is here -0.16 atomic units, or about 2.2 electron-volts, also of the right order of magnitude. As we improve our model, however, we shall see that there are many features in which this calculation must be modified, so that it is largely accidental that there is even approximate agreement with experiment.

Before leaving this model, we should mention that the energy formula used above has been used to discuss ferromagnetism in a very crude way. In the formula in terms of N_1 and N_2, it is possible for N_1 and N_2 to be different. In this case, the electron spins would not cancel each other, and the metal would have a net magnetic moment. Such a spontaneous magnetization is supposed to occur in ferromagnetic bodies. Now if we consider only the kinetic energy, we see at once that the energy will be a minimum if $N_1 = N_2$, so that the body is nonmagnetic. On the other hand, the potential energy becomes negatively larger if all electrons have the same spin. Thus at small distances, where the kinetic energy is the preponderant term, the spin will be zero, but at large distances the potential energy is more important, and beyond a certain critical distance the lowest energy will be for the state where all electrons have their spin in one direction. The distance proves to be larger than those actually found in metals, so that this method does not lead to a satisfactory discussion of ferromagnetism, and it is mentioned only for the sake of completeness. But one feature connected with it persists in more correct treatments of ferromagnetism: the interaction terms in the potential energy give a lower energy for parallel spins, and if the spins can be set parallel without too much increase of the kinetic energy, the body will have a permanent magnetic moment. This is the fundamental origin of ferromagnetism.

III. Fermi Statistics and Variable Potential Field

6. The variable potential (Thomas-Fermi) model with Fermi statistics

The model of a metal we have been discussing, in which the electrons move in a field-free space, is obviously much too simplified to represent a real metal. We have spent so much time on it simply because it presents a number of features which must be present in a more elaborate theory as well. Actually, however, the field within a metal fluctuates violently, becoming very great in the neighborhood of an atomic nucleus, but reducing to zero in the regions between atoms. In the present section, we consider the actual nature of the field, and the treatment of electrons in such a field by the Fermi statistics. This approximation comes very much nearer the truth,

---

* See, for instance, reference 72. The units used by Brillouin in this reference do not agree with ours in all respects.

and still is not too difficult to handle. It leads to electrons which in the regions between atoms behave much as the free electrons we have just considered, while near the nuclei they behave like the electrons in the Thomas-Fermi* model of the atom. We still do not use wave mechanics, and get no information about discrete stationary states, but for the broad features of metallic structure these are not necessary anyway.

The first and most important question which we must answer is, what is the nature of the field in which the electrons move? We cannot answer this question finally until we discuss the wave mechanics, but we can give an explanation which includes all the physical features of the situation. Our first approximation is very simple. We take all charge, positive nuclei and distributed negative charge, and compute the electric field of this distribution. This field, and its resulting potential, are evidently not the correct ones to use for an electron. For the system as a whole is electrically neutral, and the potential is that of a distribution of charge whose total amount is zero. On the other hand, the field which acts on an electron is that of a system whose net charge is one unit of positive charge, a positive ion. Further corrections are the ones arising from the exclusion principle, mentioned in connection with the force-free problem; and effects of image forces, important at the surface of a metal, and of polarization. We shall see that these correction effects can be looked at in a rather fundamental way, and that they may change the field decidedly. Nevertheless, the effect on the field is ordinarily not great quantitatively, and for simplicity in presentation, we shall first assume that the electric potential of the whole charge distribution is the one to use. This leads to the Thomas-Fermi method.

Having once set up the potential, we can compute from it the charge distribution. According to the Fermi statistics, at the absolute zero of temperature, the density of electrons will be the maximum allowable value in the phase space at all points where the total energy H is less than H_0, zero where H is greater than H_0. Carrying out arguments similar to those used in discussing the force-free case, we see that at every point of coordinate space we can set up a momentum space, in which the density will be uniform up to a certain sphere, zero outside it. If U represents the potential energy of an electron at this particular point of space, $H_0 - U$ is the maximum allowable kinetic energy, so that the radius of the sphere is $[2m(H_0 - U)]^{1/2}$, and its volume $(4\pi/3)[2m(H_0 - U)]^{3/2}$. The number of electrons per unit volume of coordinate space and unit volume of momentum space is $2/h^3$, if we consider both spins. Hence the number per unit volume of coordinate space, integrating over momentum space, is $(8\pi/3h^3)[2m(H_0 - U)]^{3/2}$. As we see from Fig. 4, this gives a large density in the neighborhood of the nuclei, but a decreasing density in regions where U becomes more positive. At points where U is greater than H_0, so that classically the particle has negative kinetic energy and cannot move, we assume the density to be zero. We have, then, the charge distribution of the electrons.

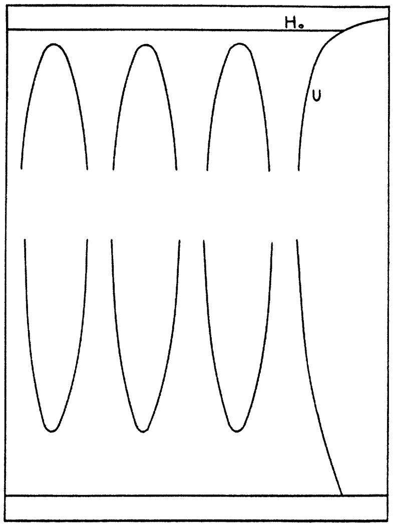

FIG. 4. Diagram for Thomas-Fermi method. Upper curve: potential energy U, maximum energy H_0. Lower curve: charge density of electrons.

Since we already know the positions of the nuclei, and have found the charge distribution of

---

* For the Thomas-Fermi model, see references 4, 5, 11, 19, 21, 41, 83, 84, 86, 102, 118. For some of the calculations in connection with the present section, the writer is indebted to Mr. H. Krutter, who will probably present them in a forthcoming paper.

electrons, we are in position to compute the electrostatic potential of this distribution of charge. We may now demand that this potential energy function should agree with the original function U. This is the requirement that the field be self-consistent, though not precisely in the sense of Hartree. It is a condition which completely determines the field, and from now on we assume that U is determined in accordance with it. Let us consider if an approximate solution for U is possible with real metals. For a single atom, our problem reduces to the Thomas-Fermi atomic model. The requirement of self-consistency then is known to lead to a differential equation in which variables can be separated in spherical coordinates. The solution depends only on the radius, and the resulting ordinary differential equation has been solved numerically. Hence this problem may be considered as solved. When we meet a metal, however, the problem does not have spherical symmetry, and leads to a partial differential equation, which at first sight seems hopeless to solve. Fortunately, however, there is a feature of the situation which we shall use constantly from now on, and which makes an approximate solution quite easy.

7. The cellular potential

We consider a regular crystal of the body- or face-centered cubic type. We imagine each atom to be the center of a polyhedral cell, so constructed that the cells of adjacent atoms touch, and taken together just fill up the space. The cells can be constructed by drawing the planes which are perpendicular bisectors of lines joining nearest or next-nearest neighboring atoms. These planes bound the cells. For the body-centered lattice, the cell is a fourteen-sided figure bounded by eight hexagons (facing toward the eight nearest neighbors) and six squares (facing toward the six next-nearest neighbors). For the face-centered lattice, the cell is a regular twelve-sided figure bounded by diamond-shaped faces, one pointing to each of the twelve equally near neighbors. The simplifying feature is that within one of these cells, the potential is very nearly spherically symmetrical.* Let us justify this statement.

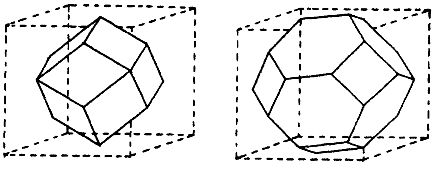

FIG. 5. Polyhedral cells for body centered and face centered cubic lattices. Dotted lines represent unit cells, containing two atoms for the body centered lattice, four for the face centered.

In the first place, as we see from Fig. 5, where the cells are plotted, they are not very different from spheres in their shape. This is natural, since the cubic lattices in question are the forms taken by close-packed spheres. In fact, if we packed the spheres closely, and imagined them flexible, like rubber balls, and then squeezed them so that no empty spaces were left between, they would squeeze into just the polyhedral cells we are considering. Now each cell, by symmetry, will contain just enough negative charge to balance the positive charge on the nucleus, so that it will be uncharged as a whole. If the charge were spherically symmetrical, it would have no field at exterior points. If it is almost spherically symmetrical, as we are stating it to be, it will have but a small field at outside points. Thus, at a point within one cell, the other cells will contribute but small fields. Further, they surround the cell in such a symmetrical way that even these small fields will tend to cancel. As a net result, the field at a point inside a cell is almost entirely produced by the charge within that cell itself. But this charge is arranged in general spherically about the nucleus, so that it produces a central field. Our arguments, then, form a closed circle: it is a consistent assumption to suppose that within each cell the field is almost central. It will be a good approximation to suppose that it is exactly central. This becomes particularly reasonable when we examine the actual form of the potential curves, which we later determine. For it then appears that the potential energy, going to infinity negatively as $-Ze^2/r+$constant, where Z is the atomic number, at the nucleus; rises to a value approximately constant rather close to the nucleus; and in the whole intermediate region between atoms is quite closely constant. If we assume a central

---

* This is pointed out particularly in reference 87, though in a different connection.

field of this character throughout the sphere inscribed in the cell, and let the potential be exactly constant between this sphere and the outer boundary of the cell, surely we shall have a good approximation, which at the same time will give a potential showing the right periodicity, and yet being continuous and with a continuous derivative everywhere.

The self-consistent problem then becomes a spherically symmetrical one within each cell, and leads to the same differential equation as for the isolated atom, and the same numerical method of solution. Only the boundary condition is different. With the Thomas-Fermi atom, the condition is that infinite space contain just the Z electrons necessary to balance the nuclear charge. With our problem, however, the Z electrons are crowded into the finite volume of the cell. As a slight variation of the statement in the last paragraph, let us take a sphere whose volume is the same as that of the cell, so that it is between the inscribed and circumscribed spheres, and let us assume a spherical Thomas-Fermi problem within this sphere. Our boundary condition is that the Z electrons be just contained in this sphere. The equation can be solved subject to this boundary condition, for arbitrary values of the radius of the sphere, and in particular for the observed value. It automatically gives a potential which has zero slope at the surface of the sphere, for the sphere encloses no net charge, and its field is zero at the boundary. As stated above, when we carry out the calculations for actual cases, the potential not only has zero slope at the surface of the sphere, but is nearly constant for a considerable range on both sides of it, so that we can join together the solutions in neighboring spheres smoothly. From now on, we shall assume the potential to be determined by this method, and shall use freely the results of the numerical calculations determining it.

8. The potential acting on an electron

It has already been pointed out that the field used in the Thomas-Fermi method is not the correct one, for it is the field of all electrons, not of all minus the one which is being acted upon. We wish instead the field at a point where an electron is known to be found, due to all the other electrons. As seen from Section 5, this is not given by $\int(\rho/r)dv$, the ordinary electrostatic formula, but by $\int(\rho_n+\rho_e')/r\,dv$, where $\rho_e'(x_2) = -(N-1)eF_2(x_1x_2)/F_1(x_1)$ is the density of the $N-1$ electrons left over when one electron is at x_1. Our problem, in discussing the field qualitatively, is to analyze the various ways in which ρ_e' differs from ρ_e, the whole average density of negative charge. In this analysis, we shall consider several effects neglected in Section 5.

A first illustration of the problem may be seen by considering the problem of a single atom. An electron of charge $-e$ finds itself in the field of the nucleus of charge Ze, and $(Z-1)$ other electrons. When the electron in question is far from the nucleus, quite outside the distribution of electronic charge of the other electrons, it is in the field not of a neutral atom, but of a positive ion, which to the first approximation is spherically symmetrical. This produces a field e/r^2, if r is the distance of the electron from the nucleus, with a corresponding potential e/r. The potential of an atom as computed by the ordinary Thomas-Fermi method, on the other hand, falls off in an exponential fashion with the distance. This potential should be corrected, so as to be approximately e/r rather than exponentially zero for large distances.

As the electron we are interested in comes somewhat closer to the ion, it polarizes it. The field of the electron at the ion is $-e/r^2$, so that if the polarizability of the ion is α, the induced dipole moment is $-\alpha e/r^2$. The field of this dipole back at the electron is $2\alpha e/r^5$, and its potential $\alpha e/2r^4$, just as truly a part of the potential acting on the electron as the Coulomb potential at large distances, though neither term is part of the potential of the total charge distribution. There is one difference, however. The polarizability of an ion or atom is a quantity which depends on the frequency of the external field which is polarizing it, on account of the ordinary dispersion effect. If the electron were standing still, or if instead it were a proton, which moves so slowly that it might almost be considered at rest, we should use the polarizability as determined from the static dielectric constant. But actual electrons are moving so rapidly that the electrons within the ion or atom do not have time to adjust themselves perfectly to its motion. In other words, we should really use the polarizability for high frequency, and for an undetermined frequency, since the external electron is not sinusoidal in its motion. The polarization term, though we can understand it in principle, and though we can fairly well approximate it, is thus not a perfectly definite thing. It depends on the electron's motion as well as its position, and not in a simple way which we can formulate directly.

Both effects of which we have spoken take place when the electron is well outside the charge distribution of the ion. When however it penetrates into this charge distribution, the situation is entirely altered. It is still in the field of the nucleus and $Z-1$ electrons, but the electrons deform their distribution entirely on account of its presence. The polarization is the beginning of this deformation, the first approximation to it for large distances; for the polarization arises from the action of the electron in pushing the cloud of other electrons away from it to the other side of the nucleus, producing thereby a dipole. But as the electron penetrates, instead of pushing the other electrons all to one side, it simply pushes them away from itself, in all directions. Thus when the electron gets very close to the nucleus, it acts like a K electron. By this time it has distorted the other electrons so that there is only one of the others close up, functioning as a K electron, while the outer shells are full of their normal complement of electrons, instead of lacking one from the outer shell, as in the original ion. In other words, by the time the electron has penetrated to the nucleus, it has cleared a hole for itself, by pushing one of the electrons which was already there out to the periphery of the atom. Further, this electron which it has pushed out is of the same spin as itself, so as to leave a hole into which it can fit, without contradicting the exclusion principle. It is this hole which we discussed in Section 5 in the force-free model. We approximated it by a spherical hole, from which all other charge of the same spin was removed, of just such a size that, with the allowed density of electrons, it would include one electron's charge. From this we could find the change in the energy of the electron, which with our crude method of calculation was $-(3/4) \times(4\pi/3)^{1/3}e^2(N/V)^{1/3}$, where N/V is the number of electrons of the same spin per unit volume. In our simple case, though this effect gave a contribution to the energy, it resulted in no force. In general, however, where the potential varies with position, the allowable density N/V is a function of position, so that if we extend our calculation of energy made with a constant potential to the case of variable potential, an approximation only partially justified, we shall find an energy term depending on position, and consequently a force resulting from this effect. Actually we shall use the potential rather than the force, just as in our applications we shall wish the potential rather than the force for the polarization and other effects.

In the last paragraph, we have really mixed up two effects, though they are closely related. An electron can keep other electrons away from it in two ways: by the exclusion principle, and by the repulsive forces coming from electrostatics. The first acts only between electrons of the same spin, and approximately the same momentum, the second between any pair of electrons. The electrostatic effect is the one responsible for polarization, since this can take place before there is any real penetration of the charge distribution. The exclusion effect is primarily responsible for the formation of the hole around a penetrating electron. But the electrostatic effect is present even here. It modifies the hole for electrons of the same spin, though only slightly. More important, it produces a somewhat similar effect on electrons of opposite spin. The accurate calculations on the structure of helium, almost the only case where the correlation between electrons of opposite spin has been very carefully studied, has shown that the probability of finding an electron of one spin near one of the opposite spin decreases, though not to zero, as the two come together. The fact that there is a finite probability of finding them in contact, with an infinitely high potential energy, is compensated by the fact that they have an infinitely great negative kinetic energy at this point, a possible thing in wave mechanics, though not in classical mechanics. Now to compensate this deficiency of charge of opposite spin in the neighborhood of an electron, there must be an excess at a somewhat larger distance, since the total charge of opposite spin must add up to

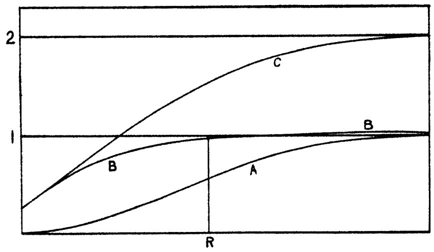

FIG. 6. Density of charge near an electron, plotted against internuclear distance. Curve *A* for another electron of same spin, *B* of opposite spin, *C* for both spins combined. One unit of density represents maximum allowable value for electron of one spin. Integrated deficiency of charge, for curves *A* and *C*, one electron unit; for *B*, zero.

the required amount, regardless of any displacement which it has. This is indicated in Fig. 6, where we plot in a schematic way the density of charge of the same and of opposite spin near an electron. For the electronic charge of the same spin, we use not the simple approximation so far employed, in which the density is zero at distances less than R, and the maximum value at larger values, but the value given by wave mechanics,* which will be described later, and which gives a smooth curve, but with the same total charge of one unit included in the hole. For the charge of opposite spin, we have drawn a schematic curve giving less than normal density at $r=0$, with larger density at intermediate distances, so arranged as to give the same total number of electrons as if there were no variation of density. It is seen that the result of this additional correction is to give a hole with more charge displaced than before, so that it will somewhat increase the magnitude of the correction to the energy, without changing its qualitative nature. In other words, it will change the constant factor in the formula, but since we know that it is not right, this effect will not result in any change in our conclusions. There is, however, one final remark to be made about Fig. 6. Suppose we find an electron at a distance r from a second electron, and ask what are the probabilities that it have the same spin, or the opposite spin. The answer is to be found from the relative heights of the two curves *A* and *B* in Fig. 6. At large enough distances, the probabilities will be equal. But at very small distances, and persisting for some considerable distance, the probabilities will be much in favor of the second electron having the opposite spin. This is the way in which, in the statistical method, we observe that pairing of electrons with opposite spin which is so important in the theory of atomic shells, and in the theory of homopolar valence. We observe further that the effect of an electron on determining the spins of its neighbors extends through only a relatively small sphere of action.

We have now investigated the difference between the ordinary average electric field, and the field at a point where an electron is known to be, both outside and inside an atom. Outside, the divergence comes because the electron, by its removal from the atom, left the atom positively charged; and further, because it polarizes the atom. Inside, the polarization effect becomes merged with the exclusion effect, by which an electron forces other electrons of the same spin and same momentum out of its range. We have found approximate formulas for these ionic and polarization effects on the energy, both outside and inside the atom. It is the author's belief that by making these corrections to the electrostatic potential outside and inside the atom, and by joining them smoothly in the intermediate region, it should be possible to set up the best approximation available for the potential to use in atomic problems. In particular, in some of its features, it should be better than Hartree's method,[7] which is often used, though for single atoms there is not a great difference. We shall anticipate the wave-mechanical treatment a little by mentioning the differences between this method and Hartree's. Hartree finds a wave function for each electron of the atom, such that the sum of the squares of all the wave functions gives the total charge density. As with our method, his first approximation to the electric field acting on an electron is just the total field of this charge density and the nucleus, which of course is better computed by his method than by the statistical method. But as with us, this would give the field of a neutral atom rather than of an ion. He corrects this, then, by sub-

---

* See reference 87, and a forthcoming paper of Wigner and Seitz (Phys. Rev. **46**, 509 (1934)) for further discussion of these questions. The writer is indebted to Dr. Seitz for the privilege of seeing his paper before its appearance.

tracting from the total charge distribution the charge connected with the wave function of the particular electron whose orbit he is finding. Thus if he is investigating a K electron, he subtracts off a K electron's charge density, while if he is investigating a valence electron he subtracts a valence electron's charge. The result is qualitatively similar to what we have used. But in general the hole from which we remove the charge is more concentrated than the wave function which Hartree uses. Thus the mean distance between the removed electron and the one in question is smaller for our method than for Hartree's, and the resulting negative potential energy contribution is larger, so that our model gives a stabler energy, and therefore by the minimum principle is nearer the truth. Similarly Hartree does not have any means of taking into account the polarization, and detects errors in his method for this reason; we should do this, again decreasing the energy. Our method is much more like that of Fock;[35, 48, 50] but even that does not take polarization into account.

The discussion of the field for an atom has been so long, only because similar problems come up in a more aggravated form when dealing with a metal. In the first place, Hartree's method when dealing with a metallic potential is quite impossible.* The wave functions of individual electrons in the metal will prove to be waves travelling throughout the metal, corresponding to equal probabilities of finding the electron on each of the N atoms of the metal. If now we corrected the total electric field by subtracting off the charge of the particular electron whose wave function we were computing, we should subtract $1/N$th of an electron from each atom, leaving a small and approximately uniform volume charge throughout the metal. This would leave a potential showing large-scale changes from point to point, and depending on the shape of the sample of metal, wholly different from what we should expect. On the other hand, when the electron is within the metal, our method suggests removing the charge from a hole closely surrounding the electron, just as we have discussed in the case of an atom. This will give a potential which has the same value at each atom of the metal, as it should. The hole in general proves to be of the same order of magnitude as an atom, so that the charge is largely removed from the atom on which the electron is located, leaving the other atoms electrically neutral, as we have assumed before. And the energy correction on account of the removal of the electron is large on account of its close concentration, resulting in an effect of binding on the metal as a whole, as we found with the force-free model. We assume, then, that within the metal we are to correct for the removed electron as described before, modifying the correction to take account of the displacement of electrons of opposite as well as of the same spin. And within the metal we are at every point in a region of approximately maximum allowable electron density, so that the polarization and ionic effects met outside atoms are not to be used. These come in only at the surface of the metal.

If the electron we are interested in is outside the metal, we must treat it differently, as with the atom. Here, however, the ionic effect has quite a different character from what it does in an atom. It becomes instead the image force, an effect which has created considerable misunderstanding, for like the other forces we have been considering, it is not part of the average field, and yet forms part of the field acting on the electron. Let us assume that a charge $-e$ is at a distance r from the free surface of the metal, where r is small compared with the dimensions of the sample of metal. The total charge e will be induced on the surface, consisting physically of the unbalanced positive charge left by the removal of the electron. We shall not have one atom carrying the whole positive charge, and the others electrically neutral, however. Rather, by electrostatics, the positive charge will on the average be distributed so that the resulting lines of force outside the metal are the same that would be produced by a point charge e at a distance r below the surface of the metal, the "image" of the charge $-e$. The field exerted by the induced charges at the electron is of magnitude $e/4r^2$, and its potential is $e/4r$.

The field of an electron outside a metal is the image field. A separate polarization field at smaller distances is not necessary in this case, for the image field takes account of the redis-

---

* This has been pointed out for example by Brillouin, as in reference 95.

tribution of the charge of the metal on account of the presence of the outer electron, and hence really takes the place of the polarization effect as well as the ionic effect. As the electron begins to penetrate the cloud of electrons of the metal, however, the formula for image force is no longer correct. Instead, the potential becomes that of the whole charge distribution corrected for the absence of the electrons in a hole surrounding the electron in question. As with the atom, these two types of potential curve must be joined at the surface. A careful discussion of this joining has not been given. Probably the details are not as simple as at first sight they would seem. For one thing, the atoms in the neighborhood of the surface are not in the symmetrical position of the inner ones, and it is hardly legitimate to suppose that the potential is centrally symmetrical within one of their cells. This question becomes particularly important in connection with the effect of surface films on metals. Another complication is that the image field is a one-dimensional field, depending only on the distance from the surface, while the field within the metal shows periodic variation from atom to atom. It would not seem easy to join these smoothly. Undoubtedly a periodicity extends out for some little distance, so that the image field is not so smooth as it seems. Nevertheless, this effect is probably small.

One qualitative feature of the joining of the field within the metal to the image field should be mentioned. The potential within a cell becomes constant at the edge of a cell. The image potential, however, continues to rise inversely proportionally to the distance. Thus the image potential will rise above the potential within the metal, and the potential at infinity will be decidedly higher than the maximum potential within the metal. This is shown in Fig. 1. In fact, it will rise several volts above the maximum energy H_0 of electrons. This difference, between the potential at infinity and H_0, represents the work function of the metal. We cannot expect any calculation based on the statistical method to give a very accurate calculation of it, but it is interesting that it gives any information about it at all. For a single atom, the Thomas-Fermi method without our corrections indicates that the maximum energy H_0 is the same as the potential at infinity, giving an ionization potential of zero.

9. Conductivity in the Thomas-Fermi model

We now have a picture of the potential within a metal, and of the electron distribution within it. Let us, as with the simpler model, ask what sort of information we can get about the two questions of conductivity and total energy. First we consider conductivity, and ask what the effect of an external electric field will be on the motion of the electrons. This is not the simple question that it was with the force-free model, for the actual motions of the electrons in the absence of an external field are now extremely complex. We are assuming that the electrons move according to classical mechanics, showing the presence of the quantum theory only in their statistics. Now an electron in a periodic field of the type we are considering can have two quite different sorts of motion. First, its energy can be less than the value corresponding to the midpoints between atoms. If we imagine the potential energy of an electron, $-e$ times the potential, plotted in the form of a relief map, there is a funnel-shaped depression of potential energy at each nucleus (using a two-dimensional form of statement), rising to a constant plateau, and then going over to the depression at the next nucleus. The electrons we are considering are those whose energy is not great enough for them to rise up to the top of the depression. Such electrons are then confined, in classical mechanics, to the particular depression in which they find themselves. They are real bound electrons. If now an external field is imposed, this will not, except for an excessively small number of electrons right at the top of the depression, make any difference in the status of the electrons; those which were bound without the field will still be bound. They cannot carry current from one atom to another. The field will merely change their orbits slightly, shifting the center of gravity of the orbit to one side or the other; it will polarize the bound electrons, as if they were in free atoms. These bound electrons, moving in a central field, will have orbits like those in Bohr's theory; only, since our Fermi model does not take account of discrete stationary states, they will be like Bohr orbits of a continuous set of quantum numbers. They will be periodic orbits, with precession, penetrating into the interior part of the atom, of the type so well known in the earlier developments of atomic theory.

