Energy Bands in Solids

After the Second World War great commercial interest in transistors and the availability of digital computers sparked a rise in experimental and theoretical studies of solids. Our knowledge of energy bands, Fermi surfaces and magnetic properties of crystals have all profited.

by John C. Slater

THE SECOND WORLD WAR interrupted work in the quantum theory almost completely, and yet it furnished a stimulus that had the most profound effect on postwar work. First, there was the effect of the microwave radar research that was carried on at the Massachusetts Institute of Technology radiation laboratory, Bell Telephone Laboratories and other institutions. Physicists who had been converted into radar engineers during the war went back to their laboratories with a greatly enhanced knowledge of electronics and microwave techniques. This situation resulted in the development of new and powerful experimental techniques of high-frequency experimentation on solids, leading to the methods of paramagnetic resonance, radiofrequency spectroscopy, cyclotron resonance, and a myriad of other ways of investigating atoms, molecules and solid state. Another outgrowth of the radar work was the increased interest in solid-state electronics. Radar techniques of the wartime period used silicon crystals as rectifiers much as early radio had used galena crystals. Toward the end of the war remarkable properties of germanium, similar chemically to silicon, were beginning to be appreciated, and an intensive research program to understand these substances was started at Purdue University under Karl Lark-Horovitz. It was realized that these substances were semiconductors, and those working on them began to recall what had almost been forgotten during the war; namely, the way in which the energy-band theory of solids explained semiconductor properties. Bell Telephone Laboratories had the foresight to realize that the electrical properties of semiconductors might have great practical value in more everyday applications than radar and put a team of their best men, including William Shockley, John Bardeen and others, to work.

Practical investigation

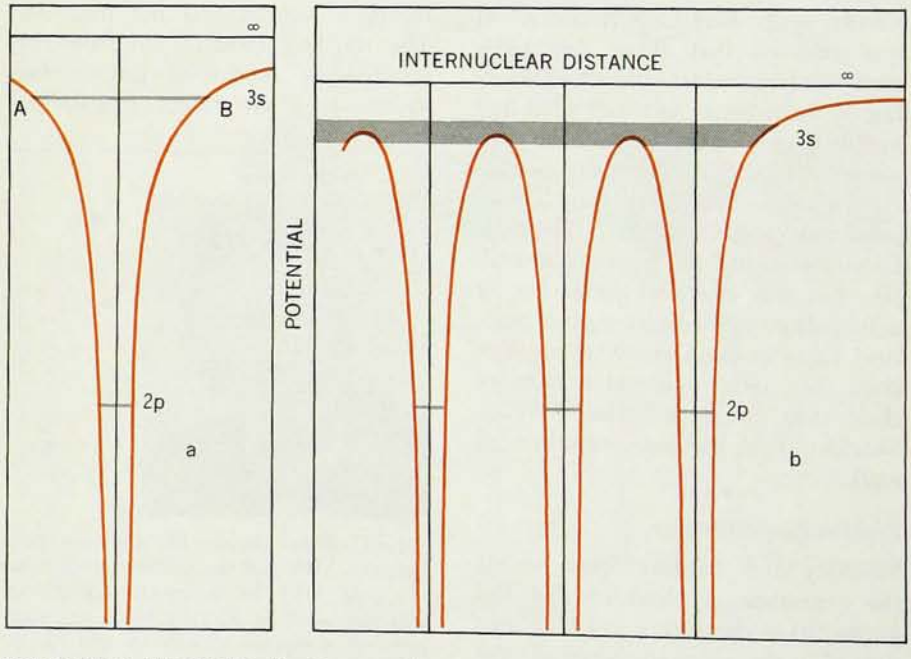
Shockley and Bardeen were among the generation of physicists that had grown up in the 1930's and had been trained in the quantum theory of solid state. Shockley was at MIT and Bardeen was at Princeton and Harvard. They saw the possibility of bridging the gap between the rather empirical state of the art with respect to semiconductors and the sophisticated theory that was already in existence and that they thoroughly understood. The result, of course, was the transistor. The startling research program that resulted in solid-state rather than vacuum-tube electronics formed the

John C. Slater divides his academic year between MIT and the University of Florida. At MIT he is Institute professor emeritus, and at Florida he is graduate research professor of physics and chemistry. In this sequel to his article in the January issue of PHYSICS TODAY, Slater continues the historical devleopment of the quantum theory of matter.

main line of practical development in electrical engineering.

The great commercial importance of these methods meant that the study of electronic properties of crystals changed from being an academic question to a practical one of the greatest importance. Many younger physicists of the postwar generation went into this field. Along with engineering and practical work there was more and more adaptation of the high-frequency research, which had come out of radar work, to investigation of properties of solids. The 1950's represented a period when these methods really reached a peak of development that proved how powerful they were and that drew great attention to the field. One secondary wartime technique helped considerably in this experimental advance: the development at MIT of the Collins liquefier for liquid helium. The most interesting solid-state resonance phenomena can only be observed at very low temperature, where the electrical resistance is low. If it had not been for the ease of reaching these low temperatures, made possible by the new liquefier, progress would have been much slower. Another wartime development that played an essential role in the rise of experimental solid-state physics was the advance in methods of producing pure crystals or crystals with well regulated amounts of impurities. These methods had been essential during wartime for production of rectifier crystals with controlled properties.

Still another wartime experimental development proved useful in a somewhat limited range of problems: neutron diffraction. This advance was an outgrowth of the development of the nuclear reactor. Neutrons are diffracted by crystals much as x rays are, but with one striking difference: Because neutrons have a magnetic moment, their diffraction depends on the magnetic properties of the crystals through which they pass. Ever since the work of Max von Laue and the Braggs in 1913, we knew that x-ray diffraction by crystals furnished a technique by which the distribution of electronic charge density within the atoms could be found experimentally, and Douglas Hartree was concerned, in his early work, with tying in these experimental results with the calculated wave functions of electrons in atoms. In a similar way, neutron diffraction can give information about the density of magnetization as a function of position in a magnetic atom. In the hands of Clifford Shull and others, this technique has supplied detailed information about problems of ferromagnetism, antiferromagnetism and other types of magnetic behavior.



POTENTIAL ENERGY affecting an electron in a self-consistent model of a sodium atom (a). For sodium crystal (b): There are maxima of potential between atoms. Shaded area indicates broadened 3s level. Potentials go from minus infinity at the nucleus to zero at infinity and are adjusted so that x-ray levels agree. —FIG. 1

