PROCEEDINGS OF THE I.R.E.

THE TRANSISTOR ISSUE

1952

PROCEEDINGS OF THE I.R.E.

Published Monthly by

The Institute of Radio Engineers, Inc.

VOLUME 40 — *November, 1952* — NUMBER 11

PROCEEDINGS OF THE I.R.E

Contents continued on page 1282

PROCEEDINGS OF THE I.R.E.®

Published Monthly by

The Institute of Radio Engineers, Inc.

The Transistor Issue

The Board of Directors of The Institute of Radio Engineers, in its continued endeavors to meet more fully the needs of the membership, has established the precedent of authorizing on a nonscheduled basis special issues of the PROCEEDINGS OF THE I.R.E. Each such special issue is devoted, in timely fashion, to new and basically important portions of the domain of communications and electronics.

This present, greatly expanded issue of the PROCEEDINGS, the second such special issue to be published, is one assigned to a novel, significant, and rapidly developing field, namely, the broad subject of semiconductors and those of their properties and applications of primary interest to the membership of the Institute. Specifically, it is concerned also with electronic devices based on germanium, thus including the whole class of components known as "transistors."

Upon its emergence four years ago, the transistor promised to be one of the truly major advances, even in an art which has been accustomed to giant-stride progress. The fulfillment of this promise, however, depended on the difficult task of determining how to transform the transistor from a laboratory oddity to a practical device which could be manufactured uniformly and in quantity. Marked strides of recent months has furthered progress rapidly toward this goal, so that the threshold of activity has now been reached where the many capabilities of the transistor may be systematically explored, developed, and widely utilized.

The subject is a vast one, and under rapid development. There have, nevertheless, been gathered here a large selection of significant and instructive original papers covering both theory and practice, which it is believed will constitute a substantial and timely contribution to knowledge and expansion of this new and important field.

—*The Editor.*

Transistor Lineage

THE EXTENSIVE and rewarding research and development work already carried out in the transistor field fully justifies the dedication of this issue of the PROCEEDINGS OF THE I.R.E. to the electronic aspects of certain semiconductors. In a similar sense, this issue may be regarded as a tribute to those effective and tenacious workers and organizations many of whose accomplishments are described herein.

It seems timely and permissible also to endeavor partly to orient the transistor in certain of its historical and evolutionary respects.

Students of scientific advances often find that these largely issue from the operation of two factors, namely, progressive cyclic history, and inventive dependence on availability of instrumentalities.

Progressive cyclic history involves the recurrence of events or material arrangements in a form strongly suggestive of, or broadly based on earlier occurrences or assemblies. In brief, history repeats itself—but always with a difference. Examples of this cyclic phenomenon are modern microwave developments as a recent vastly improved version of the original Hertzian wave experiments. Radar obviously suggests the reflections of electromagnetic waves, both in the optical and radio fields as long ago discovered. Coherers and crystal detectors seem to be the remote and primitive ancestors of the semiconductive rectifiers and amplifiers of today. There are many other examples of the repetitive tendency which will occur to the readers.

It should be noted, however, that the recurrent or later invention is always an elaboration, improvement, or inspirational modification of earlier knowledge, and forms a substantial addition to mankind's store of information. For this reason, the inventive genius involved in the later developments often equals or even transcends that involved in the earlier and analogous case, even though historically the later work may sometimes be overshadowed by the pioneer aspect of the earlier inventive work.

The second factor mentioned in connection with scientific advances is one which is frequently overlooked. Inventive dependence on the availability of suitable instrumentalities is simply the obvious need for some physical materials or methods whereby an invention may actually be embodied in usable form. Inventions, in the effective sense, can usually be made only after the necessary factors for their physical creation are at hand or, by reasonable expectation, are soon to be at the inventor's disposal.

It would hardly have been possible, for example, to have invented the picture tubes of television before electron-beam excited phosphors were known, nor could television camera tubes have been produced before controllable photosensitive surfaces were available. Yet, great genius may be displayed in proposing nonrealizable and abortive invention. Nevertheless, mankind reserves its approval for the actual accomplishment. A similar comment is appropriate to the effect that invention, which too far precedes the availability of correct guiding principles, arrangements, or materials, may tend dangerously toward "science fiction."

Considering both these historical factors, the transistor is seen to have an honorable lineage, to be a worthy and advanced member of its family, and to have been born when electronic, chemical, and metallurgical knowledge in this field was ripe. Aside from its excellent ancestry, the transistor appears to have a bright future. Its compactness, light weight, long operating life, low power consumption, novel circuitry, and other factors suggest future accomplishments of value to mankind.

At this time, it is of course impossible to predict what further devices and principles will be evolved from the transistor at the next turn of the wheel of scientific and technical history, nor yet what new discoveries and instrumentalities will be needed to open the doors to such later evolution. One can be confident only that there are many remaining treasures in nature's storehouses awaiting the enterprise of newer pioneers of a later day.

—The Editor.