The second type of electron is that whose energy is great enough so that it can get over the barriers between atoms. These electrons classically will have orbits which loop from one nucleus to another. In general, they will enter each cell with the same energy, but with a different angular momentum, except in very special cases, so that the loops around successive atoms will be of different shape. Their classical motion could be discussed, but as far as the writer knows it has not been. It involves some very interesting and fundamental questions of mechanics, questions which have a definite bearing on the problem of conduction. In particular, does an electron, started off in a given direction, retain any trace of its original momentum after encounter with many nuclei? It is obvious that this is a question of the greatest importance in treating resistance. The writer cannot pretend to answer this question, but certain features connected with it can be brought up. It is an interesting question, obviously, because if the electrons do not retain their original momentum, that means that the corresponding electric current becomes lost, or there is a resistance, while if the momentum persists after indefinite collisions, there will be no resistance. It is obvious that if the atoms are arranged at random, as in a liquid or alloy, there can be no question of persistence of momentum. But with a regular lattice it is not so obvious. And particularly is this the case because we shall find that in wave mechanics this persistence of momentum actually occurs for a regular lattice, so that such a lattice is resistanceless. The resistance arises entirely from the irregularity of the lattice caused by temperature vibration or by mechanical irregularity. And while it is generally supposed, without argument, that this is a peculiarity of wave mechanics, the writer is not convinced of this.

It is certainly true that for some regular lattices there can be groups of electrons which retain their momentum. Thus think of a cubic lattice made of material spheres, rather small in comparison with the distance between. The field of force is supposed to be the repulsive force of collision which particles going through this lattice would feel when hitting the rigid spheres. Now there will be certain open channels through such a lattice, in certain directions related to the crystal axes, the channels being larger as the spheres are smaller in proportion to the grating space. And particles directed through these channels, with an arbitrary speed, can travel clear through the lattice without collision. There are, further, certain modes of motion in which the particle collides successively with spheres on opposite sides of a channel, in just such angles as to keep moving in general along the channel. Of course, these types of motion are but a small proportion of all possible motions, and even these would not be possible in an arbitrary periodic potential, but they are enough to show that it is not entirely fanciful to suppose that momentum can be in some sense conserved.

Another and somewhat more physical way of looking at the problem is the following. Let us suppose a parallel beam of electrons, all of the same energy, incident on a certain cell of the crystal. Different electrons will be deflected by different amounts, depending on how near the nucleus they were aimed; those which were in between atoms will hardly be deflected, while those making a direct hit will be entirely changed in direction. In general, however, the beam after the one collision will retain a good deal of its original momentum, but some electrons will be definitely scattered out of the beam. But now, instead of an incident parallel beam, we could have had in addition certain incident electrons converging toward the nucleus. If these were properly chosen, we might have had just such incident electrons that they would be scattered into the original beam by collision, compensating for those scattered out. A state of affairs is conceivable in which the net effect of the scattering in the cell is zero: as many electrons are scattered into each direction as are scattered out if it. With such a situation in a cell, it is possible to set up a beam of electrons progressing through the crystal, individual electrons being scattered out of the beam, others being scattered in, in such a way as to produce equilibrium, and a beam which on the whole is unaffected as it goes through. Whether such a situation is possible mechanically is not known. It is possible in wave mechanics, and as a result a perfect lattice is resistanceless. Until the question is answered, we cannot say whether this is true in classical mechanics as well. But in any case, if the lattice is distorted by temperature vibration, it is certain that such streams of electrons will be damped out, there will be a certain mean free time during which the forward momentum will be entirely lost, and the metal will consequently have resistance. Only now, in contrast to our earlier elementary theory, the resistance, and the collisions which produce it, do not have to be introduced as extra hypotheses, but a complete carrying through of the mechanism would lead automatically to a calculation of resistance. For this reason, the problem is much more difficult than with the elementary theory. We can no longer separate the acceleration of the electrons under the action of the external field, and the collisions; for all the time the electrons are being accelerated they are also under the action of atomic nuclei, and are suffering deflections, indicated by motion in the momentum as well as coordinate space, which take the place of the sharp collisions. We shall not try further to analyze the problem of conductivity on this model, on account of the great difficulty. For fortunately the wave mechanics introduces simplifications which bring us back much nearer to the elementary field-free theory, and which permit a fairly simple treatment.

10. Free electrons and momentum distribution in the Thomas-Fermi model

One feature of conduction in the present model is worth while discussing, however: the question as to whether we can give any meaning to the energy distribution of the free electrons, as treated in the elementary theory, where the number of free electrons in a given energy range is proportional to the square root of the energy, and where there is a definite maximum of energy at absolute zero of temperature, with a rounding off according to the Fermi statistics at higher temperatures. This concept has been so useful in problems particularly of thermionic emission that one hopes that it has a significance aside from the special model on which it is based. Fortunately this seems to be the case. We have seen in the first place that there is in this theory a sharp distinction between free and bound electrons: those whose energy lies above the maximum of the potential energy in the crystal are free, those below are bound. The electrons just able to pass over these maxima, having zero kinetic energy at them, correspond in a sense to the lowest electrons of the Fermi distribution of the force-free theory. Thus the height of the maximum energy H_0 above these electrons measures the greatest kinetic energy of a free electron, in that theory. This maximum kinetic energy increases as the grating space decreases, with lowering of the height of the potential barrier. And the greater the maximum kinetic energy, the more free electrons there are per atom, so that the number of free electrons should be greater for metals with small grating space. Putting it the other way around, we have here an indication that the metals with few free electrons, the alkalies, should have large grating spaces, while those with many free electrons, as in the transition groups of the periodic table, should have small grating spaces, and at the same time large maximum kinetic energies for the free electrons. But when we look at the matter a little more closely, we seem to find a difficulty in this interpretation. As one of these free electrons moves from the region intermediate between atoms to the neighborhood of an atomic nucleus, its kinetic energy enormously increases. In the inner part of the orbit, there is no fundamental distinction between free and bound electrons; this distinction comes only in the electrons' ability to surmount the barrier between one atom and the next. We shall find when we consider wave mechanics that we are not permitted to do a number of things that are allowed with the present model. We cannot speak of the energy of a given electron; for the electrons exert forces on each other, and we have made an approximation by replacing these by an external potential field. And we cannot, on account of the uncertainty principle, inquire too closely about the speed of an electron when we limit its position too much. In particular, we cannot fix our attention only on the electrons in between the atoms, neglecting those which happen to be near the nuclei, and expect to get definite information about their energies. On account of these two facts we shall have to do two things: first, consider the kinetic energy of the electrons, which can be given a definite meaning, rather than their total energy; secondly, find the average behavior over electrons at all different parts of the atom, rather than picking out the electrons merely at the top of the potential barrier. We shall anticipate those difficulties by seeing how much information about the electrons we can get from our present model, if we restrict ourselves in this way.

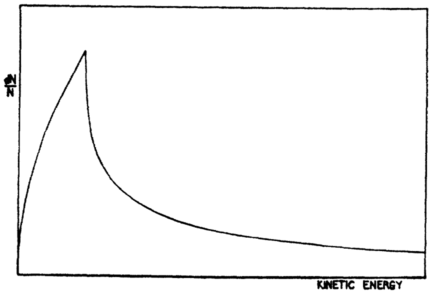

FIG. 7. Distribution in kinetic energy, Thomas-Fermi method.

First, we must consider only kinetic energy. Let us, then, consider the distribution of density in momentum space alone. Secondly, we must average over all coordinates. What we are allowed to use, then, is the integral over coordinate space of our distribution function f, which will then give an average distribution in momentum space. This is the converse of the process we have used before: integrating f over momentum space, to get the density in coordinate space. It is a very easy process: first we choose a certain magnitude of momentum, and compute from it the kinetic energy (we are not interested in the direction of the momentum, for our distribution in momentum space must be independent of direction). Then electrons of this kinetic energy will be found at all points of coordinate space for which the sum of this kinetic energy, and the potential energy, is less than H_0, and at no other points. At these allowed points, f equals $2/h^3$. Thus we simply multiply $2/h^3$ by the volume of coordinate space satisfying the condition just stated, and the result is the density in momentum space at the corresponding momentum. The result is shown in Fig. 7, and is easy to interpret. First, if the kinetic energy we are considering is less than the maximum kinetic energy of an electron at the potential barrier, all parts of the metal are available to it, and the integral is $2V/h^3$, a constant. This is the density at points in the momentum space within a sphere whose radius measures the maximum kinetic energy of free electrons. As the kinetic energy increases beyond this limit, however, the available volume very rapidly decreases, so that the density rapidly falls off, with a sharp break. The parts of momentum space corresponding to large kinetic energy, of course, give information only about the bound electrons. Though these cannot take part in conduction, still they appear in the density plot in momentum space, and the writer sees no way, consistent with wave mechanics, of leaving them out of account.

The sharp break in the distribution curve in momentum space, or in the corresponding one in kinetic energy (see Fig. 7), seems to be real, not merely a characteristic of our model, and it is this to which the theory of thermionic emission must have reference. If we now imagine the temperature to be elevated, so that we must use the Fermi statistics for a high temperature, a little consideration will show that the break will be rounded off in the familiar way. More electrons of high kinetic energy will appear, and these will not be like the ones already there, confined to the region of the nucleus, but will be free to travel about. They are the electrons which can escape from the metal. The fact that in the kinetic energy distribution they overlie certain other electrons of the same kinetic energy, however, directs our attention to the fact that this diagram, though it seems to be the only one permitted by wave mechanics, makes no clear distinction between free and bound electrons. The electrons within the sphere of discontinuity in the momentum space are not all free; many of them are bound electrons which, being at a turning point of their orbit, have momentarily a small kinetic energy. Similarly free electrons frequently have much larger kinetic energy than the "maximum" kinetic energy of a free electron, so that they are partly represented by the part of the momentum distribution outside the sphere.

But nevertheless the number of electrons within the sphere equals precisely the number of free electrons, for its radius is determined by the maximum energy of a free electron at the maximum of potential energy.

11. Energy in the Thomas-Fermi method

The total energy of our model of the metal behaves much more like the real metal than is the case with the simpler model. Thus as the atoms approach from infinite separation, instead of varying as inverse powers of the distance, the energy varies much more slowly, not differing appreciably from the sum of the energies of the separate atoms until they begin to come into contact with each other. If we compute the energy, and field in which the electrons move, without taking into account the difference between the average field and the field where an electron is, there is no minimum in the energy; it rises gradually, finally becoming infinite as the atoms are pushed into contact.* This increase is a result, as with the elementary model, of the increase of kinetic energy when the electrons are pushed into a smaller and smaller volume, with consequent speeding up on account of the Fermi statistics. This is partly counteracted by a decrease of potential energy. For as the metal is compressed, an average electron is nearer the nucleus of its atom, and hence in a region of lower potential energy. The decrease of potential energy, however, is not rapid enough to give an attraction. But when the correction is inserted corresponding to that made in Sections 5 and 8, where we assume the charge to be removed from a hole surrounding the position of an electron, with consequent decrease of potential energy, this correction appears to introduce a minimum into the curve, analogous to that found in Section 5 for the free electron model. Accurate calculations have not yet been made, but the minimum seems to be of the right order of magnitude to account for the cohesive properties of the metal. It is obvious, however, that the statistical method, which does not take account of electron shells and of the periodic properties of the elements, cannot be expected to give an accurate account of properties like grating space and heat of dissociation, which actually vary strikingly from element to element. All we can expect to get from it is a sort of average value, averaged over a group of elements of about the same atomic number but different chemical properties. Since calculations are not available, and since we shall treat the total energy much more carefully in connection with the wave-mechanical model, we postpone further discussion to that place.

IV. The Wave-Mechanical Method: Wave Functions

12. One-electron wave functions in wave mechanics

In the preceding sections we have discussed the treatment of a metal from the standpoint of the Fermi statistics, using the classical mechanics for the motion of electrons. Now we have reached the point where we must take account of the fact that the electrons really move according to the wave mechanics. But of course we cannot solve the problem of their motion exactly by wave mechanics. The most useful approximation to use in getting started is the same one we have used in connection with the Fermi statistics: we assume that each electron moves, not in the field of all other electrons, but in an averaged-out field. Then we can reduce the problem to a set of problems, one for each electron. This is essentially the method used in most approximate treatments of atomic structure, and it proves to be very useful and a good approximation there. We shall first ask, then, for the solutions for the motion of an electron in a periodic field of force. Having obtained these solutions, we can investigate further approximations.

The problem of determining the field to use is essentially the one which we have already discussed. We correct the ordinary electrostatic field of all positive and negative charges by corrections coming from the exclusion principle, image forces, and so on. For the present, we shall not consider this correction further, but shall come back to it in Section 28. It is essentially the same which we have already discussed in Sections 5 and 8. The only difference between

---

* Unpublished calculations of the writer and Mr. H. Krutter.

the present field and the previous one is that now, instead of determining the total negative charge density from the Fermi statistics, we find it by computing the wave function of each electron, squaring that wave function, assuming that to represent the charge density for that electron, and adding for all electrons. The sum, as one knows from a comparison of the Thomas-Fermi and the Hartree models of the atom, will agree rather closely with the charge density already considered. The final field, then, will be very similar to what we have discussed, differing only in unessential and minor details. One could use a self-consistent field method for determining it, in the following way. First, take the Fermi potential, corrected for exclusion, determined as described above. Find the wave functions of electrons in this field. Set up charge distribution and corrected field from this. Use that as a starting point for a second similar calculation. Repeat the process until the field was the same at the beginning and end of a step. This method, though it has not been carried out, would be practicable, by making only a few non-essential approximations. We have already pointed out in Section 8 the way in which it differs from the exact Hartree method of self-consistent fields. Briefly, the correction for exclusion is different, and is much more nearly correct in the present method than in Hartree's.

If the field is known, the next problem is to find the solution of Schrödinger's equation for an electron moving in this field. It is a problem which cannot be solved exactly, but for which we can get good approximate solutions, and derive many general properties of the solution. The problem is a purely mathematical one, but like most of the mathematical problems of quantum mechanics, it is so difficult that without a good deal of physical insight we cannot solve it. We shall go into it more carefully than we have the Fermi statistical method, for it is closer to the truth of the matter, and deserves more detailed treatment. There are three principal ways in which the results differ from those which we have already discussed. First, and rather unimportant, electrons now are capable of penetrating barriers, so that those with energy even slightly below the top of the potential barrier between atoms can go from one atom to the other; the distinction between free and bound electrons is not as definite as with the Fermi method. Secondly, the method of wave mechanics takes definite account of quantum numbers and discrete stationary states. Instead of continuous distributions of energies, we find discrete energies for the bound electrons, corresponding to the x-ray shells. And for the free electrons, though there are continuous distributions in energy, there are certain gaps, connected intimately with the quantum states. These gaps have connection with such things as electron diffraction, but even more they can have profound effect on conductivity, explaining in fact the difference between conductors and nonconductors. And the discrete quantum states naturally are the feature of the situation which explains the difference between different types of substances. There can be nothing in the statistical method explaining in detail why alkalies behave differently from alkaline earths, and why both are violently different from an inert gas or a halogen. These are characteristics of the periodic table, depending entirely on electron shells and discrete stationary states. The Fermi method can give only the vaguest sort of information about such things; we have passed over these problems tacitly in discussing that method. Thirdly, in discussing conduction, we have already mentioned that the wave mechanics gives decidedly simpler relations for accelerated electrons than classical mechanics appears to. We were unable to give a good discussion in classical mechanics, but the wave mechanics permits a number of rather accurate statements regarding conduction.

The technique of dealing with the wave mechanics is quite different from that of classical theory. We are not allowed to use a phase space, on account of the uncertainty principle, as we have already mentioned. Since this is impossible, we shall view the problem from two complementary aspects which together provide almost the same information: we shall consider the wave function in coordinate space, and in momentum space. The wave function in coordinate space is the ordinary function of Schrödinger theory, whose square gives the probability of finding the electron within given limits of coordinates. In the neighborhood of each nucleus it resembles the wave function of a bound electron rotating about that nucleus, while in the regions between nuclei it resembles the plane wave characteristic of a free electron. The wave function in momentum space is less familiar, though well known from the transformation theory of quantum mechanics.* It is a function of the three components of the momentum of an electron, such that its square gives the probability of finding an electron within given limits of momenta, irrespective of its coordinates. For considerations of conduction, the Fermi distribution, and such things, the momentum function is even more useful than the coordinate function, since it gives immediate information about the mean velocity and current of electrons, and the mean square velocity and kinetic energy. Of course, either of these functions can be found from the other, by a simple transformation, but sometimes we can get information from one more easily, sometimes from the other. The momentum wave function proves to be capable of giving more general results, but the coordinate function is more familiar and more easily visualized. Therefore we shall start our discussion with it, basing our work on certain general theorems which we shall prove later in discussing the momentum, and which we shall simply state at present.

13. Periodicity of potential and wave function, and Bloch's approximation

The potential field for our Schrödinger problem is a periodic one. The periodicity is best described by setting up the cells surrounding the nuclei which we have already mentioned, shown in Fig. 5, polyhedral cells all of the same shape and size, which just fill up the space. Then the potential has the same value at corresponding points of all cells. Let one of the nuclei be at the origin, and let the vector from this nucleus to the kth nucleus be Q_k. Then the potential is unchanged by making the translation Q_k, no matter what value k may have: $U(q)=U(q+Q_k)$. We now naturally ask, if the potential shows this simple property of periodicity, does not the wave function show some similar property? It does, as we shall be able to prove very easily when we have treated the momentum eigenfunctions, in Section 20. For the present, we give merely an intuitive discussion of it. It seems plain that in some fundamental way the wave function must have the same behavior in the neighborhood of each nucleus; it would be too much to suppose that it could be quite different near different nuclei, and yet satisfy Schrödinger's equation with the same potential in each case. We can readily imagine, however, that the function near the kth nucleus should be like that near the nucleus at the origin, but multiplied by some constant factor; for a solution of Schrödinger's equation can be multiplied by a factor and remain a solution. If in particular this factor is of the nature of $\exp(ic_k)$, where c_k is a real constant, then though the wave function is different at the kth nucleus from what it is at the origin, still the density of charge will be the same, for this depends on the wave function times its conjugate, and the product $\exp(ic_k)\exp(-ic_k)=1$. It is this case, where the density is periodic, though the wave function is not, which proves to be correct. Further, it can be proved that the exponential factor must be of the form $\exp[(2\pi i/h)p_0\cdot Q_k]$, where p_0 is an arbitrary vector, of the dimensions of a momentum, and which for the present we shall call a momentum, finding its interpretation later. In other words, the wave function satisfies the equation

$$u(q+Q_k)=\exp[(2\pi i/h)p_0\cdot Q_k]u(q).$$

This equation can be given a simple interpretation. If we were dealing with free electrons, the wave functions would be plane waves, like $\exp[(2\pi i/h)p_0\cdot q]$, where p_0 is a vector which proves to be the momentum of the electron, so that p_0/h is a vector of magnitude equal to the reciprocal of the wave-length, pointing along the wave normal. Then if we increased q by Q_k, the function would be multiplied by the factor $\exp[(2\pi i/h)p_0\cdot Q_k]$, just the same factor we have found. In other words, our function is one which, looked at in a broad way, behaves like a plane wave, but in more detail behaves like an atomic wave function. Within a cell, it can be very different from a plane wave, but if we compare only the values at corresponding points of different cells, these values will be related to each other just as in a plane wave.

---

* See for instance reference 40. For a specific application, to free atoms, see reference 26.

Bloch[25] has made an approximation to the wave function, satisfying the condition just stated, and further agreeing closely with the atomic wave function near the nucleus. If u_k is the wave function of an electron surrounding the kth nucleus if other atoms were absent, and Q_k is the radius vector from the origin to the kth nucleus, he assumes that the wave function connected with p_0 is

$$\sum_k \exp\left[(2\pi i/h)p_0\cdot Q_k\right]u_k.$$

If the atoms are far apart, so far that the wave function u_k of the kth atom has fallen practically to zero before one reaches one of the nearest neighbors, then we shall actually have this sum equal to the kth term near the kth atom, so that it will simply be a constant times the function for this atom itself. Further, it obviously changes by the proper factor in passing from one atom to another. If the functions of several neighboring atoms overlap, the correct periodicity will still be shown, for each term of the original summation must be multiplied by the same factor in going from one atom to another. But on account of the overlapping the wave function no longer will behave just like that of an isolated atom. This, however, is to be expected; surely as atoms approach the wave function must become distorted. The approximation to this distortion introduced by using our sum as a wave function is, however, a very poor approximation, and for this reason the method is only suggestive, not quantitatively correct.

In spite of the poor wave functions yielded by this method, it still is possible to obtain an approximation to the energy, for by the perturbation theory the mean of the energy over even a poor wave function is correct to one more order of small quantities than the wave function. Thus let H be the energy operator for the one-electron problem. Then the mean energy is given by

$$\frac{\int \sum_m \exp\left[-(2\pi i/h)p_0\cdot Q_m\right]u_m H \sum_n \exp\left[(2\pi i/h)p_0\cdot Q_n\right]u_n dv}{\int \sum_m \exp\left[-(2\pi i/h)p_0\cdot Q_m\right]u_m \sum_n \exp\left[(2\pi i/h)p_0\cdot Q_n\right]u_n dv} = \frac{\sum_{m,\,n} \exp\left[(2\pi i/h)p_0\cdot (Q_n-Q_m)\right]H_{mn}}{\sum_{m,\,n} \exp\left[(2\pi i/h)p_0\cdot (Q_n-Q_m)\right]\delta_{mn}},$$

where $H_{mn}=H_{nm}=\int u_m H u_n\, dv$, and $\delta_{mn}=\delta_{nm}=\int u_m u_n\, dv$. In the nondiagonal terms, we notice that the two terms corresponding to m, n and n, m combine to give a cosine multiplied by 2. Further, each atom of the crystal is identical with every other, so that the double sum is just the single sum, multiplied by the number of atoms in the crystal. Finally we assume that the atoms are far apart so that u_m and u_n overlap only slightly, the only case in which the method of treatment is justified anyway. In this case, we can approximately set $\delta_{mm}=1$, $\delta_{mn}=0$ if $m\neq n$. Hence the expression for energy may be rewritten

$$H_{mm}+\sum_{n\neq m} 2\cos\left[(2\pi/h)p_0\cdot Q_{mn}\right]H_{mn}.$$

Now we must consider the quantities H_{mn}. If U is the actual potential energy, with identical singularities at each nucleus, and U_n is the potential energy for the nth atom alone, we have by hypothesis $\left[-(h^2/8\pi^2 m)\nabla^2+U_n\right]u_n=\epsilon u_n$, where ϵ is the one-electron energy for the atomic problem, equal approximately to the binding energy of the corresponding electron. Thus

$$Hu_n=\left[-(h^2/8\pi^2 m)\nabla^2+U\right]u_n=(\epsilon+U-U_n)u_n,$$

so that $H_{mn}=\epsilon_{mn}+H'_{mn}$, where $H'_{mn}=\int u_m \times(U-U_n)u_n\, dv$. As seen from Fig. 1b, U is closely equal to U_n, as shown in Fig. 1a, within the nth atom, but becomes less than U_n outside, becoming negatively infinite at each other nucleus, where U is negatively infinite, U_n finite. Thus H'_{mn} will in general be negative if u_m and u_n are of the same sign where they overlap most strongly, positive otherwise. Remembering now that $\epsilon_{mn}=\epsilon$ if $m=n$, $=0$ if $m\neq n$, the energy becomes

$$\epsilon+H'_{mm}+\sum_{n\neq m} 2\cos\left[(2\pi/h)p_0\cdot Q_{mn}\right]H'_{mn}.$$

As a result of this equation, each atomic energy level is broadened into a continuum, the various levels being given by different values of the vector p_0. The maximum and minimum energies of such a band come when as many as possible of the cosine factors are positive or negative, respectively, and the breadth depends on the size of the matrix components H'_{mn}. Thus for atomic wave functions far down in the atom, for which the overlapping of functions from different atoms is very small, these components will also be small, and the band will be only slightly broadened, while for the outer electrons of the atom the overlapping will be large, and the broadening can well be comparable with the whole separation between energy levels. It is in such cases that the method breaks down, as we can see from several arguments. In the first place, we have already seen that for large overlapping, there is no reason to suppose that Bloch's method would give accurate wave functions in the region between atoms. Next, as the energy bands coming from different atomic configurations began to overlap, which they would do in this case, we should find that the energy had appreciable matrix components between levels of different atomic quantum number, so that a proper perturbation calculation would take into account the interaction between different atomic levels, and the final wave functions would be mixtures of different ones of the sort we have set up. Such a calculation has been made in special cases,[98, 109] but has not been carried out in general. The method is primarily one suited to the more tightly bound electrons of the atom.

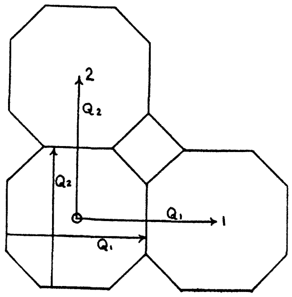

FIG. 8. Perpendicular to face of polyhedral cell. Illustrated with section of body centered cell taken through nuclei. Vector Q_1 from nucleus 0 to nucleus 1 equals vector from one face of cell 0 to opposite face.

14. The cellular method and metallic energy levels

For cases where Bloch's method is inapplicable, as well as the ones where it is applied, another method of approximation is more suitable, depending on the solution of the wave equation within the cell surrounding an atom, subject to suitable boundary conditions.[87, 110] If we knew precisely the wave function in one cell of coordinate space, and knew the fundamental momentum p_0, we could get the wave function in any other cell by multiplication by the suitable factor $\exp\,[(2\pi i/h)p_0\cdot Q_k]$, and shifting along to the required cell. All we need, then, is the wave function in a single cell. This wave function must satisfy two conditions: first, it must satisfy Schrödinger's equation; secondly, it must join on smoothly to the wave function in adjacent cells. This latter condition can be put in a form depending only on the one cell. Consider a nucleus, and a nucleus in an adjacent cell, letting the vector from one to the other be Q_k. It is plain that this is also the perpendicular vector from one face of the cell to the opposite face, as shown in Fig. 8. From a point of the surface of the cell, to the opposite point, the wave function must be multiplied by $\exp\,[(2\pi i/h)p_0\cdot Q_k]$. Thus the boundary condition is simply that the wave function, and its gradient, be multiplied by this factor on going from one point of the surface of the cell to the perpendicularly opposite point. There is, of course, a different factor for each pair of faces of the cell. The problem is to solve Schrödinger's equation in the cell subject to these boundary conditions. This problem can be discussed, and approximate solutions obtained. We have seen that it is a good approximation to suppose that the potential in a cell is spherically symmetrical. Then Schrödinger's equation is a spherical wave equation, and variables can be separated in it in polar coordinates, as with an isolated atom. Separate solutions are then of the form of spherical harmonics of angle, multiplied by functions of the radius which must be determined by numerical solution of a radial differential equation. For an arbitrary energy value, we have an infinite number of solutions, one for each spherical harmonic of angle, all satisfying the condition that they remain finite at the origin, which of course we must demand. In general these functions will become infinite at infinity, but this does not concern us, since we use the function only within the cell. Now by making an arbitrary linear combination of all these functions, with different spherical harmonics, or different angular momentum or azimuthal quantum numbers, we obtain an infinite number of arbitrary constants, with which it seems possible to satisfy the required boundary conditions. A general discussion has not been given; but it can easily be shown possible to satisfy them at a finite number of points of the surface, namely, at the midpoints of the faces of the cells, and a practicable method of doing this can be set up. The resulting wave functions have been computed for certain cases, in particular for sodium, as a function of energy and of internuclear separation. For the tightly bound electrons, one term in the linear combination is large compared with the others, so that the function resembles an atomic function of definite quantum number. This is not the case where energy bands overlap, however, corresponding in Bloch's method to the necessity of making a perturbation calculation in such cases, with combinations of several atomic states. The combinations, however, are here made exactly, rather than approximately as in perturbation methods.