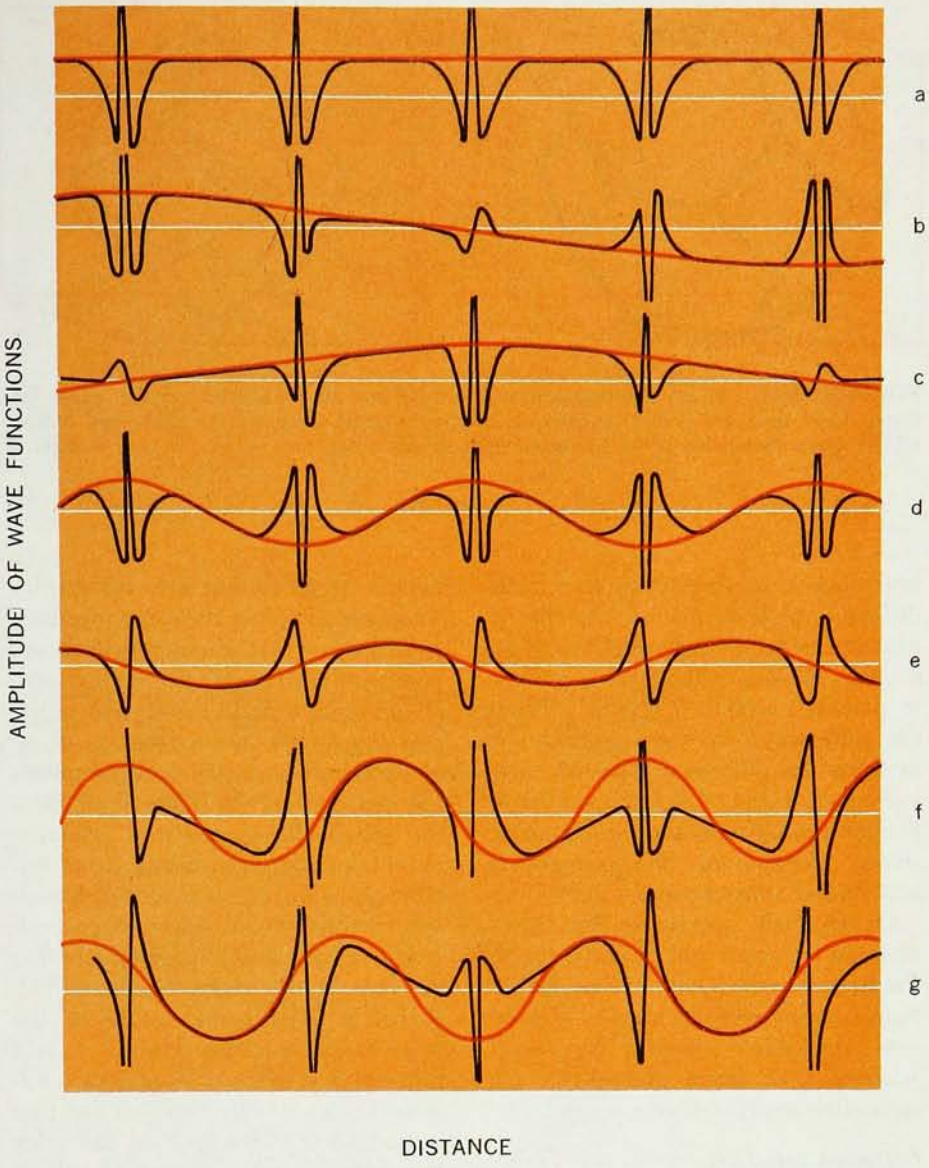
Computer techniques

All these methods, and others too, led during the period starting in the 1950's to an enormous increase in our experimental knowledge of the behavior of solids. Concurrently there was an improvement of technique that enabled one to make great advances in the application of quantum theory to matter. This advance was the development of the electronic digital computer, an outcome of wartime work that in a practical way became striking during the 1950's, and is still going on at a rapid pace. We realized that wave mechanics, invented in 1926, is the most difficult branch of mathematical physics, and this mathematical difficulty more than anything else had stood in the way of its advance in the prewar period. Theorists were devoted to finding approximations that might give some idea of the answers to problems rather than attacking them straightforwardly. Even in the prewar days one knew that mechanical methods of calculations would be extremely useful. Hartree had used the differential analyzer built by Vannevar Bush at MIT and had built his own version of the same machine in England. But this analog computer did not have the accuracy necessary for the required calculations; for most purposes it was limited to an accuracy of about one part in a thousand. The digital computers that were developed after the war had almost unlimited accuracy, and they opened a wholly new era in the application of numerical methods to quantum theory.

I remember well the first steps that students in my own research group at MIT went through in their spontaneous desire to adapt the developing computers to quantum-theoretical calculations. One of them, George Pratt, discovered that the registrar's office had a simple form of IBM punched-card machine that was used for tabulating students' grades. He had a feeling that he could adapt this machine to solve Schrödinger's equation. He received permission to try, and soon he was solving equations of the self-consistent field for atoms. Another student, Alvin Meckler, found out that there was an experimental computer under development by the electrical engineers, the Whirlwind. It tested new forms of memory units that

were being worked out at MIT. But it had remarkable capabilities and, for the time (the early 1950's), great speed. Meckler and some of the other students obtained permission to work out of hours on the machine. Thus began a number of years of extremely profitable use of this Whirlwind computer, a period in which methods of handling the quantum theory of atoms, molecules and solids were worked out. No doubt similar activities went on in other institutions. Late in the 1950's newer commercial computers—the IBM and other types—were available. By then my own research group, and a few others at various institutions through the country, had reached the point where we could make really significant wave-mechanical calculations. Only during the last ten years have computers reached the point where large-scale quantum-mechanical calculations have been practical.

Naturally those of us who had been interested in energy-band theory before the war were anxious to resume calculations as rapidly as possible. I tried from the early 1950's to encourage my students to program the method that we now know as the augmented plane-wave method. This method had been suggested in 1937 and tested by only one calculation before the war. I was convinced that it would be a valuable method for investigating energy bands, but it was beyond the capabilities of the first machine we had available, and it took most of the 1950's for Melvin Saffren and John Wood to program it on the Whirlwind. When the first large IBM computer became available, the IBM 704, we really started on energy-band calculations: This was almost 1960.

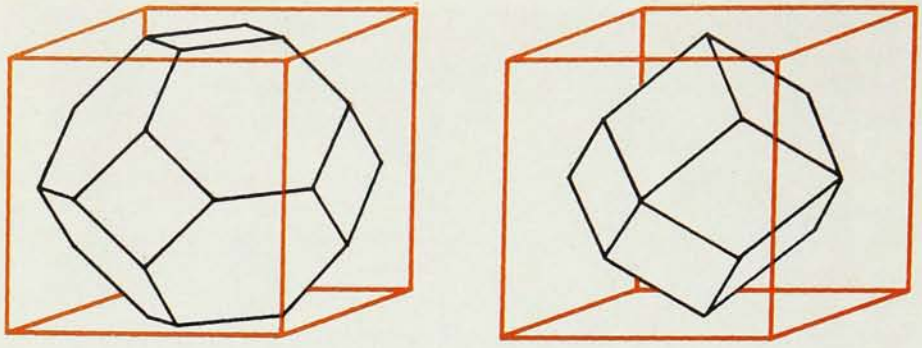


WAVE FUNCTIONS of the 3s band as a function of position going through the crystal from one atom to its nearest neighbor. The various curves show values of the wave vector k, from k = 0, in (a), to the edge of the Brillouin zone for (g). —FIG. 2

Plane-wave methods

The first real computer success with postwar energy-band calculations had come earlier with the orthogonalized plane-wave method that Conyers Herring had proposed and tried in 1940. Frank Herman of Columbia, working closely with Herring at the Bell Laboratories, programmed this method by 1952. His work, partly in collaboration with Joseph Callaway, on diamond and germanium first showed the real power of energy-band calculations. So much had been learned experimentally about germanium that it was a particularly good substance to study. I shall refer later to some of the recent and very striking comparisons between experiment and theory for this material.