Transistors in Our Civilian Economy

J. W. McRAE

At relatively long intervals there appear on the technical and industrial horizons devices of such broad scope and major significance that they profoundly affect the fields of their use. One of these epochal developments is the transistor, which bids fair to take its place beside the electron tube as one of the foundation stones of future communications and electronics.

It is accordingly timely and suitable that certain of the probable future industrial uses and effects of the transistor should be here analyzed in a guest editorial by an engineer especially qualified for this task, and who is a Fellow and Director of the Institute, and a Vice President of Bell Telephone Laboratories.—*The Editor.*

IT IS NOW a little more than four years since transistors were first announced. The youngest members of the new family, junction transistors, are not yet two years old. What can be said, at this early stage in their history, about what transistors will mean to our civilian economy and to IRE members?

Vacuum tubes had been in existence about eight years when they had their first dramatic successes as amplifiers and oscillators. Thirty-seven years ago, in 1915, they made possible the first very long distance transmission of speech, by wire line across this continent and by radio across the Atlantic. No one could then have predicted the present situation. It would have been impossible to conceive of a production of close to half a billion vacuum tubes a year or an IRE membership exceeding thirty thousand.

The fact is that vacuum tubes not only performed their original functions as detectors of radio waves; they made possible entirely new techniques for use in transmitting information. The scope of existing communications services was greatly expanded and whole new industries were brought into existence.

More recently, vacuum tubes have been applied, not only to transmit information, but also to manipulate it and put it directly to use. Electronic computers now work out intricate problems in the laboratory and solve the equations necessary to aim and fire anti-aircraft guns. Through servomechanisms, they aim the guns and control other operations involving considerable power with finesse and precision. All of this is unified with the field of information transmission through the use of similar devices and techniques as well as by a common set of theoretical principles.

Are transistors to be merely competitors for these jobs now being done by vacuum tubes? Or can they, in their turn, further extend the field of electronics? It is most important to examine this second question. There will undoubtedly be many cases in which transistors will be able, by replacing vacuum tubes, to reduce the cost and increase the convenience or effectiveness of existing services. But the benefits will be greater if transistors can penetrate into portions of our present services where vacuum tubes are not yet widely used. It will be still better if some new area of electronics can so capitalize on transistors as to be able to expand significantly.

Two examples will serve to illustrate the fact that both possibilities have substance. The first is our local telephone system in which there are essentially no vacuum tubes. The only widely used amplification between the subscriber and the central office is provided by the carbon microphone or telephone transmitter. It supplies the necessary gain, is inexpensive and reliable, and conveniently requires only a small direct current to supply its power. Because of its reliability it can be located at many scattered points without incurring excessive maintenance costs. In the great majority of cases, its power requirements can be satisfied by transmission of power from the central office, so there is no necessity for relatively expensive local sources of dependable power.

Power requirements should be considered in terms of the power levels of the signals to be handled. A good many situations in communications require that the signals be brought to high power levels. But in this particular situation one microwatt of signal power produces readily audible sound in a telephone receiver. This is thousands of times smaller than the power required just to heat the cathode in one of the smallest commercial vacuum tubes. It is a million times smaller than the cathode heater power in most of the vacuum tubes designed for suitable performance, long life, and reliability in telephone service. These vacuum tubes also require comparable amounts of plate power, usually at voltages and currents difficult to supply over local telephone lines.

The transistor promises to combine long life and reliability with the ability to amplify very small signals at high efficiency. It should prove possible to supply the small amount of necessary power over the line. Transistors may be able therefore to extend electronics into this area; which the vacuum tube has been unable to do.

Undoubtedly there are other similar situations where the need is for the *transmission* of information; but consider now the second illustration, computers and other automata which *manipulate* information. Vacuum tubes made possible electronic analogue computers and in such computers have met with little competition from other devices. On the other hand, they have met stiff competition from electromechanical devices such as relays, in digital computers, and in other automatic machines or systems which work by digital methods. Automatic business machines and telephone switching systems are examples. With few exceptions, vacuum tubes have been used only when their great speed overcame their disadvantages in cost and reliability.

These applications generally involve the use of large numbers of individual devices such as relays, switches, or vacuum tubes. Questions of the reliability and the power consumption of each individual device therefore loom large. Here again the operations to be performed can for the most part be done at very low power levels. The power required to operate relays or heat the cathodes of vacuum tubes is many times larger than the needed power of signals to be handled in each device.

Transistors and their ancestors, germanium diodes, appear to make a natural team for such applications. They ought to make it possible to close somewhat the great gap that now exists between our best electronic "brains" and the human brain, in regard to size and power consumption. At the same time, they promise better reliability and speed comparable to vacuum tubes. The future may well see a considerable expansion in the industrial use of complex automata. Perhaps it is significant that the first commercial use of transistors is expected to be in a part of the "brain" of the new telephone national toll dialing system.