The actual calculations are made in the following order. First, one decides on the energy and the internuclear distance, or the size of cell, determining the Q's. For this energy, one integrates the central field differential equation numerically, for each value of azimuthal quantum number. One determines the value of function and slope for each of these curves at the distance midway between adjacent atoms, the edge of the cell. There are then certain equations connecting the various functions and slopes with the value p_0. There are enough such equations so that the direction of the vector p_0 may be assigned at pleasure, and the equations then determine its magnitude. It is possible for p_0, however, to be real or imaginary as determined from these equations, and it is only the real values which can be used. Imaginary values would lead to damped waves, which are not allowable. Hence there are only certain ranges of energy, internuclear distance, and wave normal for which real propagation is possible.

In Fig. 9 we plot energy against internuclear distance, for sodium, as determined by the cellular method, and show those ranges of the diagram for which real propagation is possible in some direction. The ranges in which propagation is possible in all directions are somewhat more restricted. This diagram shows very clearly the relation between atomic and metallic energy levels. At infinite separation, the levels of negative energy are discrete, the energy levels of the atom, and each can be assigned to a set of quantum numbers. As the distance of separation decreases, each of these discrete stationary states spreads out into a continuum, as Bloch's approximation would indicate. The spreading becomes large for any particular state about at the point where the orbits of adjacent atoms commence to overlap. This is indicated by the dotted line AA, intersecting each energy level at the radius of the corresponding classical orbit; it is seen that to the right of this line, the spreading is comparatively small, while to the left, it is large. We observe, then, that at the actual distance of separation in the crystal, in this case, the $1s$, $2s$, $2p$ levels have spread very little, while the $3s$ has spread a good deal, so as to overlap the $3p$ and $3d$. Farther above and to the left of the dotted line, the various groups of levels overlap and interact with each other in a very complicated way. We shall discuss the detailed nature of the energy curves in much more detail later, particularly in Section 19.

15. Periodicity of energy and wave function in momentum space

The different sublevels of the continuum into which one of the atomic levels spreads are characterized by the values of p_0, which we have called the momentum. We can get a more detailed picture of the energy levels by choosing a given internuclear distance, and a given energy band, and within this band plotting energy as a function of p_0. Since p_0 is a vector, this means really a plot of energy in a three-dimensional space, in which the three components of p_0 are the coordinates. This we may call a momentum space, postponing further discussion of it until later. Now when we plot energy in the momentum

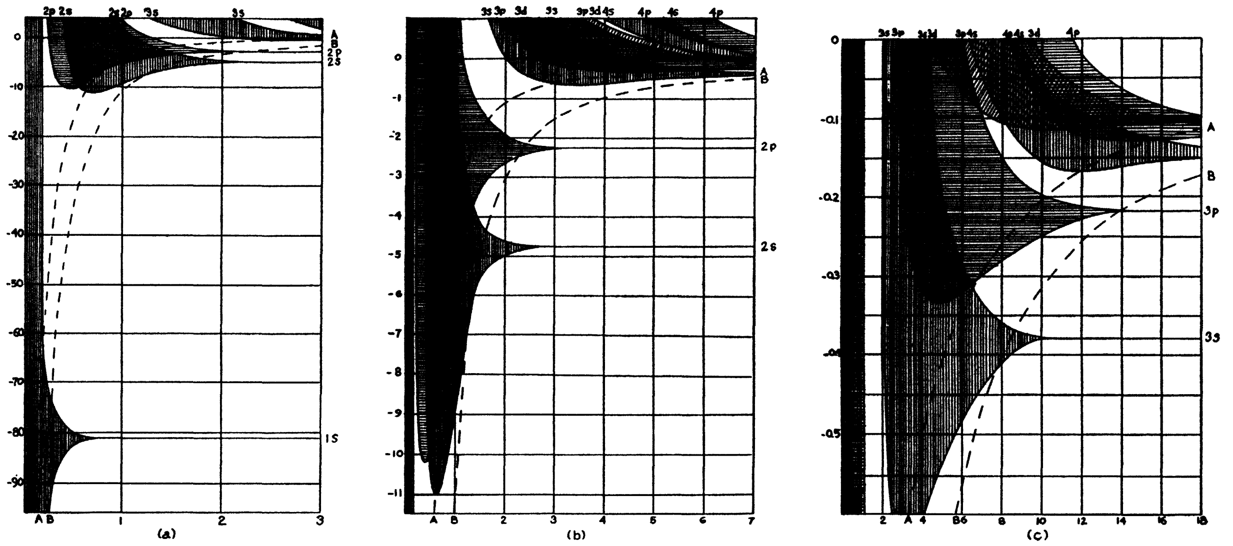

FIG. 9. Energy plotted against internuclear distance for sodium, both in atomic units. Different scales used in (*a*), (*b*), (*c*). The dotted curve *A* represents the maximum potential energy, at the edge of the cell, and *B* represents the mean potential energy throughout the cell, as a function of distance. Most of the curves are fairly accurate, but some are only estimated, so that too much confidence should not be placed in the details of the graphs, though the general relations are correct. Some states are omitted in the upper right-hand corners of all graphs, where the overlapping is really much more than is indicated. In the actual crystal, the half distance of separation is between 3.5 and 4 units.

space, for one of the bands, we discover an interesting thing: the energy is a periodic function of position in the space. This comes about mathematically because the equations connecting energy and p_0 involve trigonometric functions, cosines in Bloch's method, others in the cellular method. The periodicity is a three-dimensional one: there is a lattice of points in the momentum space, much like the lattice of atoms in the coordinate space, and if the momentum coordinates of one of these points are given by P_j, we have $E(p_0)=E(p_0+P_j)$, where E is the energy.

The periodicity in the momentum space can be seen in the following way. The quantity p_0 enters the problem only in the one statement that $u(q+Q_k)=\exp[(2\pi i/h)p_0\cdot Q_k]u(q)$. But now we can show that there are certain vectors P_j, which can be added to p_0 without changing this relation. If that is true, then the whole problem will be just the same for p_0+P_j that it is for p_0, so that in particular the energy will be the same. The condition is simply

$$\exp[(2\pi i/h)(p_0+P_j)\cdot Q_k]=\exp[(2\pi i/h)p_0\cdot Q_k]$$

for every Q_k, or

$$\exp[(2\pi i/h)P_j\cdot Q_k]=1,$$

or

$$(P_j/h)\cdot Q_k=\text{integer}.$$

But $(P_j/h)\cdot Q_k=n$, where n is an integer, is the equation of a plane perpendicular to P_j, and distant $nh/|P_j|$ from the origin, so that putting in all possible integers we have a set of parallel equidistant planes separated by the distance $h/|P_j|$. The condition then is that each of the nuclei, with coordinates Q_k, should lie in one of these planes. But such planes are precisely the ones which are of importance in the Bragg reflection of x-rays and electrons from the crystal. These planes are determined by Miller indices, as (100), (111), etc., and it is easy to show, as we do in Appendix I, that these Miller indices, in the proper units, are simply the three components of the vector P_j determining the planes. Further, it is easy to show that the extremities of the vectors P_j in momentum space form a lattice, which has simple relations to the lattice of atoms. Thus if the atomic lattice is body-centered cubic, the momentum lattice is face-centered, and *vice versa*. These are the only two types of lattice which we shall specifically consider. This lattice in momentum space is called the reciprocal lattice, because the magnitude of the vector from the origin to a point of it is proportional to the reciprocal of the distance between corresponding planes in the real lattice, since as we have seen, this distance is $h/|P_j|$. In the reciprocal lattice, as in the real lattice, it is convenient to surround each lattice point by a polyhedral cell, such that the various cells are all of the same shape and size, and taken together just fill the space. Then we have shown that the whole boundary value problem, and consequently the resulting energy levels and wave functions, have exactly the same values at corresponding points of any two cells in the reciprocal lattice. The periodicity is at once obvious in the Bloch method, where the energy depends on p_0 through the sum of terms $\cos[(2\pi/h)p_0\cdot Q_{mn}]$. If in this function we replace p_0 by p_0+P_j, we have $\cos[(2\pi/h)(p_0\cdot Q_{mn}+P_j\cdot Q_{mn})]$, which, remembering that $(P_j/h)\cdot Q_{mn}=$ an integer, is the same factor as before. In the cellular method, when one carries out the trigonometric work, it is just as obvious.

We thus have the energy given as a periodic function of p_0, for each energy band. By specializing a little we can give a good graphical representation. Thus in Fig. 10 we show two bands, computed by the cellular method, by taking the section of the cell made by the plane $p_{z0}=0$, giving a square section, and by drawing contours of equal energy in this section. The face-centered nature of the reciprocal lattice is well shown by this section. On account of the periodicity, the contours join on smoothly to the ones in adjacent squares, which means that they must cut the edges at right angles. Another way of indicating the energy by even more specialization, is to cut through the cell with a line rather than a plane, and draw energy as a function of position along this line. In Fig. 11 we do this, choosing the line in the (110) direction, or the 45 degree diagonal of Fig. 10, so that we have waves whose corresponding plane wave travels in this direction. In such a diagram, the energy bands can be all indicated on the same plot. We see that the energy has either a maximum or minimum at the center of the cell, and either a minimum or

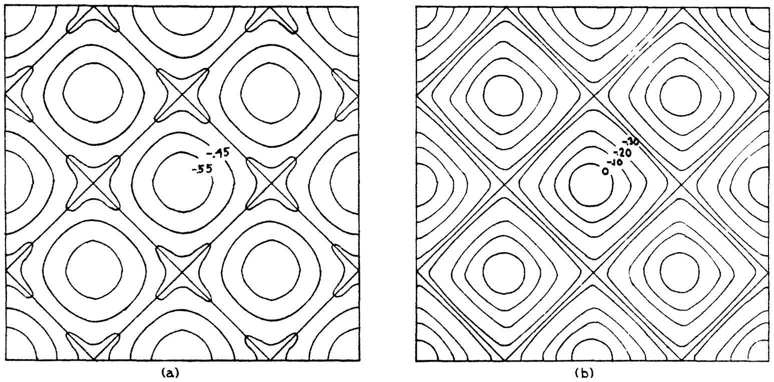

FIG. 10. (a) Energy contours in momentum space, plane $p_z=0$. Drawn for sodium, half internuclear distance = 3.8 atomic units. This represents the energy band going to $3s$ at infinite separation. (b) One of the energy bands of sodium going to $3p$ at infinite separation.

maximum at the boundary. There are furthermore certain general relations: states which go on infinite separation into s states of the atom have a minimum at the center, p states have a maximum, etc. Further, the states far below the dotted line of Fig. 9, which correspond to narrow bands of energy, show only a little variation of energy with p_0, while those far above the dotted line show large and increasing variation. In Fig. 11, it is obvious that each curve in the neighborhood of the middle of the cell shows the general form of a parabola, and the curvature of this parabola continuously increases, from a very flat one for the low energy levels, to a very sharp one for the high levels. For the low levels, where Bloch's approximation is good, the cosine-like form of the curves is apparent, but for the higher levels this is a very poor approximation. Another feature is that for the low levels, the various bands of energy are widely separated from each other, while for the high levels they can even overlap in some cases.

16. Detailed discussion of wave functions

In addition to the energy levels, it is of interest to look at the wave functions themselves. In order to do this, we can draw a straight line through the lattice, and find real and imaginary parts of the complex wave function at points along this line, drawing separate curves for the two. We do this in Fig. 12 for waves travelling in the (110) direction. The line is drawn, not along the wave normal, but in the (111) direction, inclined at a certain angle to the wave normal, and chosen because atoms are much more closely packed in this direction than along the wave normal. Along this direction, the exponential function gives a sinusoidal modulation, the real part behaving like a cosine, the imaginary part like a sine, as is very obvious from the graphs. The functions of Fig. 12 could equally well be determined by Bloch's method or the cellular method.

In looking at the wave functions in detail, it is interesting to consider the nodes, for it is well known that these are of great importance in the theory of characteristic functions, the energy in general increasing as the number of nodes increases (though the relations for a three-dimensional function are not nearly so simple in this respect as for a one-dimensional function). And we observe that nodes occur in the functions for two reasons: on account of those present in the atomic function, and on account of those introduced by the sinusoidal modulation. Thus the very lowest energy level is that coming from the $1s$ atomic function, with $p_0=0$, shown in Fig. 12a. This has no nodes at all. As we increase p_0 in the $1s$ band, however, nodes are introduced, coming from the modulating wave, not the atomic function, for the 1*s* has no nodes. In other words, as far as number (though not exact location) of the nodes is concerned, the functions agree with the plane wave functions $\exp[(2\pi i/h)p_0\cdot q]$. Such a wave is shown in Fig. 12b, c where the dashed lines indicate the modulating sine and cosine wave. This correspondence is lost only at the edge of the cell, where the half wave-length equals the distance between atoms: there, as shown in Fig. 12d there is a node between each pair of atoms, and we have only a real function, rather than a complex one, since the corresponding imaginary function would have a node at each atom, which is impossible for an *s* function. Now a complex function of position corresponds to a progressive wave, as we see by multiplying by the function of time: $\exp[(-2\pi i/h)(Et-p_0\cdot q)]$. But a real function of position corresponds to a standing wave, as $\exp[(-2\pi i/h)Et]\cos[(2\pi/h)p_0\cdot q]$, and this is what we have at this particular wave-length. There is a particular significance in this wave. The standing wave can be considered as being made up of the superposition of two progressive waves in opposite directions, and we can show that the wave-length is just such that the reflected wave is formed from the direct one by Bragg reflection on the (110) planes. We show this as follows. In Appendix II, we show that the general condition for Bragg reflection of an initial wave p_0 to a final one p_1, in the planes described by the vector P_j in the reciprocal lattice, can be stated in two parts: first, a condition of interference, the incident and reflected wave being in the same phase relation at each

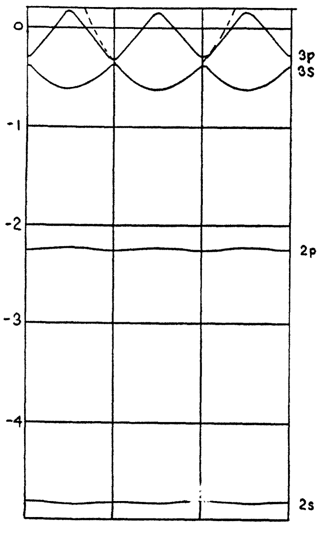

FIG. 11. Energy plotted against momentum, (110) direction. Distance as in Fig. 10. (The 2*s* and 2*p* curves are actually so nearly straight that their curvature would not show on this scale, and has been exaggerated.) The 3*s* and 3*p* curves are those shown in Fig. 10(a) and 10(b), respectively.

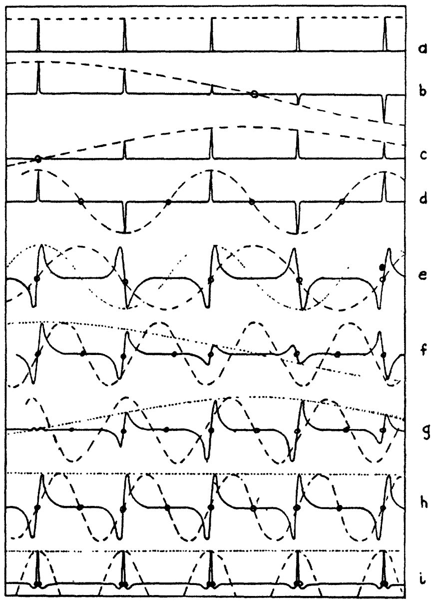

FIG. 12. Wave functions for bound electrons for sodium, as function of distance along a line in (111) direction.

a. 1*s* function, $p_0=0$. Function is real.
b. 1*s*, wave-length corresponding to p_0 is 10 atomic distances. Real part.
c. Imaginary part of (b).
d. 1*s*, wave-length =2 atomic distances, corresponding to edge of cell in momentum space. Function is real.
e. 2*p*, wave-length =2 atomic distances. Function is real, but (d) and (e) are formally related as real and imaginary parts of a function.
f, g. Real and imaginary part, 2*p*, wave-length =10 atomic distances.
h. 2*p*, $p_0=0$, real function.
i. 2*s*, $p_0=0$. Real function, but (h) and (i) are formally related as real and imaginary part of a function.

The dashed sinusoidal curves are the free electron wave functions from the free electron correspondence, having the same number of nodes as the actual function. The dotted curves are the modulating sine curves in which p_0 is chosen from the central cell. They agree for 1*s* states, but not otherwise.

plane of atoms, which is $p_1=p_0+P_i$; secondly, a condition of wave-length, the final wave having the same wave-length, and therefore momentum, as the incident, or $p_1^2=p_0^2$. Now the edge of the cell, in the (110) direction, is given by the relation $p_0=(1/2)P_{110}$. Thus if we let $p_1=-p_0$, we obviously satisfy the relation of wave-length, and at the same time we have $p_1=-P_{110}/2 =P_{110}/2-P_{110}=p_0-P_{110}$, agreeing with the Bragg law for reflection in the $(\bar{1}\bar{1}0)$ planes. We have an interesting explanation of this special case, then: at the particular wave-length for which the electron wave can be reflected from the (110) planes according to the Bragg law, it is totally reflected, and the allowed wave is not progressive at all, but standing, consisting of equal amounts of direct and reflected waves.

So far, for the particular propagation at the edge of the cell, we have found only a real, and not an imaginary, part for the wave function. It seems as if it must be possible to find this imaginary part somewhere, and as a matter of fact it is, if we look for the $2p$ wave functions. The $2p$ atomic function has a nodal plane through the nucleus, and the three degenerate substates of the $2p$ correspond to having the plane in three mutually perpendicular orientations. Correspondingly in the metallic case there are three substates. In particular for the propagation in the (110) direction, one state has the plane normal to the (110) direction, while the other two which in this particular case, though not in general, have the same energy, have nodal planes including the (110) direction. We consider the first. Further, we consider the same edge of the cell as before, so that the modulating factor is such as to change the sign, corresponding to half a wave-length, going from one atom to the next. In Fig. 12e, we see that this $2p$ function has just the same number of nodes as the $1s$ at the same point of the cell, only now the nodes are those arising from the atomic functions, rather than from the modulating plane wave. Further, the nodes of the $2p$ are midway between those of the $1s$, so that formally they are related like the real and imaginary parts of a single complex wave. The $2p$, then, furnishes the missing imaginary function. Like the $1s$ at this point, the $2p$ by itself forms a standing wave, and again the direct and reflected waves satisfy the Bragg relations.

Let us next consider the $2p$ states in which p_0 decreases from the edge of the cell to the center. These are shown in Fig. 12f and g, and we observe the remarkable fact that a decrease of p_0 brings an increase in the number of nodes, until finally for $p_0=0$, in Fig. 12h, we have twice as many nodes as for the maximum value, one midway between each pair of atoms as well as one through each nucleus. The wave function again has only a real, not an imaginary, value, and the $2s$ function with $p_0=0$, shown in Fig. 12i, is related to it as Fig. 12d and e, were related. The increase of nodes of $2p$ with decreasing p_0 seems peculiar at first, but becomes reasonable when we look at the energy diagram in Fig. 11, for there we see that the energy, as well as the number of nodes, of the $2p$ increases as we decrease p_0, verifying the general correlation between energy and number of nodes. But now another possibility of correlation with sinusoidal waves is suggested to us. We recall that the energy and wave function are really periodic functions of p_0. Instead of correlating these $2p$ functions with the plane waves having p_0 from the edge of the first cell, at $P_{110}/2$, in to the origin, we could equally well have correlated them with waves in the next cell, from $P_{110}/2$ out to P_{110}. And if we had done this, the plane waves would have had an increasing, rather than a decreasing, number of nodes, and in fact would in each case have had just as many nodes as the real wave function. In Fig. 12e and i, the dotted lines give the sinusoidal curves connected with p_0's in the central cell, while the dashed lines give the sine curves from outer cells, with the same number of nodes as the actual wave functions.

17. The free electron and atomic correspondences in momentum space

The remark we have just made is the key to an interesting method of correlating the actual solutions of Schrödinger's equation with the values of p_0 in momentum space. We shall not follow through the further details, for they can get very complicated when we consider other types of function and other directions of propagation. But if, instead of confining ourselves to the central cell of momentum space, we use the whole of momentum space, we can assign to each actual wave a value of p_0, not generally in the central cell, which has the following properties: first, it is derived from the p_0 in the central cell by adding a vector P_j of the reciprocal lattice, so that the periodicity condition of the wave can be stated in terms of this p_0 as well as in terms of that from the central cell; secondly, the corresponding plane wave has as many nodes in its real and imaginary parts as the actual wave, though in the actual wave some of these nodes come from atomic nodes. This correspondence can be called the free electron correspondence between wave functions and the momentum space. The reason is that, if the periodic potential were gradually reduced to zero, each wave function would gradually change into the wave function of a free electron, or a plane wave. The wave function would of course change violently in the process, but the one thing which would be unchanged would be the number of nodes, since this is a very fundamental property of the function. If we make this correspondence, then, the wave function which satisfies the relation $u(q+Q_k)=\exp[(2\pi i/h)p_0\cdot Q_k]u(q)$ would go in the limit when the electron became free into the function $\exp[(2\pi i/h)p_0\cdot q]$, and p_0 would go in the limit into the momentum of the free electron.

In the free electron correspondence which we have just described, each wave is assigned a different p_0, so that the p_0 can be used as a sort of quantum number, sufficient to describe a stationary state completely. This method is to be contrasted with the method we have used previously, in which we used only the central cell of the momentum space and chose all p_0's from that cell. In that case, to specify a state, we must give not only the p_0, but also the energy band to which it belongs, or what is the same thing, the quantum numbers of the atomic state to which this energy band goes on infinite separation of the atoms. This method may be usefully called the atomic correspondence. For practical purposes, it is usually much more useful, for the wave functions are much nearer those of free atoms than they are to those of free electrons, as we can see from Fig. 12. It is stretching the point a good deal to make an identification of the nodes, particularly the inner ones near the nucleus, with the nodes of a plane wave. Further, the energy levels are enormously different from those of free electrons. The free electron of momentum p_0 would have an energy $p_0^2/2m$, simply its kinetic energy. On the other hand, as we have seen in Fig. 11, the $1s$ electrons have a very low but practically uniform energy, the $2p$ much higher and again almost the same, corresponding to a great jump of energy between the two cells, and nothing like the continuous law which we should expect from free electrons. Everything indicates that for a metal like sodium, the free electron correspondence is very far from the truth, for the levels which are actually occupied with electrons. Another disadvantage of it is that it is not really nearly as simple as our description would seem to indicate. We have considered only one particular direction of propagation, and specially chosen types of wave function. If we had taken the general case, we should have met great complication. Each state which in the atomic method of correspondence lies in the central cell of momentum space will lie in some other cell in the free electron correspondence (except the $1s$ states), and a very careful consideration of the nodal structure must be made before we can tell which cell is the correct one. In general, to fit into the free electron correspondence a given atomic state which with its different p_0's fills the central cell in the atomic correspondence, it proves necessary to cut this central cell up into many small segments, as if it were a wooden block being cut up with a saw, and to shift these different segments by different vectors P_j into other cells, carrying the values of energy with them. This results in a very complicated system of zones, which has been discussed by Brillouin,[42, 61] with the energy continuous within one of the small segments, but changing discontinuously on the surfaces of the segments. These surfaces of discontinuity prove in general, as in the particular case we have considered, to be surfaces such that any vector drawn from the origin to a point of the surface represents an electron wave which is capable of being reflected according to the Bragg law. A few of these zones are shown in Fig. 13.

18. The metallic correspondence

The free electron correspondence, as we have seen, is theoretically interesting, but has no very close relation to the structure of metals. Relations

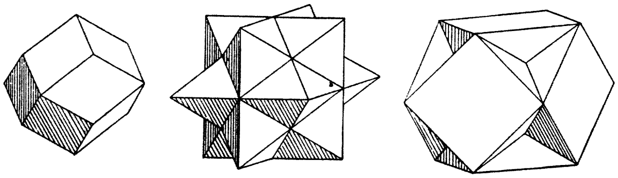

FIG. 13. Zones in momentum space, for body centered space lattice, face centered momentum lattice. First solid represents middle zone, identical with middle cell, bounded by twelve faces parallel to (110) planes. Second zone includes volume between first and second solids, also bounded by (110). Third zone included between second and third, bounded by (110) and (100).

of much more value may be obtained by considering the 3*s* electrons of sodium. Here the wave functions, as shown in Fig. 14, show quite a different relation to sinusoidal waves. In Fig. 14a, b, c, d we show not only the 3*s* electrons for several different p_0's, as before, but also the corresponding plane waves when p_0 is chosen to be in the central zone (the atomic, rather than the free electron, correspondence). We see that while, of course, the number of nodes does not agree at all, the general behavior of the wave function in the region between atoms agrees rather closely with the sine curve. It is only near the nuclei, where, of course, the wave functions have extra nodes characteristic of atomic wave functions, that the two diverge from each other, and this is not over as large a part of the space as we might think, since the volume of a small sphere around the nuclei is to the whole volume as the cubes of the corresponding radii. In perhaps ninety percent of the volume, the plane wave represents the wave function well, throughout the zone. Now these are the electrons which actually function as free electrons in the crystal of sodium, and it is decidedly a matter of interest that the wave functions have such a strong relationship to those of free electrons. The relationship is shown again by referring back to Fig. 11, which gave the energy as a function of p_0. We recall that the various bands of electrons gave energy curves resembling parabolas, but with curvature going from a very small value for the tightly bound electrons to a very high value for electrons of high energy. We have given for the 3*s* band the parabola $p^2/2m$ characteristic of free electrons, and it is seen that it agrees very well with the actual energy curve. For the lower levels, the energy does not vary as much as this with p_0, for the higher levels it varies more. We may bring this into correlation with our diagram of Fig. 9. Here, it seems that the states far below the dotted line have much less energy spread in the central cell than free electrons, those just about on the line have about the same energy spread, those far above have much more.