These two methods, the augmented plane-wave and orthogonalized plane-wave (APW and OPW) methods, were not the only ones used. J. Korringa in 1947 and Walter Kohn and Norman Rostocker in 1954, had suggested two closely related methods (now called collectively the "KKR method"), and Benjamin Segall and Frank Ham of the General Electric Company had programmed this method during the latter part of the 1950's. Segall at GE, and Glenn Burdick at MIT, were both working on the energy bands of copper and made the first striking intercomparison of results for different methods in 1961. This energy-band problem had first been handled by the prewar Wigner–Seitz or cellular method by Harry Krutter in 1935, and became the first test of the APW method by Marvin Chodorow in 1939. We arranged for Segall and Burdick to intercompare results with each other and with Chodorow's work by using their methods on Chodorow's potential. We were all greatly satisfied when the quite complicated energy bands of Segall and Burdick, worked out respectively by the KKR and APW methods, agreed in detail and coincided with the few points that Chodorow had calculated so much earlier. From further intercompari-



POLYHEDRAL CELLS for body-centered (left) and face-centered (right) cubic lattices. Each unit cell (color) contains a polyhedron that surrounds an atom and defines all the space that is closer to that atom than to any other. —FIG. 3

sons there is no doubt now that these different methods are all capable of solving the Schrödinger problem of an electron moving in the type of periodic potential met in a crystal. There are differences in the suitability of methods for different types of substances, however, that have tended to confine the OPW method to diamond, silicon, germanium, 3–5 compounds and similar substances; the APW and KKR methods appear to be better adapted to metals and in particular to the transition elements. These limitations are not very stringent, and by now the APW method has been adapted with much success for all sorts of elements and compounds.

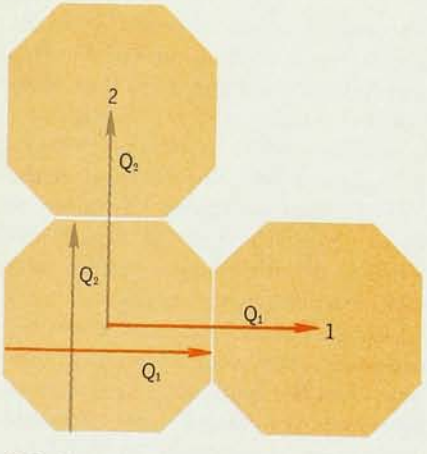
Different solutions

It is time now to discuss the differences among these methods and the way they solve Schrödinger's equation before we go further into the results. They are all facing the same problem: to find some sort of mathematical approximation that will give a good and, at the same time, a convenient representation of the wave function of an electron moving in the periodic potential in a crystal. In figure 1 we show the spherically symmetrical potential in which an electron moves in an isolated atom and the way in which such spherical potentials, spaced periodically in a crystal, lead to the problem that we face in the energy-band method. Figure 2 depicts the nature of the wave function in the crystal. It is a function that behaves around each nucleus like an atomic function but that is modulated as we go from one atom to another like a plane wave, whose complex formulation is the exponential function $\exp(i\mathbf{k}\cdot\mathbf{R})$.

Two methods of describing the wave function of an electron in a crystal were proposed in 1928, very shortly after the development of wave mechanics. Felix Bloch suggested one called the "tight-binding method." He proposed that one superpose atomic wave functions, multiplied by suitable factors $\exp(i\mathbf{k}\cdot\mathbf{R})$, where $\mathbf{R}$ represents the vector position of the given atom. This superposition gives a function of the proper sort. The only serious drawback is that exact calculations are difficult to carry out. They involve taking products of atomic wave functions on two atomic sites, multiplied by a spherically symmetrical potential on a third site, and integrating these products. The resulting integrals, called "three-center integrals," are of a type also found in the theory of polyatomic molecules. The difficulty of computing them, even with the most modern computational methods, has been so great that there have been very few really accurate calculations made by this tight-binding method. However, present improvements in the technique of calculations of such integrals may well bring this method back into a position where it can be very important.

The second method, also used by Bloch, Hans Bethe, Leon Brillouin and others, is the expansion of the wave function in plane waves, a form of three-dimensional Fourier expansion. In principle it is possible, but in practice it is so slowly convergent and requires so many plane waves that it is not practicable. Figure 2 demonstrates that the wave function resembles a plane wave in the region between nuclei, and a few plane waves suffice to describe it very accurately in this part of space. To expand the rapid oscillations of the function near the nuclei, however, requires millions of plane waves, and direct calculations with them are not possible. We shall point out later that a modification of this method—sometimes called the "almost-free-electron method" since it would work well for almost-free electrons in a periodic potential of small amplitude—has had a good deal of application.

The cellular method of Wigner and Seitz, suggested in 1933, was based on the fact that the potential surrounding a given nucleus in the crystal is very nearly spherically symmetrical. Wigner and Seitz surrounded an atom by a polyhedron in a crystal such as sodium, shown in figure 3, containing all the space that is closer to that atom than to any other. Such cells fill space, and if we can solve Schrödinger's equation within one such cell, we have solved it everywhere. The reason is that the wave function will be multiplied by a constant, $\exp(i\mathbf{k}\cdot\mathbf{Q})$, where $\mathbf{Q}$ is the vector from the center of one cell to the center of another and goes from one such cell to a neighboring cell.



CROSS SECTION through type of cell met with in sodium crystal. Vectors Q_1 and Q_2 represent displacements either from the center of one cell to another or from one face to another. —FIG. 4

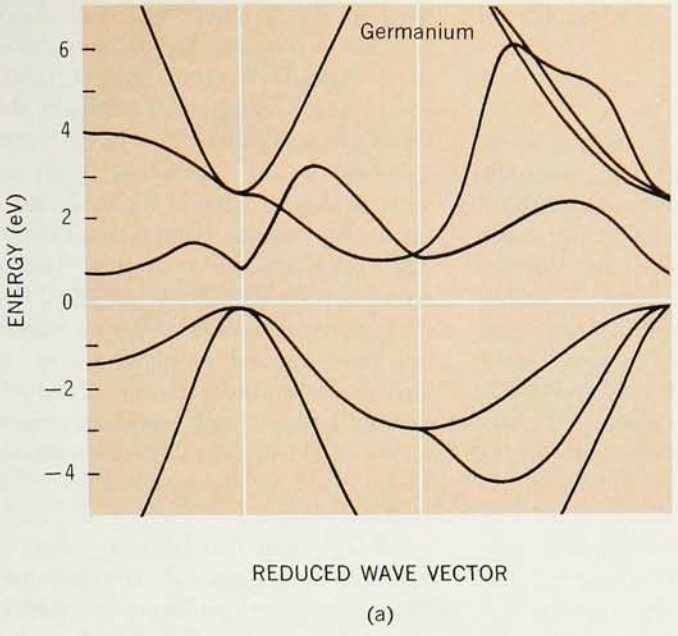
Cellular method

Wigner and Seitz suggested that Schrödinger's equation be solved within the cell as if it were really spherically symmetrical. Our knowledge of wave mechanics shows that such a solution can be written as a linear combination of products of spherical harmonics of the angles, whose properties