We may now carry out with the 3*s* and 3*p* wave functions an argument similar to what we made earlier with the 1*s* and 2*p*. As we go to the edge of the central cell, the 3*s*-like functions acquire nodes midway between atoms, as in Fig. 14d. But the 3*p* at the edge of the cell, shown in Fig. 14e, has corresponding nodes passing through the nuclei, being approximately equivalent to plane waves of the same wave-length (neglecting the other nodes near the nuclei, which here, in contrast to the free electron correspondence, do not agree for the 3*s* and 3*p* functions). If then with these 3*p* functions we let either p_0 be gradually decreased to zero, or be gradually increased to twice its value, the number of nodes continues to increase, as in Fig. 14f and g, and the function, except for the nodes near the nuclei, resembles a plane wave of shorter and shorter wave-length, corresponding in fact to the plane wave determined by p_0 in the first cell outside the central zone. This suggests then a correlation between wave functions and points in momentum space, which we can call a metallic correspondence, since it is appropriate for the actual free electrons of the metal, and which is different from both the atomic and the free electron correspondence. In the metallic correspondence, for sodium, we assign the states which go to 3*s* at infinite separation to the central cell, the 3*p*'s to cells directly outside, and so on. The dotted lines in Fig. 14 give the

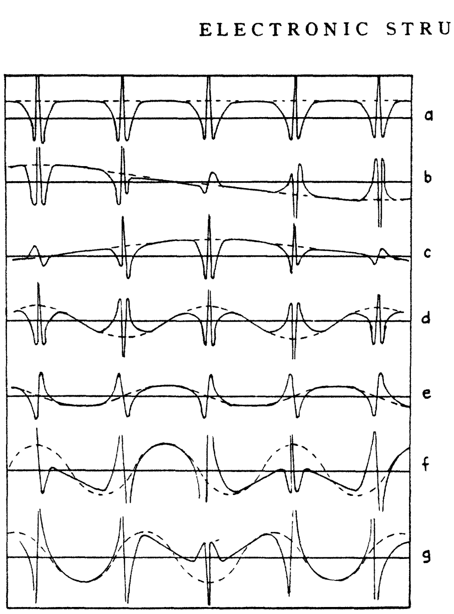

FIG. 14. Wave functions for free electrons, for sodium, as function of distance along a line in (111) direction.

a. 3*s* function, $p_0=0$. Real function.
b, c. Real and imaginary parts, 3*s*-like functions, wave-length = 8 atomic distances.
d. 3*s*-like function, edge of band.
e. 3*p*-like function, edge of band.
f, g. Real and imaginary parts, 3*p*-like function, intermediate position in band.
The dotted lines indicate sinusoidal wave functions for the metallic correspondence.

sinusoidal curves of the metallic correspondence, agreeing with the atomic correspondence for the lowest zone, *a–d*, but not otherwise. One point of considerable significance in Fig. 14 is the way in which a single function has *s*-like properties at some points, *p*-like at others, as is well shown in Fig. 14b, for instance. This indicates the fact that the functions are really linear combinations of different *l* values, as is obvious from our derivation. Another way of looking at it is that the different energy bands have broadened out so that they overlap, with consequent mixing of the stationary states. There is really very little significance, then, in denoting these levels by the atomic correspondence, which is really correct only for energy levels which do not interfere with each other. On the other hand, for the bound electrons, shown in Fig. 12, this phenomenon does not appear. In Fig. 12f, g, for instance, the wave function in the neighborhood of each nucleus is like a 2*p*. Near the left-hand nucleus of Fig. 12g, where the function is actually like a difunction, it is so small as to be negligible.

We should now expect to get good correlation between the energies of the 3*p* electrons and free electrons, and this is shown in Fig. 15, where the 3*p*'s are shown in the cell outside the central one. This curve, as one can see by comparison, is simply a part of Fig. 11, with the rest omitted. The energy curve of the 3*p*'s, as a matter of fact, agrees fairly well with the next part of the parabola beyond the central zone. At the edge of the cell, there is a discontinuity of energy, the curve showing a behavior something like an anomalous dispersion curve. Beyond that, however, for some distance, the curves agree fairly well. That the agreement is not very good, however, is indicated in Fig. 16, where we show the energy in a two-dimensional plot, made up out of our earlier Fig. 10a and b, but with the contours drawn in the zones indicated by the metallic correspondence. If the energy agreed

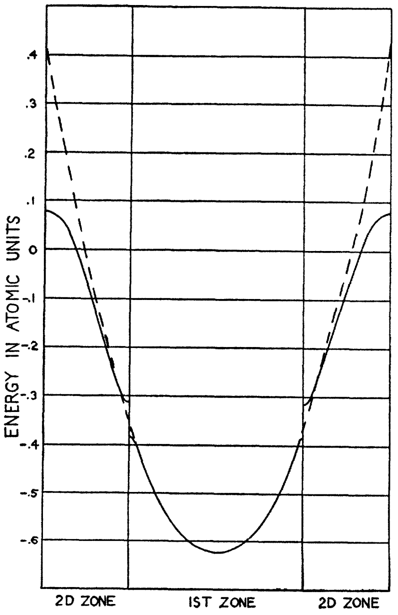

FIG. 15. Energy plotted against momentum, showing two zones.

with that for free electrons, the contours would be circles, the energy being independent of direction. In the central part of the central zone, this is the case, but by the time we get to the next zone, the curves are far from circular. The metallic correspondence is, as we infer from what has just been said, one which holds for a few bands of electrons only. As the wave-length of the corresponding plane wave becomes shorter and shorter, the oscillations of the wave function coming from the plane wave and those from the atomic part of the function about the nucleus become of the same order of magnitude, and they cannot longer be distinguished from each other. Even in Fig. 14f, g this is to some extent the case. The metallic correspondence is one which is satisfactory only when these two types of oscillation are of quite different orders of magnitude. This suggests that sodium, which we have discussed, is a particularly favorable case for finding this correspondence in a pure state, for there the inner shells of the atom are much smaller than the valence electron, and the inner oscillations of the wave function are concentrated into a small volume compared with the whole volume occupied by the atom. In metals like iron, for instance, where there are electrons, the *3d*'s, which are half bound and half free, we could expect no such clear-cut agreement between the wave functions and energy levels and those of free electrons. To put the difficulty simply, the *3d*'s would not know whether they should act as free electrons or not, and whether they should be so counted in the correspondence. But it is useful to have considered such a case as sodium, where the *K* and *L* electrons definitely are not like free electrons, and the *M* electrons are.

In any case, we may expect that as we go to much higher electron levels than these, the levels will become more and more like the real free electron levels, and that the free electron correspondence will become more accurate. This has one interesting result. The free electron correspondence of an electron like *3s* is in a cell further out than that used for the metallic electron correspondence. If we try to go in a regular series from *3s* out to electrons which demand the free electron correspondence, we must then skip over certain cells; there are not enough states to make up for the cells. As a matter of fact, there will be just five cells unaccounted for, the ones occupied by the *1s*, *2s*, and triply degenerate *2p* in the free electron correspondence, and not occupied in the metallic correspondence, where *3s* is in the middle cell. Now of course there will not be five cells, at a certain distance out, which are not occupied. Rather, we shall get to intermediate states between those for which the metallic or the free electron correspondence is more suitable, which will be ambiguous, belonging according to their properties partly to one, partly to another, cell, thus partly filling up two or more cells, and bridging the gap in a way. Later, when we consider the momentum eigenfunctions, we shall be able to understand these states in a much more definite way. A similar situation is encountered if, instead of taking one internuclear distance and one potential function and considering different energy levels, we consider one energy level and allow internuclear distance or potential function to change as parameters. In general, squeezing the lattice or reducing the periodic potential to constancy will have the effect of passing closer to the state of free electrons. It will lower the dotted line in Fig. 9 and move it to the right. As this is done, and as we consider a given energy level, we start with infinite separation or a very large potential variation, with a discrete atomic state. As the

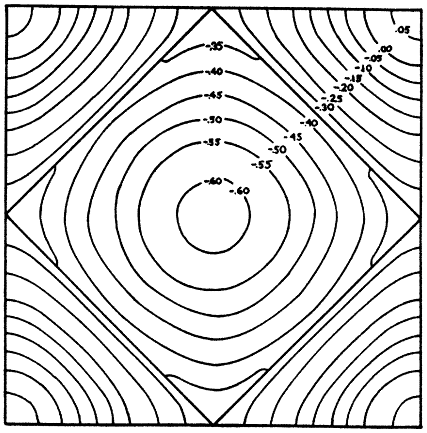

FIG. 16. Energy contours, showing two zones.

crystal is squeezed, we finally come to the point where this state lies about on the dotted line. Then it becomes appropriate to consider this state as lying in the central cell in the metallic correspondence. After a little more squeezing, however, this becomes rather inappropriate, and we assign the state partly to its central cell, but partly to an outer one. Finally the free electron correspondence becomes the correct one for this electron, and it becomes assigned to the proper cell for that, somewhat further out. This assignment is proper, no matter how much further squeezing may be done. In this process, we may say that the state has been promoted, from the inner cell to an outer one characteristic of the free electron correspondence. This promotion is very similar to what is found in diatomic molecules, where the quantum numbers appropriate for a state with considerable separation of the nuclei no longer are appropriate when the nuclei come together.

19. Promotion of stationary states

The conception of promotion of stationary states is necessary if we are going to understand the energy levels as a function of internuclear distance, as shown on Fig. 9. In Fig. 9, we have drawn a line, *BB*, indicating the mean potential energy of an electron averaged through the metal, and we observe that in general the energy levels below this line are practically discrete, those above continuous. We can easily understand the significance of this fact. Below the line, the total energy is less than the average potential energy. Since the kinetic energy is necessarily positive, this means that the potential energy is decidedly less than the average throughout the cell. The only way in which this can be achieved is to have the electronic density concentrated in the regions of low potential energy; that is, near the nuclei. Levels below the line *BB*, then, must correspond to bound electrons, and to fairly concentrated wave functions, which then cannot interfere with the corresponding wave functions in neighboring atoms to broaden the energy levels. On the other hand, levels above *BB* have a higher energy than the average potential energy. The line *AA*, near *BB*, indicates the maximum potential energy in the crystal, so that the total potential energy of the wave function cannot be greater than this value, and all excess of the energy over *AA* must be kinetic. As a matter of fact, the wave functions of Fig. 14 indicate, when computations are made from them, that the distribution of charge density throughout the cell, for an electron with metallic correspondence, is rather uniform, as in a plane wave, so that the actual potential energy is approximately the average through the cell, given by *BB*, and the excess of energy over *BB* measures roughly the kinetic energy. The kinetic energy then is practically zero as the level cuts *BB*, representing a great decrease of kinetic energy from the atomic electron state. As the level goes still further to the left, however, the kinetic energy rapidly increases again, as is shown by the rapid rise of all the curves.

Now let us ask why the energy curves show minima. This is a question which we have so far neglected, but which is of the first importance in questions of the energy of the crystal as a whole. As a matter of fact, we shall find that the minimum in the $3s$ curves corresponds in a rather close way to the minimum in the total energy of a sodium crystal, which gives its actual equilibrium, explaining its cohesion, compressibility, heat of dissociation, etc. It is of great importance therefore to understand why the minimum occurs. We observe that such a minimum is not present for the $1s$ state; as a matter of fact, the reason why all other states have minima is just that they undergo the process of promotion of which we have spoken, with consequent increase of energy as well as of quantum number, while the $1s$ state does not. As in the last paragraph, let us consider a wave function as approximately a plane wave, so that the total energy is the sum of the kinetic energy $p^2/2m$ and the average potential energy, *BB*, Fig. 9. As we compress this wave adiabatically, both terms in the energy change. For adiabatic compression, the wave-length will remain proportional to the linear dimension of the crystal, just as in a box containing standing electromagnetic or sound waves, and compressed, the number of waves will remain unchanged, the wave-length changing to compensate. Since p is inversely proportional to the wave-length, the kinetic energy will increase inversely as the square of the linear dimension of the crystal.

This is the same variation of kinetic energy which we have already found for the free electron model by the statistical method, in Section 5. The potential energy, on the other hand, will decrease as the dimension decreases, as given by *BB*, Fig. 9. For the wave with $p=0$, then, we may expect a regular decrease of energy as the internuclear distance decreases. Of course, this can be expected to hold only for distances small enough so that our approximations are reasonably good; that is, near our dotted line in Fig. 9, for to the right of that line, the atoms are practically separated, and their energy independent of internuclear distance. Now for the 1*s* state, this decrease of energy is just what we find, and for just this reason of decreasing potential energy. But if the metallic correspondence were correct, we should expect for instance that the 3*s* as well would show such behavior, for the bottom of this energy band rather accurately acts as if it consisted of electrons with no kinetic energy, as we have seen. This is, however, far from the case. As the atoms are squeezed, the 3*s* state first has a decreasing energy, on account of decrease in the mean potential energy, but at smaller distances it has a rapidly increasing energy. And the reason can only be that for a smaller distance of separation, the 3*s* becomes decidedly higher than the dotted line, enters the region where promotion takes place, and is on its way up to an energy curve which is a sum of the mean potential energy, and the kinetic energy $p^2/2m$ connected with the momentum associated with this state in the free electron correspondence, a quite high kinetic energy. We may expect the energy level to rise rapidly until it reaches this curve, then more slowly, following that curve, as the volume becomes smaller and smaller. Promotion is, then, an essential feature in understanding the energy levels. We shall be able to understand it much better when we study the momentum eigenfunctions, for then we shall see that the mean kinetic energy increases rapidly in the process of promotion.

20. Momentum eigenfunctions

We have now proceeded about as far as it is profitable to go with the coordinate wave function alone. We shall therefore study the momentum eigenfunction as well, and see in what ways it can contribute to our understanding of the problem.

The coordinate eigenfunction $u(xyz)$ is a solution of Schrödinger's equation $Hu=Eu$, where H is the energy operator, obtained from the Hamiltonian function of ordinary mechanics by replacing the momenta, p_x, etc., by $(h/2\pi i)(\partial/\partial_x)$, etc. That is, the equation is $[-(h^2/8\pi^2m)\nabla^2+U]u(xyz)=Eu(xyz)$, where U is the potential energy, E the total energy. The quantity u^*u, where u^* is the conjugate of u, is the probability of finding the electron in unit volume at xyz, if the wave function is properly normalized.

In a similar way, the momentum eigenfunction $v(p_xp_yp_z)$ is a solution of the equation $Hv=Ev$, where now the operator H is to be found from the ordinary Hamiltonian by replacing the coordinates x, etc., by $-(h/2\pi i)(\partial/\partial p_x)$, etc. It is now the potential rather than the kinetic energy which becomes an operator, and except in special cases this is a more complicated operator than the quadratic one of Schrödinger's equation. The simplest method of handling it is rather different from what we meet in the case of the coordinate eigenfunction. We first expand the potential energy in Fourier series or Fourier integral:* we assume that

$$U(xyz)=\sum_{P_xP_yP_z} W(P_xP_yP_z)\times\exp\left[(-2\pi i/h)(P_xx+P_yy+P_zz)\right],$$

where $W(P_xP_yP_z)$ is the amplitude of the Fourier component associated with the plane wave whose normal is in the direction of the vector P_x, P_y, P_z, and whose reciprocal wavelength is P/h, if $P^2=P_x^2+P_y^2+P_z^2$. In general, all values of the P's must be used, so that the series becomes an integral but we shall find that in the particular cases we are interested in it remains a series, which is the reason why we prefer that method of writing it. Now in the Fourier representation the coordinates enter in a simple way, in the exponentials. Thus the operator associated with U, operating on v, becomes

$$Uv=\sum_{P_xP_yP_z} W(P_xP_yP_z)\exp\left(P_x\partial/\partial p_x+P_y\partial/\partial p_y+P_z\partial/\partial p_z\right)v(p_xp_yp_z).$$

---

* These methods of Fourier expansion are discussed for instance in references 98, 116.

But by expanding the exponential in series, and using Taylor's theorem, it is well known that

$$\exp\,(P_x\partial/\partial p_x+P_y\partial/\partial p_y+P_z\partial/\partial p_z)v(p_xp_yp_z)$$
$$=v(p_x+P_x,\,p_y+P_y,\,p_z+P_z).$$

Hence, remembering that the kinetic energy preserves its classical form, since it does not contain the coordinates, the equation for the momentum eigenfunction becomes

$$[(p_x^2+p_y^2+p_z^2)/2m]v(p_xp_yp_z)$$
$$+\sum_{P_xP_yP_z}W(P_xP_yP_z)v(p_x+P_x,\,p_y+P_y,\,p_z+P_z)$$
$$=Ev(p_xp_yp_z).$$

In this form, generally the most convenient one, the equation is a difference equation, connecting the values of v at a number of different points, rather than a differential equation. The quantity v^*v is the probability of finding the electron in unit volume of momentum space at $p_xp_yp_z$, if the function is properly normalized.

It can be shown without trouble that the coordinate and momentum eigenfunctions can be found from each other by Fourier transformations.[40] Thus

$$u(xyz)=\sum_{p_xp_yp_z}v(p_xp_yp_z)\exp\,[(2\pi i/h)(p_xx$$
$$+p_yy+p_zz)],$$

and

$$v(p_xp_yp_z)=\left\{\int u(xyz)\exp\,[(-2\pi i/h)\right.$$
$$\left.\times(p_xx+p_yy+p_zz)]dxdydz\right\}\Big/\left\{\int dx\,dy\,dz\right\},$$

where the factor $\int dxdydz$ in the denominator can be shown to take care of the normalization, so that if u is normalized, v will also be, and *vice versa*. This and other theorems relating to momentum eigenfunctions are given in Appendix III.

The formulas which we have just written down, expressing the relation between u and v as a Fourier transformation, give the key to the interpretation of the momentum eigenfunction. We see that $u(xyz)$ can be written as a sum of plane waves, each one of the form $\exp\,[(2\pi i/h)(p_xx+p_yy+p_zz)]$, where p_x/h, p_y/h, p_z/h are the components of the vector in the direction of the wave normal, of magnitude equal to the reciprocal of the wave-length. This is just the wave which by de Broglie's relation represents a particle travelling with momentum p, of components p_x, p_y, p_z. Since the whole wave function can be written as a superposition of such waves, this means on the general principles of wave mechanics that there is a certain probability that the electron be travelling with any one of these possible values of momentum. The corresponding probability should be the square of the amplitude of the corresponding plane wave. But this amplitude, is just $v(p_xp_yp_z)$, or the momentum eigenfunction, whose square we have already stated to give the probability of finding the electron with the momentum in question.

We can now prove certain general results which follow from the fact that the potential in a metallic lattice is periodic, by use of the momentum eigenfunction. In order to set this up, we remember that we must use the Fourier resolution of the potential:

$$U(q)=\sum_P W(P)\exp\,[(-2\pi i/h)P\cdot q],$$

where q is the vector of components x, y, z. Then on account of the periodicity of the potential, only certain discrete terms of this summation will occur. Each of these terms will correspond to a certain value of the vector P, and this corresponds to a certain point of momentum space. Hence we can set up a lattice of points in momentum space, the points P which occur in the summation, give the value of W for each of these points, and the result will be a representation of the potential in momentum space. We can now show this lattice to be the same reciprocal lattice which we have already considered.

Let the points of the reciprocal lattice be P_j, and of the space lattice Q_k, where it is assumed that one atom is at the origin. Then the periodicity is expressed by the fact that $U(q)=U(q+Q_k)$. The easiest way to describe this equation in words is to set up the cells surrounding the nuclei which we have already mentioned, so that vectors from a point of one cell to the corresponding points of the other cells are the vectors Q_k. Hence our statement is that the potential has the same value at corresponding points of all cells. But to secure this, we must have just those plane waves $\exp\,[(-2\pi i/h)P\cdot q]$

which have the same value at corresponding points of all cells. That is,

$$\exp[(-2\pi i/h)P_j\cdot(q+Q_k)] = \exp[(-2\pi i/h)P_j\cdot q],$$

or

$$\exp[(-2\pi i/h)P_j\cdot Q_k]=1,$$

from which it is obvious that the expression for $U(q+Q_k)$ in the form of the Fourier expansion equals $U(q)$. But this is just the same condition for the P_j which we have already considered, in Section 15, and have shown to lead to the reciprocal lattice in momentum space.

Now let us set up the equation for the momentum eigenfunction $v(p)$ in the periodic potential field. In the equation

$$(p^2/2m)v(p)+\sum_{P_j}W(P_j)v(p+P_j)=Ev(p),$$

we see that the only points in momentum space which are concerned are first an initial point p_0, and then the corresponding points in all other momentum cells; that is, the points derived from p_0 by making a translation P_j in momentum space. To each point of the central cell of momentum space, for instance, we can set up a lattice of corresponding points, by adding to its momentum p_0 all the various P_j's. And now we see that the difference equation involves only the points of such a lattice. As far as the equation is concerned, it is entirely immaterial what values v may have at any other points. In particular, they may be zero. We shall deal with this case; others may be obtained from it by linear combination. A single stationary state, then, has a momentum eigenfunction which, rather than being defined at all points of momentum space, is defined at only one point of each cell, being zero everywhere else. One result becomes immediately apparent from this. In general, the points of the lattice of corresponding points where the eigenfunction is defined will not be symmetrical with respect to the origin in momentum space. There is no reason why the mean momentum, as defined by averaging over this eigenfunction, should be zero. But if it is different from zero at one instant, then, being a stationary state, it will continually remain different from zero. That is, we can set up stationary states corresponding to electrons having a definite mean momentum in some direction, a momentum which persists in spite of indefinitely many encounters with the nuclei. In discussing the classical problem of collisions, in Section 9, we have already mentioned this property. It is the thing which results in the resistanceless nature of a perfect lattice in wave mechanics.

The fact that the momentum eigenfunction is defined only at the discrete points p_0+P_j allows us to prove the important theorem we have used in discussing the coordinate wave function, that $u(q+Q_k)=\exp[(2\pi i/h)p_0\cdot Q_k]u(q)$. In the first place, the coordinate wave function is a summation,

$$u(q)=\sum_{P_j}v(p_0+P_j)\exp[(2\pi i/h)(p_0+P_j)\cdot q].$$

In the same way,

$$u(q+Q_k)=\sum_{P_j}v(p_0+P_j)\exp[(2\pi i/h)(p_0+P_j)\cdot(q+Q_k)]$$
$$=\exp[(2\pi i/h)p_0\cdot Q_k]\sum_{P_j}v(p_0+P_j)\exp[(2\pi i/h)(p_0+P_j)\cdot q]\exp[(2\pi i/h)P_j\cdot Q_k].$$

Remembering that from the definition of the reciprocal lattice $\exp[(2\pi i/h)P_j\cdot Q_k]=1$, this is simply $\exp[(2\pi i/h)p_0\cdot Q_k]u(q)$, proving our theorem.

21. Atomic correspondence and momentum eigenfunctions

We can now use the method of momentum eigenfunctions to discuss the correspondences already made between coordinate eigenfunctions and the points p_0 of momentum space. First we consider the atomic correspondence, in which p_0 is chosen from the central cell of momentum space. Let us pick a momentum p_0 from the central cell, and consider the lattice of corresponding points associated with it. The momentum eigenfunction is defined at each of these points, so that the set of equations for determining the momentum eigenfunctions becomes a discrete set of algebraic linear homogeneous equations for the discrete set of unknowns. For each point of the lattice, we have an equation

$$(p^2/2m+W(P_0)-E)v(p)+\sum_{i\neq 0} W(P_i)v(p+P_i)=0,$$

where P_0 corresponds to momentum zero, so that $W(P_0)$ is the average value of the potential energy, and where we have an equation for each value p of our discrete set. These equations are just like those often found in quantum mechanics, particularly in connection with the perturbation theory. In the first place, being simultaneous linear homogeneous equations, they have in general no solutions other than $v=0$. The condition that they have such solutions is that the determinant of coefficients be zero. This gives an equation for the energy E, which can be determined in such a way that the determinant will be zero. This equation is of infinite degree; in general, it is of the same degree as the number of equations. Hence there are infinitely many solutions E, and infinitely many corresponding momentum eigenfunctions. These are the various stationary states and energy values of the problem. To distinguish a stationary state, we give the value of p_0, and the quantum numbers distinguishing the atomic state into which the level goes at infinite separation. And for the states which actually are approximately like atomic states, those for which this method of correspondence is a natural one, lying below the dotted line of Fig. 9, this is a natural and unambiguous method of labelling the stationary states. It is the one following naturally from Bloch's method, where each metallic level is derived from an atomic energy level.

For the low-lying energy levels, for which the atomic correspondence is a reasonable one, we can get considerable information about the momentum eigenfunctions by comparison with the case of free atoms. The difference equation for $v(p)$ is in this case very hard to solve, but we can derive the function by Fourier transformation from the coordinate eigenfunction, whose properties we know in this case. We have

$$v(p)=\left\{\int u(q)\exp\left[(-2\pi i/h)p\cdot q\right]dq\right\}\Big/\int dq,$$

where the integration is over all space. We know, however, that $v(p)$ will be zero unless $p=p_0+P_i$, so that we need perform the integration only for these values. We know further that

$$u(q+Q_k)=\exp\left[(2\pi i/h)p_0\cdot Q_k\right]u(q),$$

and that

$$\begin{aligned}\exp\left[(-2\pi i/h)(p_0+P_i)\cdot(q+Q_k)\right]&=\exp\left[(-2\pi i/h)p_0\cdot Q_k\right]\\&\quad\times\exp\left[(-2\pi i/h)(p_0+P_i)\cdot q\right].\end{aligned}$$

Hence the integrand is unaltered if we replace q by $q+Q_k$, or go from one cell to the corresponding point of another. In other words, each cell contributes the same amount to $v(p)$, and if we wish we may define it by integrating only over one cell.