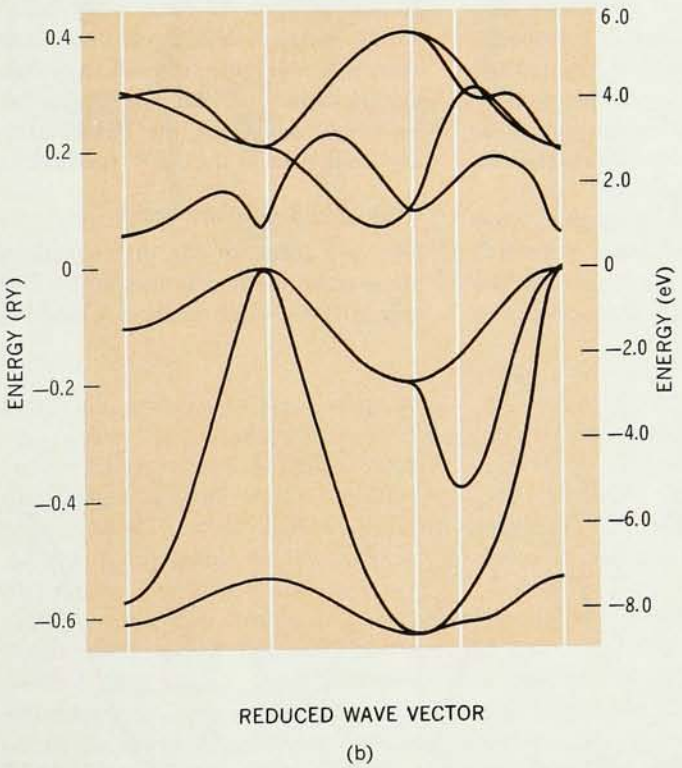
are well known, multiplied by radial functions that are solutions of the radial wave equation. This radial wave equation can be handled very easily by numerical integration of the differential equation, an easy problem for even the simplest digital computers. The only difficulty is to satisfy the proper boundary conditions around the boundary of the polyhedral cell. In going from one face of the cell to a corresponding point on the opposite face, the wave function must be multiplied by a suitable factor $\exp(i\mathbf{k}\cdot\mathbf{Q})$, where $\mathbf{Q}$ is the vector either from the center of one cell to that of another or from one face of the cell to the opposite face. In figure 4, which shows a section through the type of cells actually met with in sodium, $\mathbf{Q}_1$ and $\mathbf{Q}_2$ represent vectors of the type we are speaking of, and it is clear from the diagram how they represent either displacements from the center of one cell to another or from one face to the other.

The only trouble with this method is that it is almost impossible to apply this type of boundary condition rigorously. A very simplified version of it was used in the 1930's: One demanded that the condition be satisfied only at the centers of the various faces of the polyhedral cell. This could be done with only a few disposable constants and was achieved by using solutions of the spherically symmetrical problem corresponding to only a few values of quantum numbers l and m that appear in such solutions in wave mechanics. The result of this approximation was that the energy bands found by this method, though fairly good qualitatively, were not quantitatively reliable. Some work with this method since the war has satisfied boundary conditions at many more points on the surface of the polyhedron but becomes rather cumbersome to carry out.

To avoid this difficulty, I suggested the APW method in 1937. In this method, a sphere was inscribed in each of the polyhedra of figure 4 and surrounds each nucleus. Within each sphere the wave function was expanded in the solution used by Wigner and Seitz, but outside the spheres it was expanded in plane waves. We have mentioned that the plane-wave expansion works well for the outer part of the wave function, where it was used, and the boundary condition is transformed in this method to a problem carried out on the surface of the spheres. Here we must make the wave function and its first derivative continuous. This is a simple type of boundary condition to handle rigorously. It is easy to expand the plane waves in the form of spherical harmonics of angle times radial functions that prove to be spherical Bessel functions. A very manageable set of equations leads to a smooth matching procedure for each spherical harmonic, or each value of the azimuthal quantum number l. This APW method has, in the last few years, shown itself capa-



ENERGY-BAND MODELS for germanium. Energy is a function of the wave vector k, and all the information about energy bands is contained in the central unit cell in k space—the Brillouin zone. Theoretically determined curves (a) and experimentally determined curves (b) agree closely. —FIG. 5



ble of giving useful solutions for all sorts of crystals.

Muffin-tin potential

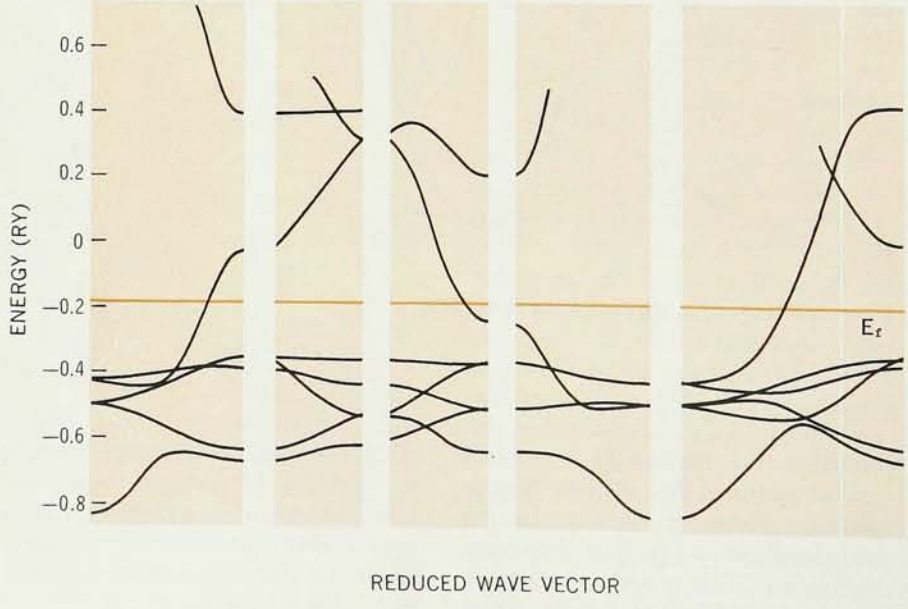
One point should be mentioned regarding this method. As originally proposed, the method was formulated for a type of potential that has come colloquially to be called the "muffin-tin," spherically symmetrical within the spheres but constant outside them. For some crystals this is a pretty good approximation. Most of the calculations that have been made with this potential have not suffered seriously in using it. But for some crystals an angular dependence of the potential surrounding an atom—the so-called crystal field—is an essential part of the problem, and this dependence demands modification of the muffin-tin potential. Fortunately the APW method is not tied in any way to the muffin-tin potential, and a nonconstant potential outside the spheres can be easily incorporated into it. But this has not been done in many cases so far and has led to the impression among some people that it is impossible. But for the tight-binding method it is natural to build up the potential as a superposition of spherical potentials, and a nonconstant potential outside the spheres comes naturally.

The KKR method, suggested after the war, uses a muffin-tin potential. The same sort of expansion of the wave function inside the spheres is used in the Wigner–Seitz and APW methods. Between the spheres, however, instead of expanding in plane waves the KKR method expands the wave function in outgoing or scattered spherical waves originating from the various atoms. Thus it is closely analogous to atomic scattering theory. The final result is mathematically equivalent to the final result of the APW method, for the scattered waves can be expanded in plane waves or vice versa, and the choice of which method to use is one based on convenience and rapidity of convergence. There is still no consensus as to which method is better, which is a pretty good indication that they are roughly equivalent in general convenience. There are some problems for which the KKR method is definitely superior such as that of an impurity atom in an otherwise perfect crystal. Here the KKR method leads directly to the wave scattered by the impurity atom, which is ordinarily what we are mainly interested in. I feel that the KKR method is likely to have many future applications. I should point out, however, that it is more difficult to include modifications of the muffin-tin assumption regarding the potential in the KKR than in the APW method.

Orthogonalized plane wave

Next we come to the orthogonalized plane-wave method suggested by Herring in 1940. This method is based on the following observation regarding wave functions such as those in figure 2. These wave functions are for sodium, and the ones shown are for the energy band arising from the 3s atomic wave functions. Their general characteristic is that, for small values of the distance from any nucleus, they resemble strongly an atomic 3s wave function (or, better, a linear combination of a 3s and a 3p wave function), while further out they resemble a linear combination of a few plane waves. But the inner oscillations of the 3s and 3p wave functions are very much like those found in the 1s, 2s and 2p atomic functions. Hence Herring reasoned that if we superposed atomic 1s, 2s and 2p functions, combined with factors $\exp(i\mathbf{k}\cdot\mathbf{R})$ to form functions with the proper behavior as we go from cell to cell (as in the tight-binding method), and add in a few plane waves, the linear combination should give a good representation of the true function. Calculations with such functions should be easy to carry out. Though we are using tight-binding types of sums, we do not experience the mathematical difficulties of three-center integrals that we mentioned earlier in connection with the tight-binding method. The reason is that the only atomic orbitals used in the combinations are those like 1s, 2s and 2p, which fall off so rapidly as we go away from the nucleus that they are quite negligible by the time they overlap similar orbitals from the next atom; the resulting three-center integrals are all negligibly small. The only integrals that must be computed are those between such a sum of atomic orbitals and a plane wave. Fortunately these are easy to handle.

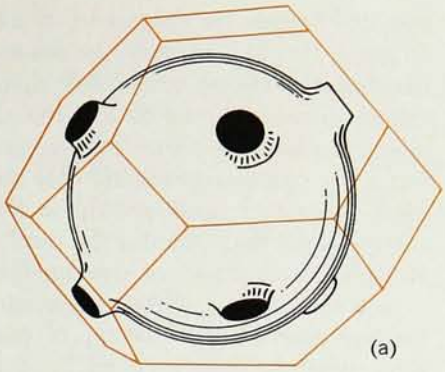
The reason why this method is called the orthogonalized plane-wave method is as follows. Any two solutions of Schrödinger's equation must be orthogonal to each other, in the quantum-mechanical sense. The linear combinations of 1s, 2s and 2p wave functions of the tight-binding type, with coefficients of the type $\exp(i\mathbf{k}\cdot\mathbf{R})$, generally called "Bloch sums," form very good approximations to solutions of the periodic-potential problem arising from corresponding atomic wave functions. Consequently the wave functions formed from the atomic 3s wave function must be orthogonal to these Bloch sums of 1s, 2s and 2p. As a step in achieving this

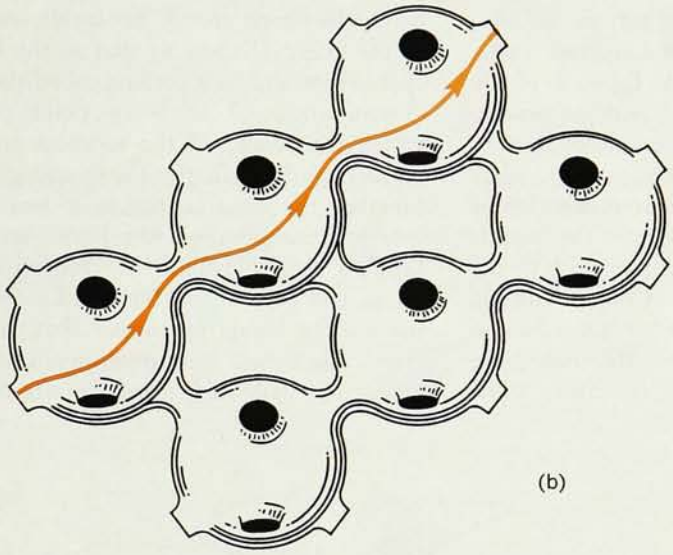


JOINING OF METHODS. Curves represent energy-band calculations for copper by both the APW and KKR methods. Fermi energy is designated by E_f. —FIG. 6

orthogonality, Herring took each plane wave and added to it linear combinations of the Bloch sums of 1s, 2s and 2p such that the resulting function was orthogonal to those Bloch sums. He called such a linear combination an orthogonalized plane wave. Then when one made linear combinations of the orthogonalized plane waves, the resulting function would automatically have the property of orthogonalization to the wave functions of the inner shells. The mathematical formulation of the OPW method could be thrown into the form of finding a linear combination of orthogonalized plane waves that best represents the 3s wave function. This is a problem easily stated in the language of matrix mechanics and easily solved by a computer.

This OPW method is very good for a crystal like sodium, but it does not work well for one like copper or the 3d transition elements that have energy bands arising from the 3d electrons. The reason is that the atomic 3d orbitals do not have entirely the characteristics of either inner orbitals, like the 1s, 2s and 2p of sodium, or outer ones, like the 3s of sodium. Their maximum density comes well inside the atom, but the tails of the wave functions extend out to the boundary of the cell. If we try, by the OPW method, to use Bloch sums of 3d orbitals as of the inner orbitals, we run into three-center integrals that cannot be disregarded, while if we expand them in plane waves, we need a great many to get a convergent expansion. There are, however, many problems for which the OPW method works well; those that concern the semiconductors such as silicon and germanium are conspicuous. The reason why it works, for example with germanium, is that this element is far enough beyond copper in the periodic system (copper is followed by zinc and gallium before we arrive at germanium) so that the 3d orbitals have shrunk inside the atom to a point where they are really inside orbitals; the Bloch sum of them in the OPW method does not involve any appreciable three-center integrals. There are many elements and compounds that have this good feature.





FERMI SURFACE for copper in single Brillouin zone is spherical but with necks extending out to the surfaces of the zone (a). Open orbit for an electron in copper (b) along Fermi surface passes from one cell to the next by means of the necks. —FIG. 7

Approximations

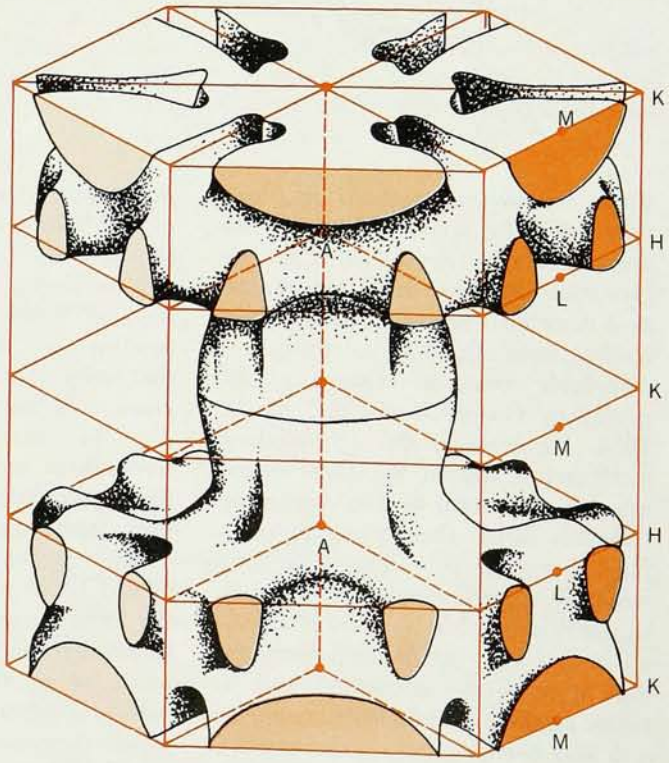
In addition to these methods that are capable of accurate energy-band calculations directly from first principles, there are several methods that are very useful for approximations to the energy bands. They fit either the information acquired from experiment or a more elaborate calculation by an analytic form containing relatively few disposable constants. One was suggested by George Koster and me in 1954. It involved the tight-binding method but treated the three-center integrals that occur in it as disposable constants rather than trying to compute them directly. This is a very workable approximation method. Let us use it to illustrate the sort of agreement that can be obtained between this approximate calculation and a more accurate theoretical treatment and also between theory and experiment. In quite recent work, Herman and his colleagues have obtained very accurate energy bands for germanium by the OPW method. This calculation makes a very slight use of experimental information to determine some points of the potential that are difficult to derive directly from theory. Herman estimates the ratio of fundamental theoretical information that goes into it as 95% and the amount of empirical information as 5%. In contrast, Gene and Mildred Dresselhaus executed a treatment by the empirical tight-binding method entirely from experimental information.