Now within a cell, $u(q)$ is a sum of terms each coming from the solution of a central field problem, of the form $\exp(\pm im\phi)P_l{}^m(\cos\theta)u_{nl}(r)$, where r, θ, ϕ are polar coordinates. In this case, if we assume that the polyhedral cell can be approximated by a sphere, the Fourier resolution can be carried out, and shows that the momentum eigenfunction is as follows, where P, Θ, Φ, are polar coordinates in momentum space, so that P is the magnitude of the momentum:

$$v(p)=\exp(\pm im\Phi)P_l{}^m(\cos\Theta)v_{nl}(P),$$

where apart from a normalization constant,

$$v_{nl}(P)=\int r^2u_{nl}(r)\frac{J_{l+1/2}(2\pi rP/h)}{(2\pi rP/h)^{1/2}}dr.$$

If $u_{nl}(r)$ is a Laguerre function, as in the discrete states of hydrogen, and if the orbit is small enough so that $u(r)$ is practically zero at the edge of the cell, the function $v_{nl}(P)$ turns out to be expressible in terms of a Gegenbauer function, as shown by Podolsky and Pauling.[26] In Fig. 17, we plot the quantity $P^2v_{nl}{}^2(P)$, proportional to the probability of finding an electron whose momentum is in magnitude between P and $P+dP$, for several quantum states of hydrogen. We see that the function is a continuous function of P, showing various maxima, with zeros between. As a matter of fact, v_{nl} shows the same number of nodes as u_{nl}. There is a sort of reciprocal relation between the two functions. Thus for s states

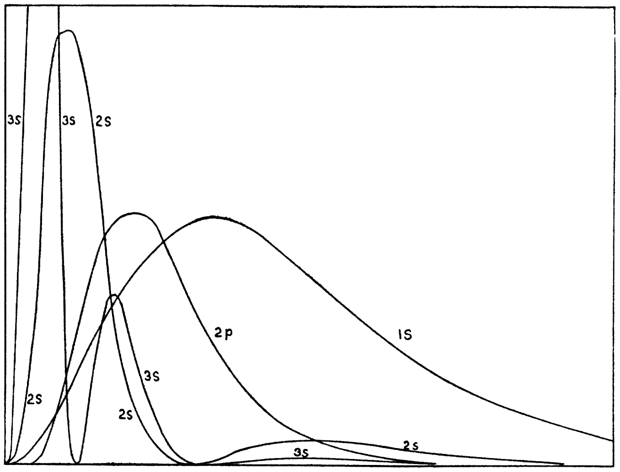

FIG. 17. Momentum eigenfunctions for free atoms.

($l=0$), the number of nodes is one less than the principal quantum number n. In the coordinate eigenfunction u, as n increases, more and more maxima are added to the function for large values of r. The inner maxima, for small r, remain approximately fixed in position, though they decrease in magnitude, while the size of the orbit as a whole increases rapidly with n. Now the inner maxima of u correspond to the outer maxima of v, and *vice versa*, for it is when the electron is far out, corresponding to large r, that it is going slowly, and in a region of large wave-length and therefore small wave number, so that this corresponds to small P. Thus as n increases, the new maxima of v are added inside the old ones, rather than outside. The outer maxima retain their approximate positions, but become reduced in amplitude, the innermost maximum having by far the greatest amplitude. Thus the effect is that the average magnitude of momentum, and consequently the kinetic energy, decrease rapidly as n increases. This of course is just what we should expect. The total energy decreases rapidly in magnitude as n increases, and for an inverse square field the kinetic energy is numerically equal to the total energy. Qualitatively similar results are found for other azimuthal quantum numbers, and the inner maxima agree roughly with those for s states, while the outer maxima, corresponding to the inner part of the orbit, are absent.

The results which we have just given are for hydrogen-like wave functions, with orbits so small, or cells so large, that the atoms can be considered as independent. When we consider the cellular structure of momentum space, we recall that instead of regarding v as a continuous function of position in momentum space, we are to give its values only at the discrete points p_0+P_j. It becomes of importance, then, to know the relative magnitudes of the cells, and of the region of momentum space in which the eigenfunction is large. It is immediately obvious that if the atomic wave function is confined to a region small compared with the cell of coordinate space, as it will be for an inner electron, like the $1s$, $2s$, or $2p$ of sodium, then the momentum eigenfunction must extend with large amplitude over many cells. The outer, and even the inner, maxima of the eigenfunction will be far from the origin. The reason is seen in the reciprocal nature of the momentum and coordinate eigenfunctions. A wave in coordinate space whose wave-length is of the order of magnitude of the lattice spacing will be represented in momentum space by a point about on the edge of the first cell. A much shorter wave will have a representative point in momentum space which is correspondingly further from the origin. Now the oscillations of a 1*s*, 2*s*, or 2*p* function, for sodium, will have an effective wave-length much smaller than the grating space, and consequently the momentum eigenfunction will be much spread out. As a result of this, there will be many different cells, or values of P_j, for which $v(p_0+P_j)$ will be appreciable and of the same order of magnitude. Also, the oscillations of $v(p)$, which we have seen in Fig. 17, will be on a larger scale than the scale of cells, so that the discrete values of the v's will in fact give a fairly adequate representation of the continuous curve.

The actual wave functions, of course, will be different from the hydrogen-like wave functions which we have been discussing. In the first place, the radial functions $u_{nl}(r)$ will be different, and in the second place it is necessary to add functions of different l values to get the complete metallic eigenfunction. For the inner electrons, however, neither of these processes will change the fundamental properties we have described, of having the momentum eigenfunction extend over many cells, with a principal maximum which is at smaller and smaller momenta for higher total quantum number. We shall not consider the eigenfunctions of these bound electrons further, for their principal interest comes simply in the contrast which they afford to the eigenfunctions of free electrons.

When we come to a state like 3*s* in sodium, at the normal grating space, the situation is quite altered. We have seen that with increasing total quantum number the maximum of the momentum eigenfunction moves to smaller and smaller momenta. For the sodium 3*s*, the coordinate wave function has two sorts of oscillation: the short wave-length oscillation about the nuclei, and the long range wave extending throughout the metal, whose wave-length is longer than the lattice spacing. The former, in Fourier analysis, will give components of high momenta, just as with the other waves we have considered, but the latter will give a single component at p_0, in the middle cell. Further, this single component will have a much larger amplitude than the others, since it represents by far the most important part of the oscillation of the wave function. It corresponds to the principal maximum of the bound electrons, which move down toward the central cell of momentum space, until finally for the first band of metallic electrons, it reaches the central cell, and becomes by far the most important component of the eigenfunction. We now see why it is reasonable, in the metallic correspondence, to assign the 3*s* state to the p_0 in the central cell: we make such assignments only when the momentum eigenfunction is much larger at one point of momentum space than at any other, so that this point can be unambiguously assigned to that state.

22. Free electron and metallic correspondence and momentum eigenfunctions

Before pursuing the metallic correspondence further, let us look at the free electron correspondence. For sufficiently high energy states, far above the 3*s*, the coordinate eigenfunctions may be expected to resemble fairly closely the plane waves of free electrons. Now a plane wave has a momentum eigenfunction which is infinite at one point of momentum space, that corresponding to the momentum of the electrons in the wave, and is zero everywhere else. Further, the point where it is infinite is at a distance p from the origin, where $p^2/2m$ equals the kinetic energy of the state. Thus as the energy increases, the infinite maximum of the function moves further and further from the origin. There are, of course, many states of the same energy, but different direction, corresponding to maxima at different points on the same sphere in momentum space. We may infer, then, that each stationary state of high energy has a momentum eigenfunction which is large at one point only of momentum space. It is then reasonable to denote this stationary state by the p_0 of that point. If the periodic potential field were reduced to zero, these states of high energy would retain approximately their same eigenfunctions, but the lower states would also approach free electron eigenfunctions. Thus for the bound electrons the values of $v(p)$ for the outer parts of their eigenfunctions would gradually decrease, as the minima of potential about the nuclei became less deep and the electrons did not acquire such great kinetic energy there, and the values in the inner parts would correspondingly increase, until finally for zero potential the 1*s* state, for instance, would be represented by a single point in the central cell of momentum space, and so on. In this way, as we have seen before, we can assign each state to a particular point of momentum space, obtaining the free electron correspondence. But now we see that for the states of high energy, even with the potential and grating space which exist, this is a reasonable correspondence, while for the low states it is not. The criterion for the reasonableness is again that the assignment can be made physically only if the eigenfunction is much larger at the one point of momentum space to which the stationary state corresponds than it is at any other point.

Now in order to discuss the metallic correspondence, we need an interesting theorem in regard to momentum eigenfunctions. We recall that the equation for these eigenfunctions is a difference equation, a set of algebraic equations of the sort met in perturbation theory. The solutions of these equations, $v(p)$, as always in such cases, can be interpreted as a set of coefficients determining an orthogonal transformation of variables. If we denote the nth atomic state by a subscript n, we are given as a result, for a definite p_0, the two parameter set of numbers $v_n(p_0+P_j)$, with the parameters n and j, giving the amplitudes for the nth state in the jth cell. The orthogonality and normalization conditions, as shown in Appendix III, can then be written in either of two forms. First, we may write

$$\sum_{P_j} v_m^*(p_0+P_j)v_n(p_0+P_j)=1 \text{ if } m=n,\ 0 \text{ if } m\neq n.$$

This means first that if we take the momentum eigenfunctions of two stationary states, multiply them together for each cell, and add, the result is zero, the orthogonality condition. Secondly, it means that we have normalized so that the sum of the squares of all the $v(p)$'s for all cells, for one eigenfunction, is unity. This is a convenient method of normalization. For the bound electrons, for instance, for which many $v(p)$'s are different from zero, it means that all these $v(p)$'s must be small, since the sum of the squares of the large number of terms is unity. For the free electrons, on the other hand, where one component is much larger than any other, this large component is almost unity, the others being very small compared with unity. But now in addition to this statement of the orthogonality condition, there is as always another form, in which we sum over the other index:

$$\sum_n v_n^*(p_0+P_j)v_n(p_0+P_k)=1 \text{ if } j=k,\ 0 \text{ if } j\neq k.$$

The orthogonality condition coming from this statement is one we shall not have occasion to use, but the normalization condition,

$$\sum_n v_n^*(p_0+P_j)v_n(p_0+P_j)=1,$$

is a very valuable theorem. It means that if we take the square of the momentum eigenfunction in a given cell, and sum, not over cells, but over all different stationary states, we get unity. It is sometimes called the completeness theorem, for if we have not the complete set of functions, with all n values, the sum will not be unity. Let us see what this means. For the outer cells of momentum space, the free electron correspondence is good. That is, for each such cell, with a given p_0, there is a state which has a contribution almost equal to unity in that cell, with very small contributions in other cells. But then our theorem tells us that, so to speak, this one state almost fills up that cell. If for it we have $v_n^*v_n$ almost equal to unity, all other states put together can have but a very small contribution to the sum, which means that all other states must have small momentum eigenfunctions in that cell. It is obvious that this theorem, can lead to important relations between the eigenfunctions of different states.

For sodium, five states, the 1*s*, 2*s*, and three substates of 2*p*, are very definitely bound electrons. For these, $v(p)$ is appreciable in many cells. The sum of v^*v, for these five states, and for all cells, is just five units. The cells where the sum of v^*v over the five states is appreciable lie in a zone in momentum space: none of them are for states of too large total momentum, as we see from Fig. 17, and none of them are right around the origin. If then we subtract the sum of v^*v for these bound electrons from unity, for each cell, we get the contribution which the other, metallic and free, electrons are allowed to make, and we see that this contribution is almost unity for the innermost cells, almost unity outside a certain zone, but less than unity between. Now to have good free electron correspondence, we must have the momentum eigenfunction practically unity at one point, practically zero elsewhere. This is only possible outside the zone occupied by the bound electrons. We can state this condition in a simple way: it is only possible for free electrons which everywhere move faster than bound electrons ever do to an appreciable extent. We can again have good correspondence, this time the metallic correspondence, in those cells within the zone occupied by bound electrons. These are for electrons moving more slowly than the bound electrons do to any appreciable extent. In between, on account of the completeness relation, it is impossible to have wave functions which correspond very closely to sine waves. This zone may be called the zone of promotion. In it, the number of stationary states apart from those of bound electrons is five less than the number of cells, so that each wave function must have appreciable contributions from several cells, and there is ambiguity as to the proper correspondence, something between the metallic and free electron correspondence being indicated. In this zone of promotion, wave functions and energy values are all abnormal, agreeing neither with metallic nor free electron models. We have already seen, for sodium, the beginning of this zone, coming just outside the first cell of momentum space. Only a few states show the metallic correspondence at all accurately, for this metal. And, as we have pointed out, this is a specially favorable case for metallic correspondence. For most metals, there would not be such a sharp contrast between free and bound electrons as for sodium. The outermost bound electrons would have larger orbits than in sodium, so that their momentum eigenfunctions would extend in practically to the origin, with the result that the zone of promotion would extend to the origin, and there would be no true zone of metallic electrons at all. We must conclude, therefore, that for most metals the electrons which actually take part in conduction have complicated wave functions and energy levels, which must be investigated separately in each case, and which are not very susceptible to generalizations. This in a way seems discouraging to theoretical progress, but in another way it is very fortunate. For it indicates that we may expect great variations between metals, with properties varying in apparently erratic ways. It indicates also the futility of obtaining general theories of such detailed things as ferromagnetism, Hall effect, thermoelectric effects, etc. These things must all depend on the peculiarities of the individual metal, and their detailed working out must be based on a study of the peculiarities of the corresponding wave functions. Other things as well must depend on these wave functions in the zone of promotion. Two examples are the optical properties of metals in the ultraviolet, and electron diffraction by slow electrons, two fields in which surprising experimental results have been obtained, with but small progress in correlating them with other properties of the metals in question.

23. Perturbation method for free electron and metallic correspondence

In cases where either the metallic or the free electron correspondence is good, the momentum eigenfunction by hypothesis has a large value in one cell, with small values in other cells. In this case, it is legitimate to solve the difference equations for the eigenfunction by perturbation theory. We choose p_0 to be the value for which the eigenfunction is large, by whichever correspondence is appropriate. Using perturbation methods,* we then have, up to second order perturbations,

$$E=p_0^2/2m+W(P_0)+\sum_{i\neq 0}\frac{W^2(P_i)}{p_0^2/2m-(p_0+P_i)^2/2m},$$

and $v(p)$ is given by a sum of terms from all points of the lattice corresponding to p_0, the amplitude of the one displaced a distance P_j from p_0 being proportional to

$$W(P_i)/[p_0^2/2m-(p_0+P_i)^2/2m].$$

We see that the components $W(P_j)$ take the place of the usual nondiagonal components of energy in perturbation theory. Now from the first expression, we see that one of the terms in the energy becomes infinite, resulting as we

---

* See, for example, references 42. 44. 59 61. 98. 116.

prove in Appendix IV in a discontinuity of amount $2\,|W(P_i)|$ in the plot of energy against p, if $p^2=(p+P_i)^2$, or if the vectors p and $p+P_i$ have the same magnitude. But this is just the condition of Bragg reflection, by which the wave represented by p is reflected in the planes represented by P_i, giving a reflected wave represented by $p+P_i$. We have already seen this in Appendix II. As shown in the figure in Appendix II, if we set up the plane which is the perpendicular bisector of $-P_i$ in momentum space, any vector p from the origin to a point of this plane satisfies the condition just stated. In other words, we can set up the infiite set of such planes in momentum space, given by the set of P_i's, and as we explore the energy as a function of position in momentum space, using the metallic or free electron classification, we shall find a discontinuity in energy every time we cross such a plane. These planes divide the space into the complicated set of zones which we have considered in Section 17.

The conditions of a Bragg reflection are two, as we have already seen: first, the reflected wave must be such as to show correct phase relations to the incident one at every plane of atoms from which the reflection is occurring, or $p_1=p_0+P_i$; secondly, the frequency or wave-length of the reflected wave must equal that of the incident wave, or $p_1^2=p_0^2$. Both conditions are satisfied, as we have seen, for the waves at points of discontinuity in momentum space. The first alone is satisfied, however, for every wave which must be combined with the wave p to get the complete eigenfunction of the state. And our perturbation theory tells us that the more nearly the second is satisfied, and at the same time the larger the component of energy $W(P_i)$ connected with the planes in question, the larger will be the coefficient of the plane wave in question. We may work the perturbation theory by the method of variation of constants, starting at $t=0$ with a wave function which is exactly a plane wave. As time goes on, the other plane waves will come in with amplitudes increasing with time, the amplitude of the one coming from reflection in the plane P_i being proportional to

$$W(P_i)/[p_0^2/2m-(p_0+P_i)^2/2m].$$

We may interpret this by saying that as a plane wave travels through the crystal, it becomes scattered, by reflection in the various planes of the crystal. Those reflections are more likely in planes in which the crystal has much density (large $W(P_i)$), and for which the reflection can be most nearly without change of energy (p_0^2 most nearly equal to $(p_0+P_i)^2$). The fact that reflections with change of energy can occur at all is a result of the fact that the states we are dealing with, the plane waves, are not really stationary states, so that they have short lives, and corresponding uncertainty in their energy. Translating back to the language of particles, a beam of parallel electrons of the same energy passing into the crystal will be scattered, more and more leaving the beam as time goes on. But a stationary state is possible, consisting of a certain mixture of all the various beams of electrons which can be obtained from the original one consistent with the interference law of wave mechanics. In this stationary state, as many electrons will be scattered out of the beam in unit time as are scattered in, so that there will be a net forward motion in the beam. This is the analogue in wave mechanics to the sort of stationary beam, conserving momentum in spite of collisions, which we discussed in connection with classical mechanics in Section 9.

24. Population of energy levels in the complete atom

We have now obtained a fairly adequate picture of the coordinate and momentum eigenfunctions of electrons in such a periodic field of force as we have in a metallic lattice. Before proceeding further, we should ask which of these stationary states are occupied in the normal metal, and what are the corresponding density distributions of all the electrons in coordinate and momentum space. Aside from certain corrections having to do with electron spin, which we shall later discuss, the lowest stationary state of the metal as a whole is found when there are two electrons, of the two possible spins, in each of the stationary states of a single electron, counted up from the lowest energy level, and including just enough levels to take care of all the electrons. Let us consider first the bound electrons. Each such electronic state is characterized by a value of p_0, which by the atomic correspondence we choose to be in the central cell of momentum space, and by quantum numbers similar to atomic quantum numbers. We have spoken so far as if p_0 could take on any value within the central cell, so that there would be an infinite number of such states. But actually, as we show in Appendix V, if there are N atoms in the crystal, there are only N possible values which p_0 can have, if we are to satisfy boundary conditions at the surface of the crystal. Further, these N possible p_0's are arranged with uniform density throughout the cell. Hence, corresponding to one state of a free atom, there are just N states of the crystal, or a possibility of $2N$ electrons, 2 per atom, if we consider both spins. This is just the number which we should have in the same number of separated atoms. When all these levels are filled up, the total density in momentum space will no longer be confined to certain points, but will fill the cells smoothly, so that the total density function will look like that of Fig. 17, being much like that for electrons in separate atoms. This will be the case for all the bound electron levels. There will then be a certain number of electrons left over for the metallic electron states. For instance, in sodium, there will be one electron per atom. Now the central cell in the metallic correspondence for sodium is large enough to hold two electrons per atom. Thus this cell will be only half filled up. Naturally the inner part, corresponding to lower energies, will be filled, with equal numbers of electrons of each spin, and we see that in this inner part, the energy levels are almost exactly those for free electrons, as we expect from the accuracy of the metallic correspondence here. Now by the completeness relation, the superposition of all these states cannot fill the momentum space to more than the constant maximum density of $2N$ electrons per cell anywhere, but in the inner half of the central cell it will be filled practically to this density, for the higher states, which are not occupied, have almost no contribution to their momentum eigenfunctions in the central cell. Outside the inner half of the central cell, however, the density will fall off sharply from this value, representing practically the density of the inner electrons, which will fall off gradually, extending out to large momenta. The net result, then, will be a distribution in momentum space very similar to that of Fig. 7, derived from the Thomas-Fermi theory. Here, as there, the sharp break in density at the outer boundary of the metallic electron region is present. And here as well, if we set up the distribution for a temperature above the absolute zero, regulating the number of electrons in each stationary state by the Fermi statistics, a rounding off of the sharp break in the distribution would occur just at this point.

For an element with more free electrons per atom than sodium, it is obvious that more than the first half cell will be filled with these electrons, so that the break in density will come for a larger momentum, as we have seen from the Thomas-Fermi method. Furthermore, in such an element probably no electrons would show such accurate metallic correspondence as with sodium. As a result, the break would probably be less sharp, and more complicated in structure, than for a monovalent metal, but since detailed calculations have not been made, it is not possible to be more specific as to its structure.

V. THE WAVE-MECHANICAL METHOD: CONDUCTIVITY AND ENERGY

25. Conductivity in the wave-mechanical model

Now that we understand the electronic structure of metals, using the picture of electrons moving in a periodic potential field according to wave mechanics, our next task is to investigate the electrical conduction in the metal. This is a problem in which the interactions of the electrons are not important, so that our model should be quite adequate to discuss it. We have seen that a given electronic eigenfunction in general corresponds to a certain mean momentum, and mean current, which persists in spite of collisions. This mean momentum is simply the center of gravity of the momentum eigenfunction. Our next question is, how does this mean momentum change under the action of an external electric field? The answer is very simple. The momentum eigenfunction, as we have seen, is defined only at certain points, given by p_0+P_j, where p_0 is a constant, P_j the elements of the reciprocal lattice. But now under the action of the external force F, these points are no longer fixed, but move bodily through the momentum space, according to the vector equation $F=dp_0/dt$, analogous to Newton's second law of motion. This can be proved easily from the equation for the momentum eigenfunction. We have ordinarily written this

$$(p^2/2m)v(p)+\sum_{P_i}W(P_i)v(p+P_i)=Ev(p),$$

but now, since we are dealing with time variation, we wish the form involving the time. If by $v(p, t)$ we were now to understand the function involving the time, equal to our former $v(p)$ times $\exp(-2\pi i/h)Et$, the equation for $v(p, t)$ would be

$$(p^2/2m)v(p, t)+\sum_{P_i}W(P_i)v(p+P_i, t) = -(h/2\pi i)\partial v(p, t)/\partial t.$$

Now we wish to add to the potential $W(P_i)$ that of a uniform field, and find the change in $v(p, t)$. To do this, we shall add a potential $-(F_xx+F_yy+F_zz)$, where F_x, F_y, F_z are the three components of the force vector. Now we could analyze this potential in Fourier series, as we have done with the other potentials, but it is more convenient to proceed directly, replacing x by $-(h/2\pi i)\partial/\partial p_x$, etc., according to our general rule. When we do this, and collect the added terms on the right side of the equation, the result is

$$(p^2/2m)v(p, t)+\sum_{P_i}W(P_i)v(p+P_i, t) = -(h/2\pi i)(\partial/\partial t+F_x\partial/\partial p_x+F_y\partial/\partial p_y +F_z\partial/\partial p_z)v(p, t).$$

But now if we imagine that the whole distribution $v(p, t)$ is flowing in the momentum space with a velocity F, as well as undergoing changes with time directly, the right side of this equation can be written as $-(h/2\pi i)Dv(p, t)/Dt$, where D/Dt represents the derivative following the streaming points in momentum space. We thus see that the equation for the momentum eigenfunction in the presence of the external field, as we follow along with the points, is just the same as for the function at fixed points of momentum space in the absence of the external field. The solution must then be the same, so that we can set up a vector p_0, which must be assumed to change with a time derivative equal to F, and the eigenfunction will have the same values at the points p_0+P_i, at any instant of time, that it would have at the same point in the absence of a field, differing only in its steady motion through momentum space.

There is one point regarding the momentum eigenfunction which must be understood. As the points p_0+P_j stream through momentum space, the values of the function v at the various points do not stay constant; they vary, so that at each instant they correspond to the eigenfunction of a stationary state in the absence of field having the instantaneous values of p_0+P_j. This variation in the first place has an effect on the eigenfunction and the energy. It is part of the time variation of v as we follow the streaming points in momentum space, so that it contributes a certain amount to $Dv(p, t)/Dt$, an amount furthermore which is different for the different points p_0+P_j. Since from our solution we have $Dv/Dt=(-2\pi iE/h)v$, where E is the same for all points and is the energy as computed for the field free case, and since part of this Dv/Dt comes from the time variation of the magnitude of v, we see that the remainder, which is available for the variation in phase which gives the energy in the ordinary sense, is not precisely the same for all points. Hence our ensemble does not precisely correspond to a set of particles all of the same energy. But since the time rate of change of the magnitude of v is proportional to the external field F, and since for all ordinary cases this is very small compared with atomic fields, this correction is ordinarily negligible. The other point, however, is of real physical importance. We recall the periodicity of the whole problem in momentum space, such that corresponding points of two cells give just the same eigenfunction and energy. As a result of this, if we allow the external force to act long enough so that p_0 has travelled from one cell into another, and if it happens to be travelling in such a direction that it reaches a corresponding point of the second cell, the eigenfunction and energy will come back to the values which they originally had. If for example we consider the average momentum, which is proportional to the current, we shall find that as the field is applied, the momentum begins to increase in the direction of the force, but at a certain point, a remarkable change will occur, the momentum will decrease below its original value, and then will commence an increase again, so that after traversing a complete cell it will have just reached its original value again.

To understand the curious behavior of which we have just spoken, let us consider the case of an almost free electron. There the momentum eigenfunction is practically unity at one point, say p_0 (if we use the free electron correspondence), and very small at the other points p_0+P_j. The average momentum is then practically p_0, so that its time rate of change is the external force, and the particle accelerates as in classical mechanics. Let it be accelerated outward from the origin in momentum space. Then as p_0 approaches a surface of discontinuity connected with a Bragg reflection, we know that the eigenfunction approaches a standing rather than a progressive wave, and the Fourier representation of this consists of equal amplitudes for a direct and reflected wave, or at the face of the zone which p_0 is approaching, and at the opposite face. In other words, near the face, there is a remarkable change in the amplitudes, the component p_0 decreasing to half its intensity, and the component p_0-P, where P corresponds to the face we are considering, building up to a corresponding value. As the process continues, the component p_0 decreases still further, rapidly approaching zero, while the other component, which now is in the central cell, builds up to a value unity. What has physically happened is very simple: the particle has been accelerated until it has the proper speed to make a Bragg reflection in the planes represented by P. At the proper moment it makes this reflection, reversing its momentum. As the field continues to act, this negative momentum is decreased numerically by the field, the particle being slowed down. As this goes on, the original component p_0 has been moving outward in the zone outside the one it started in, but it has lost practically all its intensity, while the component which originally in the opposite zone on the other side has taken on practically all the intensity. And physically, instead of following the one or the other component, the particle has been reflected, going from one component to the other.

These relations are even clearer if we plot the average momentum and average energy as functions of p_0, which is chosen to be a vector perpendicular to one of the faces of the cell, so that we can get the periodicity of which we have already spoken. These are shown, for the (110) direction in sodium, in Fig. 18. The curve for energy is the same one we have already obtained, in Fig. 11. At the edge of the cell we can see the sudden change in momentum, arising from the Bragg reflections. Correspondingly there is a sudden change in the slope of the energy curve, the energy stopping its increase with p_0, as we have when the particle is travelling in the direction of the field, and commencing to decrease, as it does when the particle is opposed

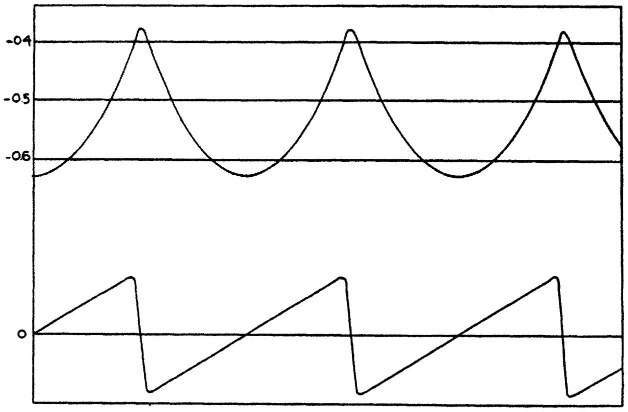

FIG. 18. Energy and momentum as functions of p_0.

by the field. One result is suggested by this figure: the curve for average momentum looks like the derivative of the one for average energy. As a matter of fact this is true, and it is an important theorem, for it allows us to compute the momentum, and current, from the energy curves, without making a separate computation.

The theorem which we have just suggested is the following: if E is the energy, as a function of p_0, then $\partial E/\partial p_{0x} = x$ component of the velocity, averaged over the eigenfunction.[112, 116] This can be proved without trouble by direct computation, as we show in Appendix VI. It seems almost obvious that it must be true, however, from the following simple argument. Let a constant force F act in the x direction. Then the rate of working of the force should be F times the x component of velocity. This, on the other hand, should be the time rate of change of energy, or

$$(\partial E/\partial p_{0x})(dp_{0x}/dt) = (\partial E/\partial p_{0x})F,$$

from which the relation follows at once. Some interesting qualitative results follow from this theorem at once. The bound electrons, as we have seen, have energy curves which vary only slightly with p_0. Thus they must correspond to very small mean velocity or mean momentum. In momentum space, where the eigenfunction for these electrons is spread out through wide ranges of momentum, this means that the distribution is so nearly spherical that the center of gravity, giving mean momentum, lies almost at the origin. In coordinate space, the potential barriers between atoms are so high that very little current can flow over the top. The situation is just reversed for free electrons, where the momentum eigenfunction is almost all concentrated at one point, which can be far from the origin, giving large mean momentum, and where the energy is above the potential barrier, allowing free flow of current.

Having investigated the eigenfunctions of individual electrons in the presence of an accelerating field, we can next inquire what would happen to a whole collection of electrons, as we have in a metal. In the first place, suppose we have an energy band entirely full of electrons. At the origin of time, we may take the central cell of momentum space to be just filled to its maximum density with representative points. As a result of symmetry, the mean momentum will be zero. But now let the field act. After a certain lapse of time, some points will have moved out of the central cell, leaving gaps in that cell. But the gaps will be just offset by the parts of other cells which are now filled, and by periodicity, the states which have been vacated have just the same energy, momentum, etc., as those which have been filled. Thus all properties, as mean energy, mean momentum, etc., will be independent of time. In particular, the mean momentum, or current, of such a filled band of electrons is always zero. This means in the first place that the bound electrons of a metal, in which the bands are always full, do not contribute to the electrical conduction. But it also has significance as applied to nonconductors. It may be that in a solid all occupied bands are just filled with electrons, and that there is a gap in energy between the last occupied band and the first empty one. Then the net current must be zero, and the material is a nonconductor.[59] In a conductor, on the other hand, there must be a band of energy levels which is only partly filled. In the absence of a field, the electrons will arrange themselves to give no current, as we can see because they will fill the states of lowest energy, and since these are symmetrically placed with respect to the origin, they will take up a symmetrical distribution, with cancellation of the current. But as a field acts, the distribution will displace itself bodily in the direction of the field, with consequent production of a current. This current will continue to grow until electrons begin to undergo Bragg reflection; this will put a stop to the increase, leading to a maximum current, which will then decrease, change sign, and come back to zero after an interval of time.

In actual conduction, however, other complicating features will come in long before the complete cycle of which we have just spoken is complete. We have been dealing throughout with a perfect lattice, but actually at any temperature above the absolute zero the lattice will be distorted by temperature vibrations. These irregularities will produce scattering in which the mean momentum is not conserved, but decreases, and will therefore limit the increase of momentum, as we have already described in connection with the classical mechanics. The amount of scattering depends on the mean square deviation of the atoms from their regular periodicity, and hence is proportional to temperature, since the kinetic energy of the atoms, proportional to the square of their amplitudes of oscillation, is proportional to temperature. As a result, a resistance is introduced, proportional to temperature, as is observed experimentally. We shall not treat the detailed theory of conduction in the present report. It is a long and complicated subject, and even at the present time has not been treated perfectly satisfactorily, but the essentials are as we have described, and the theory has been given adequately in several comprehensive reports.*

26. Total energy in the wave-mechanical model

The other interesting result to be obtained from our solution for electronic motions in a metal is the evaluation of the total energy, with applications to cohesion, compressibility, heat of vaporization, etc. To discuss this important question, we shall proceed in a rather fundamental manner. So far, we have not set up anything of the nature of a wave function for the whole metal; we have used merely wave functions for single electrons in a periodic field. Let us first ask how we should compute the energy if we had the exact wave function at our disposal. Let there be N electrons, with coordinates from $x_1y_1z_1$ to $x_Ny_Nz_N$ and momenta from p_{x1} to p_{zN}. If the problem were completely solved, we should have a coordinate eigenfunction $u(x_1\cdots z_N)$, and a momentum eigenfunction $v(p_{x1}\cdots p_{zN})$. We should be able to get the total energy by averaging over either of these eigenfunctions: $E=\int u^*Hu\,dx_1\cdots dz_N=\int v^*Hv\,dp_{x1}\cdots dp_{zN}$, where H in the first expression is the operator in which momenta are replaced by differentiations, in the second in which coordinates are replaced by differentiations. It is simplest, however, though unconventional, to use the coordinate eigenfunction for finding the mean potential energy, and the momentum eigenfunction for the mean kinetic energy. For in this case no operators except simple multiplicative ones are necessary, and we can deal with easily understood averages over density distributions. Then, if T is the kinetic energy, U the potential energy we have for the mean values $\overline{T}=\int v^*Tv\,dp$, $\overline{U}=\int u^*Uu\,dx$, where T is a function of the p's, U of the x's, in the ordinary algebraical sense.

27. Kinetic energy

Let us consider the kinetic energy first, as the simpler function. We have, remembering that all our particles are electrons of mass m, $T=(1/2m)(p_{x1}^2+\cdots+p_{zN}^2)$. This is a sum of terms, one for each momentum, or three for each electron. Now on account of the antisymmetry which must be present in the wave function, on account of the exclusion principle, we must get the same mean kinetic energy for each electron. Hence we have the total kinetic energy equal to N times that of the first electron, or

$$\overline{T}=N\int\int\int[(p_{x1}^2+p_{y1}^2+p_{z1}^2)/2m]$$

$$\times dp_{x1}dp_{y1}dp_{z1}\int\cdots\int v^*v dp_{x2}\cdots dp_{zN}.$$

But now $\int\cdots\int v^*v\,dp_{x2}\cdots dp_{zN}$ is a function of $p_{x1}p_{y1}p_{z1}$, and gives the probability that electron number one be found in unit volume of momentum space at $p_{x1}p_{y1}p_{z1}$. Let us call it $G_1(p_{x1}p_{y1}p_{z1})$, or $G_1(p_1)$ for short. It is a function analogous to $F_1(x_1)$ for the coordinate distribution, which we have introduced in Section 5. Then

$$\overline{T}=N\int(p^2/2m)G_1(p)dp_xdp_ydp_z,$$

the average of the kinetic energy for a density distribution $NG_1(p)$ in a three-dimensional momentum space. In other words, we do not need the whole momentum eigenfunction for calculating the kinetic energy, but only this three-dimensional density.

Now we have found separate momentum eigenfunctions for each type of electron in the lattice, and corresponding to each we can set up the density by squaring. Adding these together for the different types of electron, we get a density

---

* A great many of the references in the Bibliography deal with conduction, which as a matter of fact has received more attention than most of the topics treated in this report. An excellent review of the present status of the theory is given in Sommerfeld and Bethe's article in the Handbuch der Physik, reference 98. It is hardly worth while enumerating the large number of papers on the subject, since they can be found from the titles in the Bibliography.

for the whole system, and it seems reasonable to postulate that this is a good approximation to the quantity $NG_1(p)$, the actual density in three-dimensional momentum space. If the electrons really were independent of each other, this would be just correct, even though we take account of the antisymmetry of the wave functions. Thus let the N wave functions be $v_1(p), \cdots v_N(p)$. Then to set up an antisymmetric function of the momenta of all N electrons, it is well known[2, 27] that we form the determinant

$$\frac{1}{\sqrt{N!}}\begin{vmatrix} v_1(p_1) & v_1(p_2) & \cdots & v_1(p_N) \\ v_2(p_1) & v_2(p_2) & \cdots & v_2(p_N) \\ \cdots & \cdots & \cdots & \cdots \\ v_N(p_1) & v_N(p_2) & \cdots & v_N(p_N) \end{vmatrix}.$$

If now we multiply this determinant by its conjugate, and integrate over all momenta but p_1, then as shown in Appendix VII, the result can be easily proved to be

$$(1/N)[v_1^2(p_1)+v_2^2(p_1)+\cdots+v_N^2(p_1)],$$

a result following from the orthogonality and normalization of the v's. But the determinant multiplied by its conjugate is just the probability of finding the first electron in unit volume of momentum space at p_1, the second at p_2, and so on, so that the integral over all momenta but p_1 is the probability of finding the first in unit volume at p_1, independent of the others, or our $G_1(p)$. Hence we have $NG_1(p)=v_1^2(p)+\cdots+v_N^2(p)$, the sum of densities of the separate electrons, as we have stated above. Actually, of course, the electrons are not independent of each other, so that this determinant is not a correct eigenfunction for the whole system. Nevertheless, in a function like the kinetic energy which depends only on a sum of terms one coming from each electron, and which contains no terms depending on mutual action of the electrons, it does not seem likely that neglect of these interactions should be of great importance. Thus we suspect that for computing kinetic energy, the sum of one-electron densities in momentum space gives a good density function. The same argument would apply to the use of this density for finding total current, as in discussing conduction. These arguments are reinforced by another slightly different consideration. The interactions which we are neglecting are the collisions of electrons with each other. Now in such a collision the kinetic energy and momentum of the individual electrons change, but on account of the conservation principles the sum for the two colliding electrons remains unchanged. Since we are interested only in total kinetic energy and total momentum or current, such collisions should not be of importance.

28. Potential energy

The potential energy, which we consider next, is of three parts, as we have already seen in Section 5: the potential energy of interaction between pairs of nuclei, between a nucleus and an electron, and between pairs of electrons. The first term involves no use of wave mechanics at all: between a nucleus of charge Z_1e and one of charge Z_2e at distance r, the energy is simply $Z_1Z_2e^2/r$. The second term, however, is more complicated. The term in the energy is

$$\sum_{Jk} -Z_Je^2/r_{Jk},$$

where r_{Jk} is the distance between the Jth nucleus and the kth electron. This term must be averaged over the electronic wave functions. On account of the equivalence of electrons, the result is N times the average for the first electron, or

$$N\int\int\int\sum_J(-Z_Je^2/r_{J1})dx_1dy_1dz_1$$

$$\times\int\cdots\int u^*u\; dx_2\cdot d_{zN}.$$

But just as for the momenta, $\int u^*u\; dx_2\cdot dz_N$ is the probability that electron one be in unit volume at $x_1y_1z_1$, which we have called $F_1(x_1y_1z_1)$. Then the mutual term in the energy is

$$N\int\sum_J(-Z_Je^2/r_J)F_1(x)dv,$$

which is exactly the interaction energy, as computed by ordinary electrostatics, between the nuclear charges Z_Je, and the average charge density $-NeF_1(x)$. This charge density is the usual one considered in quantum theory. And by arguments just like those used for the momentum, if we assume the electrons to be independent, and set up a wave function for the whole system in the form of a determinant, the result is $NF_1(x)=u_1^2(x)+\cdots+u_N^2(x)$, the sum of the densities of the separate electrons. Like the kinetic energy, this term involves no interactions between electrons, and we may assume that the one-electron eigenfunctions give a good approximation to the real density.

For the last sort of potential energy term, however, the interactions between pairs of electrons, the situation is quite different. The term in the energy is

$$\sum_{\text{pairs } ij} e^2/r_{ij},$$

and on account of antisymmetry the result will be the same as for the one term e^2/r_{12}, multiplied by $N(N-1)/2$, the number of pairs. That is, it will be

$$N(N-1)/2\int\int\int\int\int\int e^2/r_{12}dx_1\cdots dz_2 \times\int\cdots\int u^*u\; dx_3\cdots dz_N.$$

Here $\int\cdots\int u^*u\; dx_3\cdots dz_N$ is a function of the coordinates of the two electrons 1 and 2, and it gives the probability that simultaneously electron 1 will be in unit volume at x_1, electron 2 in unit volume at x_2, which we have called $F_2(x_1x_2)$. In terms of it, we have seen that the interaction energy is

$$N(N-1)/2\int\int e^2/r_{12}F_2(x_1x_2)dv_1dv_2.$$

A feature of the interaction energy which must be considered is the effect of the electron spin on it. The wave function u is really a function of the spins as well as the coordinates of the N electrons, and integrations over coordinates should include summations over spins. Thus $F_2(x_1x_2)$ is a function in which x_1 symbolizes three coordinates and a spin, x_2 three coordinates and a spin. And the whole function consists of four separate functions of coordinates, for the four cases (1) in which both electrons have + spin, (2) where electron 1 has +, electron 2 −, (3) electron 1−, 2+, (4) both −. Similarly $F_1(x_1)$ really consists of two functions, one for the case where the electron has + spin, the other for − spin. Integrating F_1 over the coordinates, without summing over spin, would lead to two numbers, the first giving the probability that the electron have + spin, the second that it have − spin. If the system as a whole had balanced spins, with no net moment, each of these numbers would be 1/2, but if the spins were unbalanced, the numbers would be different from 1/2, but always adding to unity. When now we divide F_2 by F_1, to get $-(N-1)eF_2(x_1x_2)/F_1(x_1)$, the charge density of other electrons at x_2 if there is an electron at x_1, we again have a function with four parts. It may be interpreted as follows: two parts give the probability that, if the first electron have + spin, the second have + and − spin, respectively, and lie within a certain volume element; the other two give similar probabilities for the case where the first electron has − spin. In general, all four functions can be different, and so give different contributions to the integral, on account of the exclusion principle, as a result of which the relative behavior of two electrons depends greatly on whether they have the same or opposite spin. Thus it becomes of importance to consider the spin. For the other terms in the energy, on the other hand, the kinetic energy and the interaction with the nuclei, it is easy to see that all effect of the spin cancels out, so that we do not have to consider it.

The two-electron probability $F_2(x_1x_2)$ which is needed for finding the interaction between electrons cannot be expected to be accurately determined by our method, which involves one-electron wave functions in which interactions are explicitly neglected, except for their average value.* A qualitative discussion of the distinction between the potential of the whole charge distribution, which at point x_1 is $\int -NeF_1(x_2)/r_{12}\,dv_2$, plus the nuclear potential, and the potential of the charge distribution minus one electron, when that electron is at x_1, which is

$$\int -(N-1)eF_2(x_1x_2)/F_1(x_1)(1/r_{12})dv_2,$$

---

* For discussion of this question, see reference 87, and also a forthcoming paper of Wigner and Seitz, Phys. Rev. **46**, 509 (1934). The author is indebted to Mr. Seitz for the privilege of seeing this paper before publication.

plus the nuclear potential, has already been given in Section 8. There it has been shown that the second differs from the first by terms arising first from the fact that the second potential was of a singly charged positive distribution rather than a neutral one, so that at a large distance from the distribution the potential fell off as e/R rather than exponentially, while at points inside the distribution the potential differed from that of the whole charge by having one electron's charge removed from the neighborhood of x_1. The difference arose secondly from polarization, image forces, and other effects arising from the electrical action of the electron at x_1 in producing a change in the distribution of the other electrons. These are all effects depending on electronic interaction, and from our methods of discussion by a one-electron problem we cannot improve on our previous qualitative treatment. There is only one exception to this, in which the one-electron functions are of some value: the case of the exclusion principle. It is recalled that if x_1 is in a region of the distribution where the density, as reckoned in the Fermi method, approaches its maximum allowable value, the other electrons of the same spin have a distribution which is like the distribution of the electrons as a whole, but with the charge removed from a hole surrounding x_1, just large enough to include the one electron which is at x_1, and therefore lacking from the distribution of other electrons. We can improve that approximation by a calculation using antisymmetric wave functions.

Let us assume that the wave functions $u_1 \cdots u_n$ correspond to $+$ spins, those from $u_{n+1} \cdots u_N$ to $-$ spins. Let $F_{1+}(x_1)$ and $F_{1-}(x_1)$ be the probability densities for the case of $+$ and $-$ spins, respectively, so that

$$\int F_{1+} dv_1 = n/N, \quad \int F_{1-} dv_1 = (N-n)/N.$$

Similarly let $F_{2++}(x_1x_2)$ be the probability density for the case where both electrons have $+$ spin, with similar interpretations for the other three components, so that

$$\int F_{2++} dv_1 dv_2 = n(n-1)/N(N-1),$$

$$\int F_{2+-} dv_1 dv_2 = n(N-n)/N(N-1) = \int F_{2-+} dv_1 dv_2,$$

$$\int F_{2--} dv_1 dv_2 = (N-n)(N-n-1)/N(N-1).$$

Then as shown in Appendix VII we have

$$F_{2++}(x_1x_2) = [1/N(N-1)] \sum_{i=1}^{n} \sum_{j=1}^{n} [u_i^*(x_1)u_i(x_1)u_j^*(x_2)u_j(x_2) - u_i^*(x_1)u_j(x_1)u_j^*(x_2)u_i(x_2)]$$

$$F_{2+-}(x_1x_2) = [1/N(N-1)] \sum_{i=1}^{n} \sum_{j=n+1}^{N} [u_i^*(x_1)u_i(x_1)u_j^*(x_2)u_j(x_2)]$$

$$F_{2-+}(x_1x_2) = [1/N(N-1)] \sum_{i=n+1}^{N} \sum_{j=1}^{n} [u_i^*(x_1)u_i(x_1)u_j^*(x_2)u_j(x_2)]$$

$$F_{2--}(x_1x_2) = [1/N(N-1)] \sum_{i=n+1}^{N} \sum_{j=n+1}^{N} [u_i^*(x_1)u_i(x_1)u_j^*(x_2)u_j(x_2) - u_i^*(x_1)u_j(x_1)u_j^*(x_2)u_i(x_2)].$$

Also we have

$$F_{1+}(x_1) = (1/N) \sum_{i=1}^{n} u_i^*(x_1)u_i(x_1), \qquad F_{1-}(x_1) = (1/N) \sum_{i=n+1}^{N} u_i^*(x_1)u_i(x_1).$$

We observe that F_{2++} and F_{2--} differ from F_{2+-} and F_{2-+} in that they have two sorts of terms. If the second sort were omitted, we should have $N(N-1)F_{2++}(x_1x_2)=(NF_{1+}(x_1))(NF_{1+}(x_2))$, with similar values for the other subscripts. That is to say, dividing through by $F_1(x_1)$, we should find that the probability of finding the second electron at x_2, when the first is at x_1, would equal just $F_1(x_2)$. Since there are no correction terms if the electrons have opposite spins, we have this result in that case: if electron 1 has one spin, electron 2 the opposite spin, and electron 1 is at x_1, the distribution of electron 2 will be just the fractional part, $(N-n)/N$, of the total charge distribution of electrons of this second spin. In other words, the electron 1 does not affect the distribution of charge of opposite spin at all. Of course, we know that this is not actually the case, as we have mentioned in Section 8, but we cannot expect to get any other result from this method of calculation. On the other hand, if electrons 1 and 2 have the same spin, there is a correction term in the density. Thus if they both have the + spin, and if the first electron is at x_1, the density of other electrons of + spin will not be $-neF_{1+}(x_2)$, but this diminished numerically by a correction

$$\frac{e\sum_{i=1}^{n}\sum_{j=1}^{n}u_i{}^*(x_1)u_j(x_1)u_j{}^*(x_2)u_i(x_2)}{\sum_{i=1}^{n}u_i{}^*(x_1)u_i(x_1)}.$$

This correction represents a charge density which integrates to a total of one electronic charge. For if we integrate over x_2, on account of orthogonality the only terms which will not integrate to zero will be those for which $j=i$, which will give unity, leading to a numerator equal to the denominator. It is now interesting to consider this correction charge density, for it evidently represents the hole which we have spoken about as resulting from the exclusion principle. We can see its qualitative behavior easily. When $x_2=x_1$, it becomes

$$e\sum_{j=1}^{n}u_j{}^*(x_2)u_j(x_2)=neF_{1+}(x_2),$$

so that the correction cancels the first term, and the density of other electrons is zero at the position of the first electron. On the other hand, when x_1 and x_2 are very different, the numerator proves to be a sum of products of oscillatory terms in practically random phase relations, which almost cancel, giving practically zero. This qualitative behavior agrees with what we have stated before, the correction charge being mostly removed from the neighborhood of x_1.

To get a quantitative calculation of the correction charge, we should have to take the actual wave functions which we have calculated, and work out the double summations. This could be done, though it would be a good deal of work. We can, however, get an idea of its value by assuming that the wave functions are plane waves, as in a field-free region. In this case, the function proves to depend only on the distance from x_1 to x_2, or r_{12}. It can then be shown[87] that the charge density of the removed charge at distance r_{12} from x_1 is

$$-e\left\{3\frac{\sin(r_{12}/d)-(r_{12}/d)\cos(r_{12}/d)}{(r_{12}/d)^3}\right\}^2$$

where $d=(V/3\pi^2n)^{1/3}$. This is shown in Fig. 6, curve A, and represents a concentrated density which can be easily shown to give a total charge of $-e$. In Fig. 6, A, we also give the radius R of the hole as calculated in Section 5, which is equal to $(3V/4\pi n)^{1/3}=1.92d$, and we see that the density attains half-value at just about this radius. With this good agreement, we may assume that the final calculation, using the correct wave functions, would not differ much from the value for plane waves, or in turn from the approximation we have used before.

We have now seen, at least in outline, how to calculate the total energy of a metal, or for that matter of any other atomic or molecular system. For the kinetic energy, we average $p^2/2m$ over a density function in a three dimensional momentum space, which can be obtained with fair accuracy by adding the densities of the momentum eigenfunctions of the various electrons. For the potential energy, we may summarize the process in a slightly different way from what we have used above. We first set up a charge density, consisting of the point charges of the nuclei, at their appropriate positions, and a volume distribution of charge of the electrons. This volume distribution can be found with fair accuracy by adding the densities of the coordinate wave functions of the various electrons. Next we set up a potential at each point of space. We really need two potential functions, however, one for the potential acting on a nucleus, the other for the potential acting on an electron. The potential acting on a nucleus is the potential, as computed by electrostatics, of all other nuclei, and of the whole electronic charge distribution above mentioned. The potential acting on an electron is the electrostatic potential of all the nuclei, and of all other electrons. This last term, the potential of the other electrons, is difficult to find, for it does not really follow from one-electron wave functions, and we have to make approximations and assumptions to get it. But we have described the way in which the charge distribution of all the electrons but one differs from that of all the electrons, and in a qualitative way at least we see how to compute it. Having found the resulting potential as a function of position in space, we can compute the potential energy by the electrostatic formula $(1/2)\int(\rho_n U_n+\rho_e U_e)dv$, where ρ_n and ρ_e are the charge densities due to nuclei and electrons respectively, and U_n and U_e are the potentials acting on nuclei and electrons.

29. Energy as function of internuclear distance

Having found how to compute the energy, we next ask what sort of function it is of the nuclear positions, for that is the real information we wish in the case of the metal. This cannot be answered until we know how the momentum and coordinate densities vary as the nuclei are moved. We consider only uniform compressions of the lattice, for these are the only distortions in which our assumptions of a regular lattice continue to be satisfied. Let us imagine first that the atoms are at infinite separation, all in their lowest stationary states. Then the density in momentum space will be of the nature shown in Fig. 17, a distribution going to zero at zero and at infinite momentum, with a maximum between. The density in coordinate space will be of the familiar sort for a single atom, in the neighborhood of each nucleus, going to zero in between. The total kinetic energy will be large, the total potential energy large and negative. The virial theorem tells us that the potential energy is twice as large numerically as the kinetic energy, so that the total energy is negative, and equal numerically to the kinetic energy. This total energy represents the amount necessary to pull all the electrons off all the atoms, removing all the particles of the crystal to an infinite distance from each other. All changes in energy which occur in the formation of the crystal, and in its compression by experimentally realizable amounts, change this total energy by only a small fraction of itself.

As the crystal is compressed, both momentum and coordinate densities change, with resulting change of energy. In discussing these changes, it is useful to keep in mind the virial theorem, which for the case of hydrostatic pressure tells us that $T+E=3pv$, where E is the total energy, T the kinetic energy, p the pressure, v the volume.[91] At infinite separation, the pressure is zero, so that $T=-E$, as we have mentioned above. At the equilibrium volume, the same thing is true. In between, however, the pressure is negative, so that the kinetic energy is less than $-E$, while at volumes less than the equilibrium volume the pressure is positive, and the kinetic energy is greater than $-E$. If one uses in addition the relation $p=-\partial E/\partial v$, one can eliminate the pressure, and obtain the kinetic energy, and hence also the potential energy, directly from the total energy. In Fig. 19 we see the form which T and U must have if the total energy is to show a minimum, as it is known to do experimentally. It is these resulting changes of kinetic and potential energy which must be explained from our model.