The tight-binding calculation involved 13 disposable parameters, and the Dresselhauses have found 13 different experiments relating to germanium, all of which have been done with good experimental accuracy, from which these constants can be determined. These experiments involve effective masses, optical absorption, cyclotron resonance and many other types of experiment. In figure 5a we

show the energy bands determined by Herman and his colleagues, and in figure 5b those found by the Dresselhauses. Such agreements between theory and experiment give a great deal of confidence in the fundamental correctness of the energy-band method.

I should discuss briefly what is being plotted in figures 5a and 5b and in energy-band work in general. Quantity **k** in the expression $\exp(i\mathbf{k}\cdot\mathbf{R})$ is the wave vector, and the energy is a function of the components of this vector. The energy is a periodic function of position in **k** space, and all the information about the energy bands is contained in the central unit cell of this space, called the "central Brillouin zone." To give a graphical description of the three-dimensional variation of energy with **k**, for each of the energy bands, the expedient usually adopted is to go along various lines in **k** space (for instance, along radius vectors from the origin out in different directions or along lines on the boundary of the Brillouin zone) and give the energy as a function of the magnitude of **k** along these lines. Figures 5a and 5b plot energy against the magnitude of **k** in different directions. The essential feature for the reader to examine is first the accurate agreement between the almost completely theoretical curves of figure 5a and the almost completely experimental ones of figure 5b. Second, he should note the great amount of detail present in the curves, detail that includes the results of many different sorts of experiments.

In a semiconductor like germanium the levels are filled to the top of one band called the "valence band;" above that is a gap, and above the gap is an empty band, the conduction band. This gap is indicated in figures 5a and 5b. The Fermi energy, the top of the occupied levels, lies within the gap. In a metallic conductor, on the other hand, the Fermi energy lies inside one of the energy bands, so that in the **k** space there will be a surface called the "Fermi surface," at every point of which the energy of the topmost occupied band equals the Fermi energy. Most of the experiments in a metal give information about the Fermi surface. I shall illustrate this first with the case of copper. In figure 6 I show the energy bands of copper that, as previously noted, were determined, in agreement with each other, by Segall with the KKR method and by Burdick with the APW method. Thus figure 6 represents both sets of calculations. The Fermi energy is in this figure. If we find the Fermi surface in the Brillouin zone, it is as given in figure 7a: more or less spherical but with necks extending out to the surfaces of the zone. When we remember that the energy is a periodic function of position in **k** space, we see that this surface extends from cell to cell, as in figure 7b, with necks joining the sphere in one cell to that in the next.



DOUBLE-ZONE REPRESENTATION of complete Fermi surface for holes in metallic gadolinium. Surface was determined by APW method and is shown for one unit cell. —FIG. 8

Fermi surfaces

Let us describe a few of the types of experiment that can give information about such Fermi surfaces. If the crystal is placed in a constant magnetic field, one can show that an electron will describe an orbit that is the intersection of a constant-energy surface like the Fermi surface with a plane perpendicular to the magnetic field. Such an intersection is shown in figure 7b, a case in which the path of the electron extends indefinitely, going from one cell of **k** space to the next. If, however, the magnetic field were in another direction, the path might be a closed one, for instance like a circulation around one of the necks or around one of the open spaces between spheres shown in the figure. The frequency of such rotation will be proportional to the magnetic field but will obviously depend also on the field orientation and on the orbit size in **k** space. These frequencies can be found by experiments such as cyclotron resonance, and the areas of the orbits can be discovered by other experiments such as the de Haas–van Alphen effect. With such experiments the form of the Fermi surface for copper had been found experimentally before Segall's and Burdick's calculations were made; all the dimensions were worked out with great accuracy. The resulting surface is identical in all respects with that from the calculations.

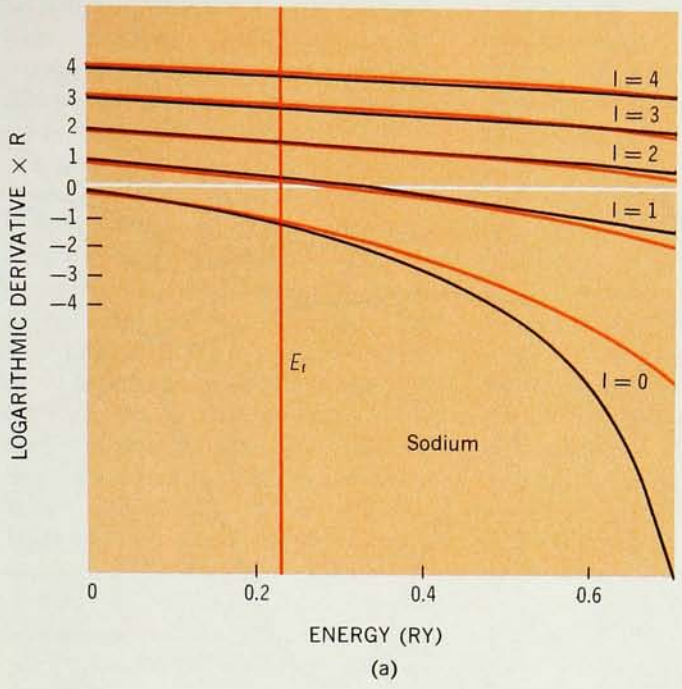
The Fermi surface of copper is one of the simple ones. As an example of a more complicated one, we see in figure 8 the Fermi surface of metallic gadolinium, as determined with the APW method by Arthur Freeman, John Dimmock and Richard Watson. This surface is shown in only one unit cell of the hexagonal structure, analogous to figure 7a; it would be too complicated to show if it were re-

peated, as in figure 7b. But here again, as in copper, there are many experiments that verify the correctness of the energy-band calculations that have led to this Fermi surface. A new breed of solid-state physicists has grown up calling themselves "Fermiologists," who specialize in finding Fermi surfaces from experiment and theory and comparing them. As the reader will suspect, they could hardly operate without having draftsmen who are experts in perspective drawing to prepare their illustrations.