As the atoms approach, the first effect will be on the outer electrons. At infinite separation, these are atomic electrons, with momentum eigenfunctions which are large for a considerable range out from the origin in momentum space. At shorter distances, however, we have seen that these outer electrons approach the metallic correspondence, in which the momentum eigenfunction is large only at one point in momentum space, generally for smaller momentum than the maximum at infinite separation. Taking all these outer electrons together, the general effect will be at first a decrease of mean momentum, and hence of kinetic energy. This is what we should expect from Fig. 19. As the atoms are still further

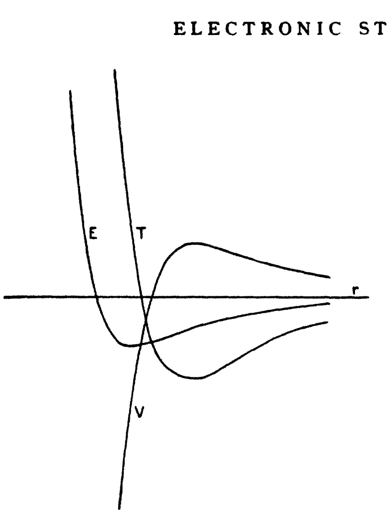

FIG. 19. Kinetic, potential, and total energy by the virial theorem.

compressed, however, the effect of promotion comes in for the outer electrons. In consequence of this effect, their kinetic energy increases. This is partly compensated, however, by the inner shells of electrons, which at the same time are decreasing their kinetic energy, preparatory to becoming free electrons. The net result, however, is an increase of kinetic energy, as Fig. 19 shows. This change in kinetic energy can be briefly described as follows. The process of creating free electrons is one of decrease of kinetic energy, for the electrons become concentrated in the region between atoms, where the potential energy is high, and consequently the kinetic energy is low. This goes on as the atoms are brought together. It is counteracted, however, by the process of compression of the free electrons already produced, which on account of the exclusion principle is accompanied by an increase of kinetic energy. At large distances, the creation of the free electrons is the more important process, while at smaller distances the number does not vary so fast, but the increase of energy on account of compression becomes the governing factor.

The changes in potential energy as the atoms approach can also be understood from the model. The coordinate wave functions of the outer electrons become rapidly larger in the region between the atoms, as they come together, and as a result the charge density between the atoms increases, with compensating decrease in the immediate neighborhood of the nuclei. Negative charge consequently is removed from the region of low potential energy about the nuclei, and transferred to the region of higher potential energy, with increase of the total potential energy. This increase, for distances of separation greater than the equilibrium distance, almost balances the decrease in kinetic energy. For both these effects come from a shift of charge distribution within the atom, and by the variation principle a first order change in the wave function, and charge distribution, makes only a second order change in the energy, though it can make a first order change in potential and kinetic energy separately. The net effect of change in potential and kinetic energy, however, must be a decrease of energy, or an attractive force, as we can see from the virial theorem. For we may write this $(3/2)pv=(T+E)/2=T+U/2$, where U is the potential energy. Let the kinetic and potential energy at infinite separation be T_0 and U_0, respectively, where we know that $T_0+U_0/2=0$. Then we have $(3/2)pv=(T-T_0)+(U-U_0)/2$. But if $(T-T_0)+(U-U_0)$ is a small quantity compared with the change of either kinetic or potential energy separately, we have approximately $(3/2)pv=(T-T_0)/2$, and since this is negative, the pressure is negative, indicating attraction.

As the atoms approach even closer, however, it is no longer true that the change in total energy is small compared with the change of kinetic or potential energy, so that the conclusion we have just stated no longer holds. Once the charge of the free electrons is pulled out into the region between the atoms, the principal further effect on the potential energy comes from the change in the potential rather than the charge distribution. In Fig. 1 were plotted values of U_e, the potential acting on an electron, as a function of the distance from a nucleus, within the cell surrounding that nucleus. There is, of course, an additive constant in U_e which is arbitrary, but it is most convenient to choose this constant so that the inner part of the curve is independent of internuclear distance. This has the result that U_n remains constant, so that the part $(1/2)\int \rho_n U_n\, dv$

is independent of distance, and the only change in potential energy comes from the electronic term $(1/2)\int \rho_e U_e \, dv$. It also leads to x-ray energy levels for the one-electron problem which are independent of internuclear distance, which is convenient. With this convention, we see that $-U_e$ decreases more and more in the region between atoms, as the atoms are squeezed together, and consequently that the potential energy must decrease more and more, as Fig. 19 shows that it must. This decrease of $-U_e$ is also shown in Fig. 9, where it is given by curve AA. This decrease of $-U_e$ comes from two sources. First is an effect of pure electrostatics. As the negative charge surrounding the nucleus is crowded into smaller and smaller volume, the negative potential of the whole charge distribution at the surface of a cell must decrease, if that near the nucleus remains fixed. For at any given distance from the nucleus, the attractive force on a negative charge, or repulsive force on a positive, will be less with the concentrated charge than otherwise, since there will be more negative charge within a sphere of the corresponding radius, more nearly cancelling the positive charge of the nucleus. Thus the U_e curve will be less steep for small internuclear distance than for large, resulting in the type of curve shown in Fig. 1. The second effect is that of the exclusion effect. The result of this effect, we remember, is a negative term in the potential energy, equal to the interaction of an electron with a distribution of charge, of net amount one electron, removed from its neighborhood. Further, the size of this hole is smaller as the density of charge becomes greater, so that the term in the energy is correspondingly greater. Thus as the atoms are squeezed closer together, and the free electron density becomes greater, this effect increases, reducing the value of $-U_e$ at points between the atoms. This effect, and the purely electrostatic one, are of the same order of magnitude. As the distance of separation becomes quite small, both these effects on the potential energy tend to go inversely as the internuclear distance, just as they would if they represented the potential energy of interpenetrating lattices of positive and negative charge, as in an ionic crystal. On the other hand, the kinetic energy tends toward the free electron behavior, being inversely as the square of the internuclear distance. Thus the repulsive kinetic energy will outbalance the attractive potential energy at small distances, producing net repulsion.

30. Relation of one-electron energies to total energy

We have now given a qualitative interpretation of the energy of the lattice, the energy curve which leads to the calculation of the grating space, or equilibrium distance, the heat of formation of the crystal, the compressibility, etc. The terms which enter into this are all things which can be computed, when once the electron wave functions are known, and for the cases for which they have been worked out, the agreement with experiment is satisfactory. There is another point of view regarding the total energy, however, which is also correct, and which is worth considering, since it makes closer connection with the one-electron functions. We have already seen that each of these one-electron wave functions has an energy level which is a function of internuclear distance, and it is natural to ask what relation these energy levels have to the total energy of the crystal. We may reasonably expect a connection, for it is known that in atomic models, the one-electron energies are rather accurately equal to the ionization potentials of the corresponding electrons. In the method which we have used for computing the energy of the crystal, we note in the first place that the kinetic energy of the whole crystal is assumed to be the sum of the kinetic energies of the various one-electron problems, since the momentum density is taken to be the sum of the densities for the various electrons. Also the potential energy of interaction between nuclei and electrons, for the whole crystal, is the sum of these terms for the various electrons, for a similar reason. The potential energy of interaction between pairs of electrons, however, involves a different situation. For a one-electron problem, the potential energy is the sum of the potential energy in the field of the nuclei, plus that in the field of the other electrons. This latter term is the integral of the charge density of the one electron, times the potential of all other electrons, integrated over all space. When we add these terms together for all electrons, we have the integral of the total charge density times the potential of all other electrons. But this is just twice the correct interaction energy, for it counts each pair of electrons twice, once in determining the charge, once in determining the density. The correct interaction energy is 1/2 the integral of total charge density times the potential of other electrons. Thus we have the following result: the sum of the one-electron energies of all electrons equals the correct kinetic energy for the whole crystal, plus the correct interaction energy between electrons and nuclei, plus twice the interaction energy between pairs of electrons. Conversely, the total energy of the crystal equals the sum of the one-electron energies, minus the interaction energy between pairs of electrons, plus the interaction energy between pairs of nuclei. This statement is true to the same approximation to which our whole calculation of energy is correct.

For a single atom, the interaction between pairs of nuclei is not present. We see then that for an atom the total energy equals the sum of one-electron energies (which in this case are approximately the ionization potentials, with negative sign), minus the interaction energy between electrons. It is easily understandable that the total energy is lower than the sum of one-electron energies, for each of these represents approximately the energy required to remove an electron from the neutral atom, while in completely breaking up the atom, successive electrons are removed from more and more highly charged ions, requiring much more work. Now for the whole crystal at large separation, the one-electron energy levels are almost exactly what they are for the separate atoms, only there are N times as many levels of each type. Thus the sum of one-electron energies is just N times that for a single atom. The interaction energy between pairs of electrons is of two parts: first, the interactions between pairs in which both electrons are on the same atom, of which the sum is just N times that for a separate atom; secondly, interactions between pairs in different atoms. This latter interaction is that between the net negative charge of each atom and of each other atom, and on account of the spherical distribution of each atom, it just cancels the interaction between the nuclei, leaving the total energy as N times that of a single atom, as we should expect.

As the atoms approach, we are interested only in the change of total energy. We have seen the nature of the change in one-electron energy. For the inner electrons there is no change, while the energy of the outer electrons decreases, goes through a minimum, then increases again, as we have seen in Fig. 9. The outer electrons then contribute the only change to the sum of the one-electron energy, resulting in a minimum of energy at some equilibrium distance. Let us see how the other correction terms affect this energy. We know that each cell remains electrically neutral and approximately spherical. Thus the interaction energy between the electrons of one cell and those of another continues, even in the compressed crystal, to cancel almost exactly the repulsion between the nuclei of the corresponding cells. The only corrections of any magnitude, then, come from the change of the interaction energy between electrons of the same cell, as the cell is compressed. This correction would be small for a univalent element like sodium, in which it involves only the interaction between the valence electron and the inner electrons, but larger for other elements, where it involves the interaction between pairs of valence electrons. Thus for sodium the sum of one-electron energies should give fairly accurately the variation in the total energy, which calculation shows it to do, while for atoms with several valence electrons, there should be additional terms in the energy, resulting in tighter binding than the sum of one-electron energies would indicate. In any case, however, we see that the minima in the one-electron energies are closely connected with the minima in total energy of the crystal.

31. Total energy and magnetic properties

A final question connected with the energy is that of the magnetic properties of the metal, particularly ferromagnetism. We have assumed that the lowest state of the metal was always that in which the spins exactly cancelled, so that the metal was nonmagnetic. Empirically we know, however, that ferromagnetic metals appear to have a magnetic moment in their normal state, with spins parallel to each other. We must ask how this can be interpreted. At first sight it would appear impossible, for the kinetic energy certainly has its lowest value when each of the lowest energy levels is occupied by two electrons of opposite spin. The potential energy, however, depends on the spin, and under certain circumstances it can be lower for states of nonvanishing total spin, by an amount sufficient to make these states actually lower than those of balanced spin. We remember that the potential U_e in which an electron finds itself is the potential of the distribution $-(N-1)eF_2/F_1$, so that it really consists of two parts: the potential in which an electron of + spin finds itself, and a similar potential for an electron of − spin. In the case of balanced spins, these are equal, but if for example more electrons have + spin than −, the hole surrounding an electron with + spin will be smaller, on account of their larger density, than for a − spin, and correspondingly U_e for the + spin will correspond to a lower potential energy than for the −. This is a situation in which there will be a lowering of the total potential energy if electrons of − spin change over to + spin, with a minimum potential energy if all electrons have their spins parallel. They will do this, then, if and only if the change is not accompanied by an even greater increase of kinetic energy. In a case like sodium, for example, the change in kinetic energy certainly will be greater than in potential, at ordinary distances of separation, as one can compute from the relative sizes of the Fermi energy and the potential energy terms. But in iron, for example, the situation is quite different. There we have the whole set of 3*d* and 4*s* levels, available for outer electrons. These levels are only about half filled with electrons. Further, since the 3*d*'s overlap so little, their energy band will be only slightly spread out, so that it will involve only slight increase of kinetic energy if levels are filled with only one + electron, instead of with a + and a −. These 3*d*'s, of course, do not take much part in the cohesion, since they do not overlap enough to result in a large free electron density between atoms, but they are exactly suited to produce a permanent magnetic moment. From this example, we see that the requirement for ferromagnetism is the existence of only partially filled shells, so as to allow the spins of these electrons to be all parallel, and far enough inside the atom so that the resulting energy bands are only slightly broadened, so that the increase of kinetic energy is far less than for the outermost electrons. These conditions are fulfilled in the iron group of elements better than in any other elements of the periodic table, consistent with the ferromagnetic properties of these elements.[38]

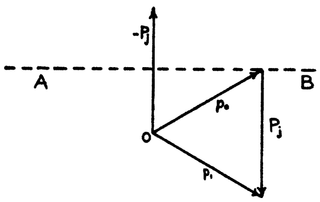

Fig. 20. Momentum vectors for Bragg reflection.

The connection between ferromagnetism and the interaction of electron spin resulting from the exclusion principle was first pointed out by Heisenberg.[22] His calculations, and most of the subsequent theoretical work on ferromagnetism, have been based on an analogy to the method of Heitler and London as used for molecular structure.[36] This method is not based on the use of solutions of the one-electron problem for the periodic lattice at all, but instead starts with the electrons as if in the isolated atoms, with wave functions concentrated about a single nucleus. That method had one considerable advantage over the earlier forms of the type of theory presented in the present report: it gave a good account of electron interactions, of the probability of finding one electron at a given distance from another, and of the relation of this to electron spin. Since these are essential features in the treatment of both ferromagnetism and valence, the method has found much application there. At the time the applications were made, the methods of incorporating electron interactions in the present form of the theory, by means of the distribution functions F_2 and F_1, were not well worked out. The writer believes that in the future these latter methods will prove to be of more value than the method of Heitler and London, not only for discussing metals in general, but for ferromagnetism and for molecular structure as well. For metals, they have great advantages in problems of electrical conductivity; they now prove to be more suitable for total energy and cohesion. Further, they have a great technical advantage, in that they deal with orthogonal one-electron functions, which the Heitler and London method does not. The writer commends the general method outlined in this report, with one-electron functions, and separate discussions of kinetic and potential energy by use of the momentum and coordinate eigenfunctions and the virial theorem, as the most hopeful line of attack at the present time on the structure of solids and of molecules.

APPENDIX I. MILLER INDICES AND RECIPROCAL LATTICE

The Miller indices h, k, l of a set of planes are determined as follows, for a simple cubic lattice: of a set of parallel equidistant planes including all atoms, of the sort we have described, we choose the one which comes nearest the origin without cutting it, given by our equation $(P_j/h)\cdot q=1$. (Do not confuse the h, Planck's constant, with the h Miller index). We take the intersections of this plane with the x, y, z axes, h/P_{jx}, h/P_{jy}, h/P_{jz}, and then define h, k, l, by setting these three intersections equal to d/h, d/k, d/l, where d is the lattice spacing. That is, $h=(d/h)P_{jx}$, $k=(d/h)P_{jy}$, $l=(d/h)P_{jz}$. Now we shall prove that h, k, l must be integers, which means that P_{jx}, P_{jy}, P_{jz} are arranged on a lattice of spacing h/d, reciprocal to the spacing of the coordinate lattice. To do this, we need only prove that if and only if h, k, l are integral, every atom of the crystal lies on one of the planes. The equations of the planes become $(hx+ky+lz)/d=n$. Let $x=qd$, $y=rd$, $z=sd$, where q, r, s are integers; this is the condition that they lie on a simple cubic lattice. Then the equations become $(hq+kr+ls)=n$. But obviously if h, k, l, q, r, s are integers, n is also an integer, proving that for every lattice point the equation of one of the planes is satisfied. Conversely, if at least one of the indices h, k, l is not an integer, it is obvious that q, r, s can be given integral values which will make n non-integral, showing that some lattice points would not lie on the planes, so that non-integral indices are not allowed.

In a body centered lattice containing atoms at the corners of a cube of side d, and one at the center of each cube, we evidently have q, r, s all integers (for the corners of the cube), or all half integers (for the centers). Then if $hq+kr+ls=n$, an integer, h, k, l must all be integers (to satisfy the condition for integral q, r, s), but furthermore $h+k+l$ must be an even number (so that, for the half integral q, r, s, the factor 1/2 may be multiplied by an even number, and give an integral n). Thus the reciprocal lattice will have points only for $h+k+l=$ even number, or in the momentum space $(P_{jx}+P_{jy}+P_{jz})=$ even number $x(h/d)$. But this is the condition for a face centered lattice in momentum space, the faces being of width $2h/d$, with points in the center of each face. Thus the reciprocal lattice to a body centered space lattice is a face centered momentum lattice.

Similarly the reciprocal to a face centered space lattice is a body centered momentum lattice. Let the atoms be at the corners of a cube of side $2d$, and at the center of each face. Then q, r, s are all integers, and $q+r+s=$ an even integer. We can then satisfy $hq+kr+ls=n$ with h, k, $l=$ integers, or h k, $l=$ half integers, resulting in a body centered momentum lattice, of side h/d.

APPENDIX II. CONDITIONS FOR BRAGG REFLECTION

Let the initial wave be represented by exp $[(2\pi i/h)(p_0\cdot q)]$ the final wave by exp $[(2\pi i/h)(p_1\cdot q)]$. Then for Bragg reflection in the first place these must have the same wave-length, giving $p_1^2=p_0^2$. Secondly, they must both be in the same phase at the lattice planes $(P_j/h)\cdot q=n$, in order to have interference. This demands that the exponentials be equal on these planes, or $p_1\cdot q=p_0\cdot q+nh$, where n is an integer, or $[(p_1-p_0)/h]\cdot q=n$. For these planes to be the same as those represented above, we must have $p_1-p_0=P_j$, or p_1, the reflected wave momentum, $=p_0+P_j$. This is shown in Fig. 20, where we also see that any vector p_0 from the origin O of momentum space to the plane AB, perpendicular bisector of $-P_j$, satisfies this condition.

APPENDIX III. PROPERTIES OF MOMENTUM EIGENFUNCTIONS

The properties which we shall describe can be put in various different forms, depending on whether u and v are continuous or discontinuous functions, and on whether they are confined to small regions of space, or extend throughout all space. We shall state them in a convenient form for the case we use most, namely, that in which u is a continuous function, extending throughout space in such a way that u^*u repeats itself periodically, and v is a discontinuous function defined at discrete points, and confined to the part of momentum space surrounding the origin. In this case it is convenient to normalize v to unity: $\sum_p v^*(p)v(p)=1$. The coordinate function u, however, cannot be normalized to unity, but on account of its periodicity, we can set its average value over many cells equal to unity: $\lim\left[\int u^*(q)u(q)dv/\int dv\right]=1$, where the limit is taken as the volume of integration becomes larger and larger. Let us then assume that

$$u(q)=\sum_p v(p)\exp\left[(2\pi i/h)p\cdot q\right].$$

We can show that if v is normalized in the sense described, u is also normalized. To do this, multiply u by its conjugate, and average over coordinates, obtaining the relation

$$\frac{\int u^*(q)u(q)dv}{\int dv}=\sum_p\sum_{p'}v^*(p)v(p')\frac{\int\exp\left[(2\pi i/h)(p'-p)\cdot q\right]}{\int dv}.$$

But now the average value of $\exp\left[(2\pi i/h)p\cdot q\right]$ is zero unless $p=0$, in which case it is unity. Hence the right side of the expression above becomes $\sum_p v^*(p)v(p)=1$, proving the normalization. Similarly, multiplying the equation for $u(q)$ by $\exp\left[(-2\pi i/h)p'\cdot q\right]$, and averaging, we obtain the relation

$$v(p)=\lim\left\{\int u(q)\exp\left[(-2\pi i/h)p\cdot q\right]dv\right\}\Big/\int dv.$$

Using the definition of u in terms of v given above, we can prove that the v so determined satisfies the same difference equation given in the text, so that it is the same function used there. To do this, substitute u in Schrödinger's equation. We have

$$\begin{aligned}\left[(-h^2/8\pi^2m)^2+U\right]u(q)&=\sum_{p'}v(p')(p'^2/2m+U)\exp\left[(2\pi i/h)p'\cdot q\right]\\&=\sum_{p'}v(p')(p'^2/2m)\exp\left[(2\pi i/h)p'\cdot q\right]+\sum_{p'}\sum_{P}v(p')W(P)\exp\left[(2\pi i/h)(p'-P)\cdot q\right]\\&=Eu(q)=E\sum_{p'}v(p')\exp\left[(2\pi i/h)p'\cdot q\right],\end{aligned}$$

where we remember the expansion of U in Fourier series. Now we multiply both sides of the equation by $\exp\left[(-2\pi i/h)p\cdot q\right]$, and average over coordinates. The result is $(p^2/2m)v(p)+\sum_P W(P)v(p+P)=Ev(p)$, the same equation given in the text.

The equation which we have just written may be given more explicitly, for the periodic lattice,

$$\left[(p_0+P_j)^2/2m-E_n\right]v_n(p_0+P_j)+\sum_{P_k}W(P_k)v_n(p_0+P_j+P_k)=0,$$

where p_0 represents the value in the central cell for which the function is defined, and where n refers to the nth stationary state. The orthogonality of the v's is easily proved. Thus let the same equation be written for $v^*_m(p_0+P_j)$:

$$\left[(p_0+P_j)^2/2m-E_m\right]v_m^*(p_0+P_j)+\sum_{P_k}W^*(P_k)v_m^*(p_0+P_j+P_k)=0.$$

Multiply the first equation by $v_m^*(p_0+P_j)$, the second by $v_n(p_0+P_j)$, and subtract. The result is

$$(E_m-E_n)(v_m^*(p_0+P_j)v_n(p_0+P_j))+\sum_{P_k}W(P_k)v_m^*(p_0+P_j)v_n(p_0+P_j+P_k)$$
$$-\sum_{P_k}W^*(P_k)v_m^*(p_0+P_j+P_k)v_n(p_0+P_j)=0.$$

Now sum over P_j. The last term can then be rearranged, by noting that $W^*(P_k)=W(-P_k)$, if the potential is real, and by letting $p_0*P_j+P_k=p_0+P_j'$, $p_0+P_j=p_0+P_j'+P_k'$, so that $P_k=-P_k'$. When this is done, the last summation is seen to be equal to the preceding summation, so that they cancel leaving

$$(E_m-E_n)\sum_{P_j}v_m^*(p_0+P_j)v_n(p_0+P_j)=0.$$

That is, if $m\neq n$,

$$\sum_{P_j}v_m^*(p_0+P_j)v_n(p_0+P_j)=0,$$

the orthogonality condition. We have assumed that if $m=n$, the normalization condition

$$\sum_{P_j}v_n^*(p_0+P_j)v_n(p_0+P_j)=1$$

is satisfied, which can be done by proper choice of the multiplicative constants in v.

The orthogonality and normalization conditions stated above can be interpreted in terms of a function space. Let us first consider only a finite number of values of P_j, large enough to include many cells surrounding the origin in momentum space. We could take these as a finite number of rectangular coordinates, and plot the values of $v_n(p_0+P_j)$ for different P_j's as the components of a vector in this space. The orthogonality and normalization conditions show that this is a unit vector, and that if we draw the vectors for different values of n the various vectors are normal to each other. Of course, with a finite number of cells, we have only a finite number of stationary states and values of n, in fact just as many as there are cells. These vectors $v_n(p_0+P_j)$ for different n's then form a set of orthogonal unit vectors in the space, which could be used as coordinates as well as the P_j's. In terms of them, $v_n(p_0+P_j)$ for different n's, but a given P_j, represent the components of a vector representing P_j in this new set of coordinates. We can now express the condition of orthogonality and normalization of these P_j's, in the new coordinates: $\sum_n v_n^*(p_0+P_j)v_n(p_0+P_k)=1$ if $P_j=P_k$, $=0$ otherwise. This is the completeness relation of which we speak in Section 22. Passing to the limit of infinitely many cells, it continues to hold in the present case, though not in all sorts of problems.

APPENDIX IV. ENERGY DISCONTINUITIES BY PERTURBATION METHOD

If either the metallic or free electron correspondence is good, that is an indication that $v(p)$ is much larger for one value of p than for any other. Let this one be p_0. But on the surfaces of zones in momentum space, one, and generally just one, other component also becomes large. Let this be p_0+P_j, where the surface in question is then the bisector of $-P_j$. Then in the difference equations for $v(p)$, there will be just two large components. To a first approximation, we can replace the infinite set of difference equations by the two simultaneous equations involving only these two v's:

$$(p_0^2/2m-E)v(p_0)+W(P_j)v(p_0+P_j)=0, \qquad W(-P_j)v(p_0)+[(p_0+P_j)^2/2m-E]v(p_0+P_j)=0.$$

Setting the determinant of these equations equal to zero, we have a quadratic for E, whose solutions are

$$E=\tfrac{1}{2}(p_0^2/2m)+((p_0+P_j)^2/2m)\pm\{\tfrac{1}{2}[(p_0^2/2m)-((p_0+P_j)^2/2m)]^2+W^*(P_j)W(P_j)\}^{1/2},$$

where we have used the relation $W(-P_j)=W^*(P_j)$. Solving for the case where $p_0^2/2m-(p_0+P_j)^2/2m \gg W^*(P_j)W(P_j)$, so that we are far from the edge of the zone, we have the same solution obtained by the second order perturbation method. On the other hand, when

$$p_0^2/2m-(p_0+P_j)^2/2m \ll W^*(P_j)W(P_j),$$

near the edge, we have

$$E=(1/2)\{(p_0^2/2m)+[(p_0+P_j)^2/2m]\}\pm\{W^*(P_j)W(P_j)\}^{1/2}$$
$$\pm(1/8)\{(p_0^2/2m)-[(p_0+P_j)^2/2m]\}^2/\{W^*(P_j)W(P_j)\}^{1/2},$$

showing that just at the edge the discontinuity in energy is $2[W^*(P_j)W(P_j)]^{1/2}=2|W(P_j)|$, and permitting us to derive the shape of the curve near the discontinuity.

Appendix V. Stationary States for Finite Crystals

A wave $\exp[(2\pi i/h)p_0\cdot q]$ cannot be a solution for a finite crystal unless it satisfies boundary conditions. For instance, for a rectangular crystal bounded by the planes $x=0$, $x=X$, $y=0$, $y=Y$, $z=0$, $z=Z$, we must expect that the wave should take on the same value over each pair of opposite faces. The demands that $p_{0x}X/h=\text{integer}$, $p_{0y}Y/h=\text{integer}$, $p_{0z}Z/h=\text{integer}$, so that the allowable points in momentum space form a rectangular lattice, of sides h/X, h/Y, h/Z. The volume of momentum space per point is then $h^3/XYZ=h^3/V$, if V is the volume of the crystal. With other shaped crystals, the detailed arrangement of points in momentum space would be different, but the number of points per unit volume would be the same.