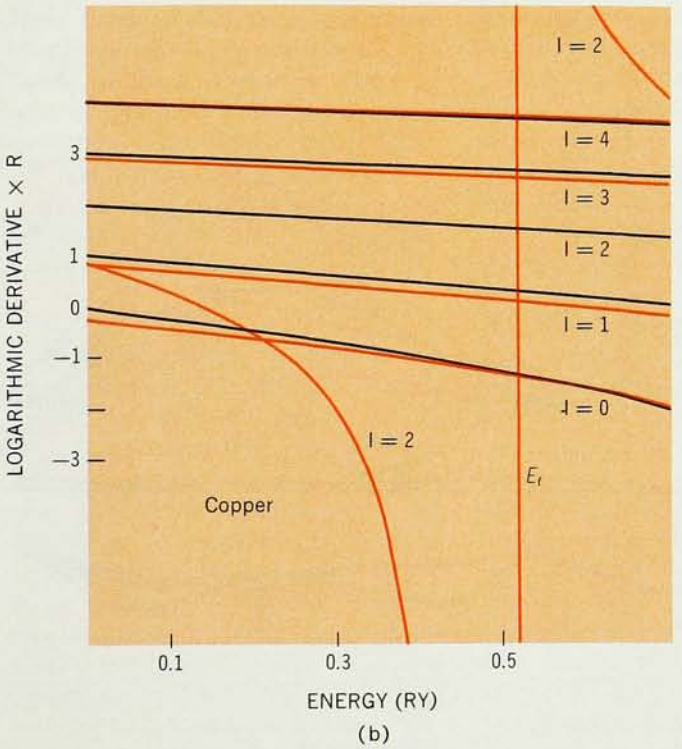
There are a great many additional points that I could discuss on the large number of interesting calculations that have been made for energy bands and Fermi surfaces. However the one that I shall use to conclude with is the role of the 3d bands in ferromagnetism. In the energy bands of copper, in figure 6, we find a large collection of levels lying below the Fermi energy (denoted by E_f). On the left-hand side of the diagram, which corresponds to the center of the Brillouin zone, these levels are denoted by Γ_{12} and $\Gamma_{25'}$, symbols indicating d-like symmetry. These levels form the 3d band and lead to a large density of levels concentrated in a narrow range of energies. This density is found in the ferromagnetic elements, iron, cobalt and nickel—much as in copper, the next element in the periodic system—and is what leads to the ferromagnetic properties of those elements.

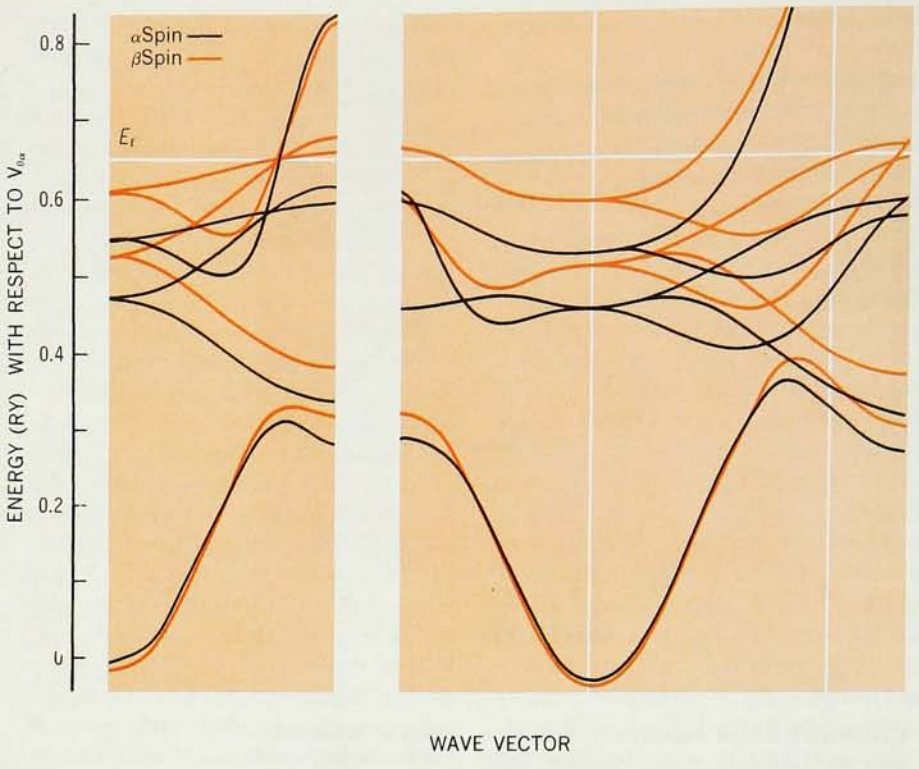
Location of energy bands

First let us see why these bands come where they do. It is essentially because the 3d atomic levels lie close to the 4s and 4p levels that lead to the conduction band in copper. More technically, we can study the bands from the point of view of the APW method. In that method, we match boundary conditions between an expansion in terms of spherical harmonics of angles and radial functions inside the spheres and an expansion in terms of plane waves outside. The plane wave can be expanded in terms of spherical harmonics and radial functions, which in this case are spherical Bessel functions. It could happen that a single plane wave outside the spheres would join smoothly onto the solution inside. Such a situation is very nearly found for sodium. If it were exactly true, each of the spherical Bessel functions representing the plane wave, one for each value of the azimuthal quantum number l, would join smoothly onto the corresponding radial function inside the sphere. This result can be tested by comparing the logarithmic derivatives, the ratio of derivative to function, for the functions inside and outside the sphere, at the radius of the sphere. Such a comparison is made for sodium (in figure 9a) as a function of energy. It is clear that the spherical Bessel functions, represented by black lines, match quite closely with the functions

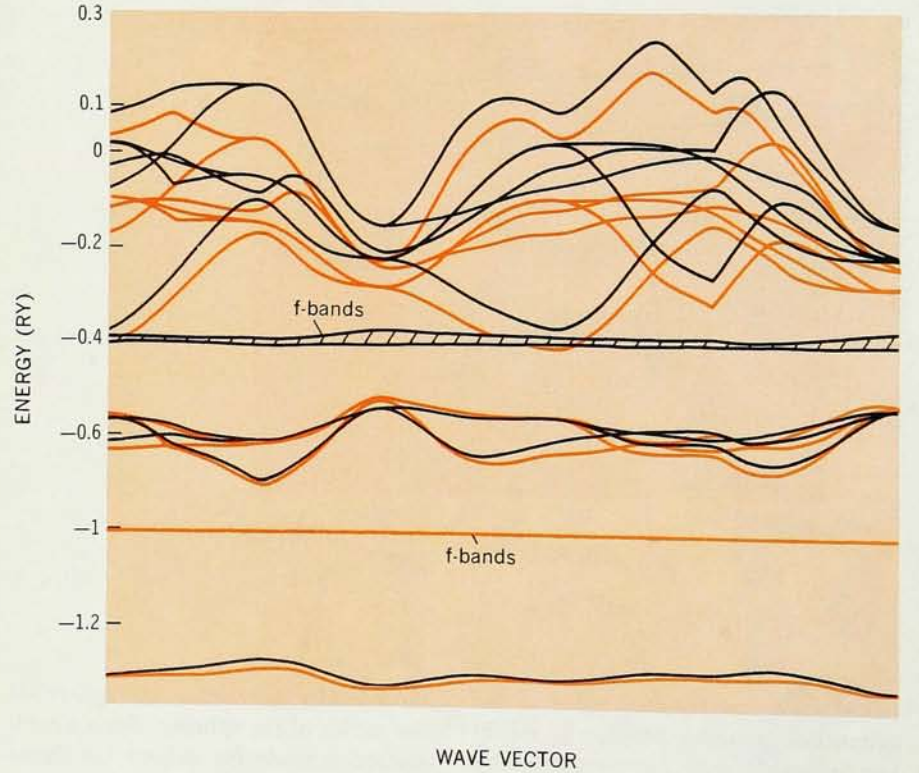


COMPARISON of logarithmic derivatives (ratio of wave-function derivative to wave function itself) inside (color) and outside (black) a sphere. Bessel-function expansions inside and plane-wave expansions outside sphere are matched at the boundary. (a) Sodium atom and (b) copper atom. —FIG. 9





BAND SPLITTING. Each energy band is doubly degenerate—one component corresponding to spin-up electrons, and the other to spin down—with spin up (black) lying below spin down (color). These calculations were made for a nickel atom. —**FIG. 10**



PROPERTIES OF EuS are determined by energy-band calculations. 4f band of spin up, lying below Fermi energy, is occupied, and 4f band of spin down, lying above Fermi energy, is empty. Energy gap from −0.4 to −0.5 Rydbergs accounts for observed semiconductor properties. Spin-up bands are in color. —**FIG. 11**

inside the sphere given by the colored lines.

This agreement has a number of consequences. First, the wave function between spheres is represented rather accurately for sodium by a single plane wave. But further, one can show that energy as a function of **k** is very nearly given by a parabolic relation equal to $k^2\hbar^2/2m$. Still a third consequence arises when we consider the equations, either in the APW or OPW method, for finding energy bands. They are very much like those for the almost-free-electron case that we obtain when we expand the solution of a problem with a periodic potential of very small amplitude in plane waves. These equations make possible a form of treatment called the "pseudopotential method," in which one would set up such a periodic-potential problem of small amplitude, called a "pseudopotential," that leads to the same energy bands as the correct calculation. This pseudopotential method has had a good deal of use as an interpolation method in the same spirit as the tight-binding calculation of figure 5b, in which the coefficients representing the pseudopotential are treated as disposable constants to be used in fitting, as well as possible, the energy bands determined either from experiment or from more accurate calculations.

Now let us see what happens in a case like copper, in which the situation is very different. In figure 9b we show the curves, like those of figure 9a, for copper. Here we see one curve, that for $l = 2$, for which there is no agreement at all between the curves for the functions within the sphere and the spherical Bessel functions. Now $l = 2$ corresponds to the d states, so that this disagreement is a result of the fact that the 3d band comes in the middle of the conduction band. If one looks through the periodic system, he finds that a situation like that in figure 9b appears whenever a new band of levels is added to the atoms. Because of the addition of the 3d band we have the situation of figure 9b all the way from calcium or scandium up to copper. We see that the $l = 2$ curve in figure 9b goes through an asymptotic behavior close to the Fermi energy. As we go beyond copper to zinc, gallium and germanium, the same sort of behavior is still observed, but it goes to negative

energies (the zero of energy in this figure comes at the bottom of the energy band arising from the 4s and 4p atomic orbitals) corresponding to the fact that the 3d level at this point is becoming bound inside the atom. In germanium, the part of the $l = 2$ curve that is coming down from infinity (in the right upper corner of figure 9b) moves down and comes into close agreement with the $l = 2$ curve from the spherical Bessel functions; thus we are back in a situation similar to that for sodium in figure 9a. In such cases the pseudopotential method works well, or we can do well with the OPW method, explaining the success of that method for germanium calculations such as those in figure 5a.

Exchange interactions

Let us now see how these facts fit in with the ferromagnetism of iron, cobalt and nickel, which have energy bands similar to copper. The only difference is that the Fermi energy lies lower, so that the 3d bands are partly empty. This circumstance can lead to the possibility of having more electrons of spin up than spin down, thus leading to a decrease in energy. A particularly interesting way to look at this possibility is called the "method of spin-polarized energy bands." To understand it, we must think a little about the nature of exchange energy; it is a reduction in energy because of exchange that would make the ferromagnetic state stable.

The Fermi hole surrounding an electron is a sphere from which charge of the same spin as that in question is removed because of the Pauli exclusion principle or Fermi statistics. Other electronic charge excluded from this sphere reduces the energy since there is less electronic charge at small radii to exert a repulsive interaction on the electron in question. Now the denser the electronic charge, the smaller the radius of this Fermi sphere, for it is of just such a size to hold one electron. And the smaller the sphere, the larger the magnitude of the exchange interaction since this interaction is inversely proportional to the radius of the sphere.

In a magnetized crystal, there will be different densities of electronic charge for electrons of spin up (that is, parallel to the magnetization) or spin down. The spin-up electrons will then have an exchange effect that will be larger in magnitude than the spin-down electrons. And since the exchange comes in with a negative sign, this means that spin-up electrons are more stable than spin-down ones. In simple language this is an explanation of the exchange effect that tends to line up electrons parallel to each other in a ferromagnetic crystal. But this process has the effect of splitting each energy band into two, with one corresponding to spin-up electrons and the other to spin-down electrons, the spin-up lying below the spin-down. We illustrate this by John Connolly's recent calculations on ferromagnetic nickel (figure 10); we see the similarity to the bands of copper shown in figure 6, aside from the doubling.

The Fermi energy in figure 10 demonstrates that there will be more occupied states for spin up below this energy than for spin down. The reason is that the complete spin-up d band lies below the Fermi energy, while Fermi energy cuts into the spin-down d band and leaves vacancies in the band. This results in net magnetization. A type of self consistency is possible: Here there is enough magnetization to split the spin-up from the spin-down bands so there is enough excess of spin-up electrons to account for magnetization. Connolly has achieved this self consistency. The net result is very similar, though described in some what different language, to the conclusions in my 1936 paper.

Magnetic properties

Calculations, similar to Connolly's in figure 10, have been made for a number of ferromagnetic metals, by S. Wakoh, Jiro Yamashita, Peter DeCicco and various other workers. In each case the results agreed with our knowledge of magnetic properties of crystals as well as leading to Fermi surfaces consistent with other types of experiments on crystals. As a last example, we show figure 11, Sang-Jean Cho's calculations not on a ferromagnetic metal but on EuS, an ionic compound with sodium-chloride structure; EuS is known experimentally to be ferromagnetic and also a semiconductor. Cho's calculated bands manifest both these properties. Let us explain the essential features of the energy bands.

The Eu^{2+} ion has a half-filled shell of 4f electrons holding seven electrons. The ground state of this ion has the spins of all seven electrons parallel, thus leading to a strong magnetic moment. Below the Curie temperature, all magnetic moments in the crystal set themselves up parallel to each other. Then because of the exchange effect, the energy bands for 4f electrons with spin up will lie far below those for spin down. However, the 4f orbital is so far inside the atom that there will be practically no broadening of the band in the crystal. In figure 11, we see the 4f band of spin up (so narrow that its broadening cannot be shown) well below the Fermi energy. The 4f spin-down band, however, lies above the Fermi energy, so that it is empty and therefore in agreement with experiment.

Above the 4f spin-up band there is another occupied band. Then there is a gap between this occupied band and the next higher band that is unoccupied. This gap leads to the semiconducting properties of the crystal. Cho's calculation has not been carried to self consistency, as has Connolly's for nickel, and some details may be changed when this work is carried out, However, it appears likely that the essential features are as I have described them here. It is interesting to inquire about the physical meaning of the large splitting of energy between the occupied spin-up 4f band and the empty spin-down band. Closer analysis shows that the transition that would carry an electron from the lower to the upper band, with reversal of its spin, would be a transition inside the Eu^{2+} ion. One of the 4f electrons reverses its spin so that instead of being in the multiplet of highest multiplicity, an ^8S spectroscopic state, it would be in a sextet state with lower multiplicity. The energy separation between these bands corresponds roughly to the energy difference that would be encountered in such a transition.

In these magnetic examples we have shown some, but by no means all, of the types of application that have been made for energy-band theory of the magnetic behavior of crystals. Other recent applications are to antiferromagnetism and to the remarkable helical arrangements of magnetism that are being found in some of the rare-earth metals. Here, too, it has been possible to make energy-band calculations that have thrown a great deal of light on the actual behavior of these magnetic crystals. □