For a simple cubic crystal, the volume of a cell of momentum space is h^3/d^3, if d is the lattice spacing. This then contains $(h^3/d^3)(V/h^3)=(V/d^3)$ allowable points p_0. But V/d^3 is just the total number of coordinate cells of the crystal, and since there is one atom per cell, it is the number of atoms N, which was to be proved.

For a body centered cubic crystal, the reciprocal lattice is face centered, of side $2h/d$, containing four points in the cube $(2h/d)^3$, or one point per volume $2h^3/d^3$. This then is the volume of one cell of momentum space. Thus there are $2(h^3/d^3)(V/h^3)=2V/d^3$ allowable points per cell. But the unit cell in coordinate space, of volume d^3, contains two atoms, so that the crystal contains $2V/d^3$ atoms, again the same as the number of allowable points.

For a face centered crystal, the cell of momentum space contains 2 points in a volume h^3/d^3, so that the cell has the volume $(1/2)h^3/d^3$, and contains $(1/2)(h^3/d^3)(V/h^3)=(1/2)V/d^3$ values of p_0. The unit cell in coordinate space, of volume $(2d)^3$, contains four atoms, so that the crystal contains $4V/(2d)^3=(1/2)V/d^3$ atoms, again the same as the number of allowable points.

Appfndix VI. Relation Between Average Momentum and Energy

Multiply the equation for $v(p)$ by $v^*(p)$, and sum over all values of p. This gives us the following value for the total energy:

$$E=\sum_{P_k}v^*(p_0+P_k)[(p_0+P_k)^2/2m]v(p_0+P_k)+\sum_{P_j}\sum_{P_k}v^*(p_0+P_k)W(P_j)v(p_0+P_k+P_j).$$

Similarly for the average velocity in the x direction, $\dot{x}$, we have

$$\dot{x}=(1/m)\sum_{P_k}v^*(p_0+P_k)(p_{0x}+P_{kx})v(p_0+P_k).$$

Now let us differentiate E with respect to the x component of p_0, p_{0x}:

$$\partial E/\partial p_{0x}=\sum_{P_k}v^*(p_0+P_k)[(p_{0x}+P_{kx})/m]v(p_0+P_k)+\sum_{P_k}[(p_0+P_k)^2/2m]\left\{\frac{\partial v^*(p_0+P_k)}{\partial p_{0x}}v(p_0+P_k)\right.$$
$$\left.+v^*(p_0+P_k)\frac{\partial v(p_0+P_k)}{\partial p_{0x}}\right\}+\sum_{P_j}\sum_{P_k}W(P_j)\left\{\frac{\partial v^*(p_0+P_k)}{\partial p_{0x}}v(p_0+P_k+P_j)+v^*(p_0+P_k)\frac{\partial v(p_0+P_k+P_j)}{\partial p_{0x}}\right\}.$$

The first term is just $\dot{x}$, and we shall show that the other terms vanish, proving the theorem. To prove that the second and third terms cancel, we note that

$$[(p_0+P_k)^2/2m]v(p_0+P_k)+\sum_{P_j} W(P_j)v(p_0+P_k+P_j)=Ev(p_0+P_k).$$

As in Appendix III, we rearrange the second term of the double summation, letting $p_0+P_k+P_j=p_0+P_k'$, $p_0+P_k=p_0+P_k'+P_j'$, $P_j=-P_j'$, so that this term becomes

$$\sum_{P'_j}\sum_{P'_k} W^*(P_j')v^*(p_0+P_k'+P_j')\partial v(p_0+P_k')/\partial p_{0x}.$$

We then drop the primes, and note that

$$[(p_0+P_k)^2/2m]v^*(p_0+P_k)+\sum_{P_j} W^*(P_j)v^*(p_0+P_k+P_j)=Ev^*(p_0+P_k).$$

The two terms which we wish to vanish now can be rewritten as

$$E\sum_{P_k}\left\{\frac{\partial v^*(p_0+P_k)}{\partial p_{0x}}v(p_0+P_k)+v^*(p_0+P_k)\frac{\partial v(p_0+P_k)}{\partial p_{0x}}\right\}.$$

But differentiating the normalization relation

$$\sum_{P_k} v^*(p_0+P_k)v(p_0+P_k)=1$$

with respect to p_{0x} shows that the bracket above vanishes, which we wished to prove.

Appendix VII. Properties of Determinantal Eigenfunctions

Let the coordinate eigenfunctions of a single electron be $u_i(x_i)$. If the eigenfunction is associated with + spin, and the electron spin is s_i, which can be + or −, the eigenfunction with spin can be written $u_{i+}(x_is_i)$, which by definition equals $u_i(x_i)$ if s_i is +, but is zero if s_i is −. Then if the eigenfunctions from 1 to n are associated with + spin, those from $n+1$ to N with − spin, an antisymmetric eigenfunction for the whole system is[2, 27]

$$\frac{1}{\sqrt{N!}}\begin{vmatrix} u_{1+}(x_1s_1) & \cdots & u_{n+}(x_1s_1) & u_{(n+1)-}(x_1s_1) & \cdots & u_{N-}(x_1s_1) \\ u_{1+}(x_2s_2) & \cdots & u_{n+}(x_2s_2) & u_{(n+1)-}(x_2s_2) & \cdots & u_{N-}(x_2s_2) \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ u_{1+}(x_Ns_N) & \cdots & u_{n+}(x_Ns_N) & u_{(n+1)-}(x_Ns_N) & \cdots & u_{N-}(x_Ns_N) \end{vmatrix}.$$

To get $F_2(x_1s_1,\ x_2s_2)$, we multiply this function by its conjugate, and integrate and sum over all coordinates and spins from 3 to N. Then F_{2++} is the value in case s_1, s_2 are both +, etc. Now the eigenfunction is a sum of terms, each formed by taking one factor from the first row, another from another column of the second row, and multiplying by the subdeterminant obtained by crossing out the two rows and columns already used. The conjugate is a similar sum of terms, and in multiplying them and integrating over all coordinates from 3 to N, we integrate products of such subdeterminants. But on account of the orthogonality of the u's, such a product of subdeterminants integrates to zero, unless it is the product of a subdeterminant and its own conjugate, in which case on account of normalization it integrates to $(N-2)!$ Thus the only nonvanishing terms are those in which the factors involving x_1 and x_2, in both the eigenfunction and its conjugate, come from the same two rows and columns, and each such term is multiplied by the factor $(N-2)!/N!=1/N(N-1)$. There are just two sorts of such terms, one of the form $u_i^*(x_1s_1)u_i(x_1s_1)u_j^*(x_2s_2)u_j(x_2s_2)$, where we take just the same terms from both the determinant and its conjugate, and the other $-u_i^*(x_1s_1)u_j(x_1s_1)u_j^*(x_2s_2)u_i(x_2s_2)$, where the two rows are interchanged in the conjugate with respect to what they are in the original deteminant, with consequent change of sign. If now i and j both symbolize quantum numbers associated with the same spin, both terms occur, whereas if they have opposite spin, the second form is absent, since in that case it is impossible for s_1 to have simultaneously the spin necessary to give nonvanishing values to both u_i and u_j. Thus we arrive at the formulas for F_2 given in Section 28.

To get $F_1(x_1)$, we integrate $F_2(x_1x_2)$ over x_2, and sum over s_2, noting that the second form of term contributes nothing, on account of orthogonality, except when $i=j$, when it cancels the first term. The first term then gives $(N-1)$ times the sum $[1/N(N-1)]\sum_i u_i^*(x_1)u_i(x_1)$, giving the values of $F_1(x_1)$ in Section 28.

By replacing the coordinate eigenfunctions $u(x)$ by momentum eigenfunctions $v(p)$, we get the momentum probability (G_1p_1), used in Section 27, in exactly analogous ways.

Bibliography

1916

1. H. A. Lorentz, *The Theory of Electrons*, Teubner

1926

2. P. A. M. Dirac, *On the Theory of Quantum Mechanics*, Proc. Roy. Soc. **A112**, 661
3. E. Fermi, *Zur Quantelung des idealen einatomigen Gases*, Zeits. f. Physik **36**, 902

1927

4. L. H. Thomas, *The Calculation of Atomic Fields*, Proc. Camb. Phil. Soc. **23**, 542
5. E. A. Milne, *The Total Energy of Binding of a Heavy Atom*, Proc. Camb. Phil. Soc. **23**, 794
6. W. Pauli, Jr., *Über Gasentartung und Paramagnetismus*, Zeits. f. Physik **41**, 81

1928

7. D. R. Hartree, *The Wave Mechanics of an Atom with a Non-Coulomb Central Field*, Proc. Camb. Phil. Soc. **24**, 89, 111
8. R. H. Fowler, *The Photo-electric Threshold Frequency and the Thermionic Work Function*, Proc. Roy. Soc. **A118**, 229
9. R. H. Fowler and L. W. Nordheim, *Electron Emission in Intense Electric Fields*, Proc. Roy. Soc. **A119**, 173
10. L. W. Nordheim, *On the Kinetic Method in the New Statistics and its Application in the Electron Theory of Conductivity*, Proc. Roy. Soc. **A119**, 689
11. J. E. Lennard-Jones and H. J. Woods, *The Distribution of Electrons in a Metal*, Proc. Roy. Soc. **A120**, 727
12. A. T. Waterman, *The Effect of Electric Fields in the Emission of Electrons from Conductors*, Proc. Roy. Soc. **A121**, 28
13. L. W. Nordheim, *The Effect of the Image Force on the Emission and Reflection of Electrons by Metals*, Proc. Roy. Soc. **A121**, 626
14. R. H. Fowler, *The Thermionic Emission Constant A*, Proc. Roy. Soc. **A122**, 36
15. H. Bethe, *Theorie der Beugung von Elektronen an Kristallen*, Ann. d. Physik **87**, 55
16. L. Nordheim, *Zur Theorie der thermischen Emission und der Reflexion von Elektronen an Metallen*, Zeits. f. Physik **46**, 833
17. A. Sommerfeld, W. V. Houston and C. Eckart, *Zur Elektronentheorie der Metalle auf Grund der Fermischen Statistik*, Zeits. f. Physik **47**, 1
18. J. Frenkel, *Zur wellenmechanischen Theorie der metallischen Leitfähigkeit*, Zeits. f. Physik **47**, 819
19. E. Fermi, *Eine statistische Methode zur Bestimmung einiger Eigenschaften des Atoms und ihrer Anwendung auf die Theorie des periodischen Systems der Elementa*, Zeits. f. Physik **48**, 73
20. W. V. Houston, *Elektrische Leitfähigkeit auf Grund der Wellenmechanik*, Zeits. f. Physik **48**, 449
21. J. Frenkel, *Elementare Theorie magnetischer und elektrischer Eigenschaften der Metalle beim absoluten Nullpunkt der Temperature*, Zeits. f. Physik **49**, 31
22. W. Heisenberg, *Zur Theorie des Ferromagnestimus*, Zeits. f. Physik **49**, 619
23. J. Frenkel and N. Mirulobow, *Zur wellenmechanischen Theorie der metallischen Leitfähigkeit*, Zeits. f. Physik **49**, 885
24. J. Frenkel, *Über die elektrische Oberflächenschicht der Metalle*, Zeits. f. Physik **51**, 232
25. F. Bloch, *Über die Quantenmechanik der Elektronen in Kristallgittern*, Zeits. f. Physik **52**, 555

1929

26. B. Podolsky and L. Pauling, *Momentum Distribution in Hydrogen-like Atoms*, Phys. Rev. **34**, 109
27. J. C. Slater, *The Theory of Complex Spectra*, Phys. Rev. **34**, 1293
28. W. V. Houston, *Temperature Dependence of Electrical Conductivity*, Phys. Rev. **34**, 279
29. R. de L. Kronig, *The Quantum Theory of Dispersion in Metallic Conductors*, Proc. Roy. Soc. **A124**, 409
30. T. E. Stern, B. S. Gossling and R. H. Fowler, *Further Studies in the Emission of Electrons from Cold Metals*, Proc. Roy. Soc. **A124**, 699
31. H. Bethe, *Termaufspaltung in Kristallen*, Ann. d. Physik **3**, 133
32. F. Bloch, *Zur Suszeptibilität und Widerstandsänderung der Metalle im Magnetfeld*, Zeits. f. Physik **53**, 216
33. R. Peierls, *Zur Theorie der galvanomagnetischer Effekte*, Zeits. f. Physik **53**, 255
34. F. Bloch, *Bemerkung zur Elektronentheorie des Ferromagnetismus und der elektrischen Leitfähigkeit*, Zeits. f. Physik **57**, 545

1930

35. J. C. Slater, *Note on Hartree's Method*, Phys. Rev. **35**, 210
36. J. C. Slater, *Cohesion in Monovalent Metals*, Phys. Rev. **35**, 509
37. P. M. Morse, *The Quantum Mechanics of Electrons in Crystals*, Phys. Rev. **35**, 1310
38. J. C. Slater, *Atomic Shielding Constants*, Phys. Rev. **36**, 57
39. J. Frenkel, *On the Electrical Resistance of Contacts between Solid Conductors*, Phys. Rev. **36**, 1604
40. P. A. M. Dirac, *Quantum Mechanics*, Oxford University Press.
41. P. A. M. Dirac, *Note on Exchange Phenomena in the Thomas Atom*, Proc. Camb. Phil. Soc. **26**, 376
42. L. Brillouin, *Les électrons libres dans les métaux et la rôle des réflexions de Bragg*, J. de physique **1**, 377
43. F. Bloch, *Über die Wechselwirkung der Metallelektronen*, Leipziger Vorträge, 67
44. R. Peierls, *Zur Theorie der elektrischen und thermischen Leitfähigkeit von Metallen*, Ann. d. Physik **4**, 121
45. R. Peierls, *Zwei Bemerkungen zur Theorie der Leitfähigkeit*, Ann d. Physik **5**, 244
46. F. Bloch, *Zum elektrischen Widerstandsgesetz bei tiefen Temperaturen*, Zeits. f. Physik **59**, 208
47. J. Frenkel, *Über die Formel für das mittlere Gitterpotential*, Zeits. f. Physik **59**, 649
48. V. Fock, *Näherungsmethode zur Lösung des quantenmechanischen Mehrkörper problems*, Zeits. f. Physik **61**, 126
49. F. Bloch, *Zur Theorie des Ferromagnetismus*, Zeits. f. Physik **61**, 209
50. V. Fock, *"Self-consistent field" mit Austausch für Natrium*, Zeits. f. Physik **62**, 795
51. N. H. Frank, *Zur Theorie der galvanomagnetischen, thermomagnetischen, und thermoelektrischen Erscheinungen in Metallen*, Zeits. f. Physik **63**, 596
52. L. Landau, *Diamagnetismus der Metalle*, Zeits. f. Physik **64**, 629
53. N. H. Frank, *Über die metallische Widerstandsänderung in starken Magnetfeldern*, Zeits. f. Physik **64**, 650

1931

54. J. Frenkel, *On the Transformation of Light into Heat in Solids*, Phys. Rev. **37**, 17
55. W. V. Houston, *The Structure of Soft X-Ray Lines*, Phys. Rev. **38**, 1797
56. A. Sommerfeld and N. H. Frank, *Statistical Theory of Thermoelectric, Galvano- and Thermomagnetic Phenomena in Metals*, Rev. Mod. Phys. **3**, 1
57. R. de L. Kronig and W. G. Penney, *Quantum Mechanics of Electrons in Crystal Lattices*, Proc. Roy. Soc. **A130**, 499
58. R. de L. Kronig, *The Quantum Theory of Dispersion in Metallic Conductors, II*, Proc. Roy. Soc. **A133**, 255
59. A. H. Wilson, *The Theory of Electronic Semi-Conductors*, Proc. Roy. Soc. **A133**, 458
60. A. H. Wilson, *The Theory of Electronic Semi-Conductors, II*, Proc. Roy. Soc. **A134**, 277
61. L. Brillouin, *Quantenstatistik*, Springer, Berlin
62. L. Nordheim, *Zur Elektronentheorie der Metalle, I*, Ann. d. Physik **9**, 607; II, **9**, 641.
63. R. Peierls, *Zur Theorie der magnetischen Widerstandsänderung*, Ann. d. Physik **10**, 97
64. E. Teller, *Der Diamagnetismus von freien Elektronen*, Zeits. f. Physik **67**, 311
65. R. de L. Kronig, *Zur Theorie der Feinstruktur in den Röntgenabsorptionsspektren*, Zeits. f. Physik **70**, 317
66. H. Bethe, *Zur Theorie der Metalle. I. Eigenwerte und Eigenfunktionen der linearen Atomkette*, Zeits. f Physik **71**, 205

1932

67. P. S. Epstein, *On Ferromagnetism and Related Problems of the Theory of Electrons*, Phys. Rev. **41**, 91
68. J. Frenkel, *Wave Mechanics, Elementary Theory*, Oxford
69. A. H. Wilson, *The Theory of Metals I*, Proc. Roy. Soc. **A138**, 594
70. R. de L. Kronig and H. J. Groenewold, *Lorentz-Lorenz Correction in Metallic Conductors*, K. Akad. Amsterdam **35**, 974
71. L. Brillouin, *Les problèmes de perturbations et les champs self-consistents*, J. de physique **3**, 373
72. L. Brillouin, *Les électrons libres dans un réseau cristallin. Equation ondulatoire et propriétés magnetiques*, J. de physique **3**, 565
73. R. Peierls, *Elektronentheorie der Metalle, Ergebnisse der Naturwissenschaften*, Springer, Berlin
74. R. Peierls, *Zur Frage des elektrischen Widerstandsgesetzes für tiefe Temperaturen*, Ann. d. Physik **12**, 154
75. E. Kretschmann, *Beitrag zur Theorie des elektrischen Widerstandes und der Supraleitfähigkeit der Metalle*, Ann. d. Physik **13**, 564
76. F. Hund, *Zur Theorie der schwerflüchtigen nichtleitenden Atomgittern*, Zeits. f. Physik **74**, 1
77. F. Bloch, *Zur Theorie des Austauschproblems und der Remanenzerscheinungen der Ferromagnetika*, Zeits. f. Physik **74**, 295
78. R. Schachenmaier, *Wellenmechanische Vorstudie zu einer Theorie der Supraleitung*, Zeits. f. Physik **74**, 503
79. R. de L. Kronig, *Zur Theorie der Feinstruktur in den Röntgenabsorptionsspektren II*, Zeits. f. Physik **75**, 191; III, **75**, 468
80. V. Fock, *Konfigurationsraum und zweite Quantelung*, Zeits. f. Physik **75**, 622
81. Y. Fujioka, *Zur Dispersionstheorie im Metallischen Leiter*, Zeits. f. Physik **76**, 537
82. W. Kroll, *Ein Beitrag zur Theorie der thermoelektrischen Effekte*, Zeits. f. Physik **77**, 322
83. W. Lenz, *Über die Anwendbarkeit der statistischen Methode auf Ionengitter*, Zeits. f. Physik **77**, 713
84. H. Jensen, *Die Ladungsverteilung in Ionen und die Gitterkonstante des Rubidiumbromids nach der statistischen Methode*, Zeits. f. Physik **77**, 722
85. W. Kroll, *Zur Theorie der thermoelektrischen Effekten bei tiefen Temperaturen*, Zeits. f. Physik **80**, 50

86. V. Fock, *Über die Gültigkeit des Virialsatzes in der Fermi-Thomasschen Theorie*, Phys. Zeits. d. Sow. **1**, 747

1933

87. E. Wigner and F. Seitz, *On the Constitution of Metallic Sodium*, Phys. Rev. **43**, 804
88. J. Frenkel, *On a Possible Explanation of Superconductivity*, Phys. Rev. **43**, 907
89. F. Zwicky, *Superconductivity*, Proc. Nat. Acad. Sci. **19**, 818
90. H. S. Taylor, H. Eyring and A. Sherman, *Binding Energies in the Growth of Crystal Nuclei from Metallic Atoms*, J. Chem. Phys. **1**, 68
91. J. C. Slater, *The Virial and Molecular Structure*, J. Chem. Phys. **1**, 687
92. R. H. Fowler, *Electronic Properties of Conductors and Insulators*, Proc. Roy. Soc. **A141**, 56
93. L. Brillouin, *La méthode du champ self-consistent*, Actualités scientifiques, Hermann, Paris, No. 71
94. L. Brillouin, *Champs self-consistents et électrons métalliques III*, J. de physique **4**, 1
95. L. Brillouin, *Le champ self-consistent, pour des électrons liés; la supraconductibilité*, J. de physique **4**, 333
96. F. Bloch, *Remarque sur un théorème de conservation dans le théorie des métaux*, J. de physique **4**, 486
97. L. Brillouin, *Sur la stabilité du courant dans un supraconducteur*, J. de physique **4**, 677
98. A. Sommerfeld and H. Bethe, *Elektronentheorie der Metalle*, Handb. d. Physik **24**, 2nd edition
99. R. de L. Kronig, *Zur Theorie der Supraleitfähigkeit*, Zeits. f. Physik **78**, 744; II, **80**, 203
100. R. Peierls, *Zur Theorie des Diamagnetismus von Leitungselektronen*, Zeits. f. Physik **80**, 763
101. H. Frölich, *Über die Absorption der Metallen im sichtbaren und ultravioletten*, Zeits. f. Physik **81**, 297
102. H. Jensen, *Über die Gültigkeit des Virialsatzes in der Thomas-Fermischen Theorie*, Zeits. f. Physik **81**, 611
103. R. Peierls, *Zur Theorie der Metalle, Erwiderung auf eine Arbeit von A. H. Wilson*, Zeits. f. Physik **81**, 697
104. H. Hellmann, *Zur Rolle der kinetischen Elektronenenergie für die zwischenatomaren Kräften*, Zeits. f. Physik **85**, 180
105. H. Bethe and H. Fröhlich, *Magnetische Wechselwirkung der Metallelektronen. Zur Kritik der Theorie der Supraleitung von Frenkel*, Zeits. f. Physik **85**, 389
106. W. Kroll, *Zur Theorie der Druckabhängigkeit der elektrischen Leitfähigkeit der Metalle*, Zeits. f. Physik **85**, 398
107. L. Landau, *Theorie der Supraleitfähigkeit*, Phys. Zeits. d. Sow. **4**, 43
108. Y. Fujioka, *Dispersion Theory in Metallic Conductors*, Inst. Phys. and Chem. Res., Tokyo, Sci. Papers **459** 202

1934

109. H. Jones, N. F. Mott and H. W. B. Skinner, *A Theory of the Form of the X-Ray Emission Bands of Metals*, Phys. Rev. **45**, 379
110. J. C. Slater, *Electronic Energy Bands in Metals*, Phys. Rev. **45**, 794
111. N. F. Mott and C. Zener, *The Optical Properties of Metals*, Proc. Camb. Phil. Soc. **30**, 249
112. H. Jones and C. Zener, *A General Proof of Certain Fundamental Equations in the Theory of Metallic Conduction*, Proc. Roy. Soc. **A144**, 101
113. H. Jones, *The Theory of Alloys in the γ-Phase*, Proc. Roy. Soc. **A144**, 225
114. S. Schubin and S. Wonsowsky, *On the Electron Theory of Metals*, Proc. Roy. Soc. **A145**, 159
115. H. Jones and C. Zener, *The Theory of the Change in Resistance in Magnetic Fields*, Proc. Roy. Soc. **A145**, 268
116. L. Brillouin, *Les électrons dans les métaux du point de vue ondulatoire*, Actualités scientifiques, Hermann, Paris, No. 88
117. L. Brillouin, *Conductibilité électrique et thermique des métaux*, Actualités scientifiques, No. 89
118. L. Brillouin, *Le modèle d'atomes de Fock-Dirac et l'existence des potentials d'ionisation*, J. de physique **5**, 185

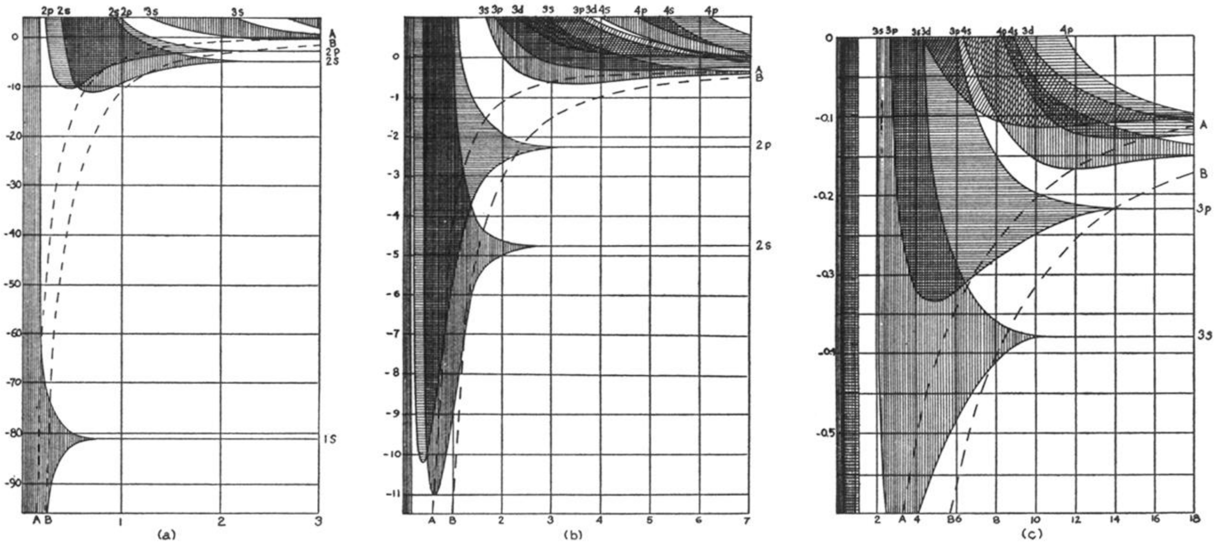

FIG. 9. Energy plotted against internuclear distance for sodium, both in atomic units. Different scales used in (*a*), (*b*), (*c*). The dotted curve *A* represents the maximum potential energy, at the edge of the cell, and *B* represents the mean potential energy throughout the cell, as a function of distance. Most of the curves are fairly accurate, but some are only estimated, so that too much confidence should not be placed in the details of the graphs, though the general relations are correct. Some states are omitted in the upper right-hand corners of all graphs, where the overlapping is really much more than is indicated. In the actual crystal, the half distance of separation is between 3.5 and 4 units.