These two illustrations demonstrate the possibilities for extensions in electronics as the result of the advent of transistors. However, it must not be assumed that the engineering job ahead will be easy. Transistors are not vacuum tubes, and many of our present circuit design practices, developed with vacuum tubes in mind, will not be applicable to transistors. It will often be necessary to analyze these new circuit problems in terms of those fundamental principles which have more general application.

Those directly concerned with the design and manufacture of transistors will get little assistance from the lore of vacuum tubes. They will have to turn to the fundamental physics of conduction in semiconductors. While this may be unfamiliar territory, they will find it well charted. Indeed, our present physical understanding of semiconductor devices is remarkably mature, very much better than our corresponding understanding of vacuum tubes in 1915. This gives us a good basis for confidence in future progress and it should help us to accelerate that progress.

Transistors and the Military

I. R. OBENCHAIN, JR. AND W. J. GALLOWAY

Many useful applications of the transistor will obviously be found in science and engineering, as well as in certain specialized fields. The anticipated technical applications will mainly be of industrial and military nature.

It is accordingly both appropriate and helpful to present here a guest editorial dealing with the military uses of transistors and prepared by experts in that field.

At the time of the preparation of this editorial Lieutenant Colonel I. R. Obenchain, Jr., Signal Corps, was Assistant to the Commanding Officer for Research, Signal Corps Engineering Laboratories; and First Lieutenant W. J. Galloway, Signal Corps, was a member of the staff in the Office of the Director of Research, Signal Corps Engineering Laboratories. The authors are respectively a Senior Member and an Associate Member of the Institute.—*The Editor.*

THE TRANSISTOR, and the unquestionable impact it may exert on electronic design as we know it today, have rapidly become major centers of interest. This impact will be important from two points of view: engineering and economics.

Having at least some bearing on both are questions of transistor application in large-scale national rearmament. For example: Exactly what are the military services doing about transistors? Can the Army use them for a wrist-watch radio? Can transistors substantially reduce the weight of electronic equipment in the Air Force's bombers? Will the Navy equip torpedoes with "transistorized" electronic brains?

Answers to the last questions necessarily stem from the results of service transistor programs, but the first is readily answered since military interest in this field has been close and continuous since 1948. Due to the difficulty of quantity production of transistors with uniform characteristics, immediate application to military equipments was impractical up to the spring of 1951. Evolution of the junction transistor in 1951, plus quantity production ability with uniform characteristics for both these new type low-noise units as well as the older point contact types, changed the previous somewhat restricted approach of the military services to an active and expanding program. This uniformity of performance is of course a necessity for military use, and consequently, the services are interested only in transistors that can be built to meet prescribed specifications.

Thereupon, the military services embarked upon substantial transistor programs in the summer and fall of 1951, which emphasized continued solid-state research, specific research to improve transistors, circuitry research for application of transistors, and the preparation of facilities for transistor production. The eventual impact of transistors on military electronic equipments was considered sufficient to form within the Defense Department's Research and Development Board an Ad Hoc Group on Transistors. This Ad Hoc Group was established under the Committee on Electronics as a high-level group to set broad policies for the co-ordination of transistor programs within the three services. Shortly afterwards a Subpanel on Semiconductor Devices was formed under the RDB's Panel on Electron Tubes. This latter group performs the normal subpanel functions of co-ordinating service projects and then passing them on to the Ad Hoc Group for final approval. On eventual dissolution of the Ad Hoc Group the Semiconductor Subpanel will take its normal position with respect to the Panel on Electron Tubes such as is now held by the other subpanels of that group—RF tubes, receiving tubes, and so on.

To speed active interest for continued research and development work on transistors and related devices (for instance, barrier rectifiers, and the like) including investigations of the use of materials other than germanium, the military services are presently supporting substantial contractual programs. Most of the contracts are joint service (equal support from the Army, Navy, and Air Force) in nature, with the services alternating as administrators for each new contract. The military are also aware of and intensely interested in the large amount of work being applied to this problem by industry without government support. A modest program in this general field is maintained on an internal basis within the various military research and development laboratories. It is hoped that this broad approach, especially the nongovernment supported work of private industry and educational institutions and the effort under special research and development contracts supported by military funds, will soon provide transistors with higher power ratings, higher frequency response, lower noise, and the ability to operate satisfactorily over wide ranges of temperature.

"Battle is the Payoff" is a slogan within the military. This thought is applied to the service transistor programs in determining where and how these new units may be used to the maximum advantage in military

equipments. The favorable characteristics of transistors, namely, no requirement for heater operation, high efficiency, resistance to shock, long life, and small size have already been widely publicized but are here re-emphasized as exceptionally desirable characteristics for active elements to be included in military equipments. The possible savings from a power standpoint alone are impressive when one remembers the considerable quantity of batteries required for hand-carried equipments during World War II—for example, the thirty-day supply of batteries for a certain type of amphibious communications company totaled forty-three tons. In general, the military services' transistor applications programs are aimed at improving existing types of equipments, making them lighter or more reliable, or providing new types of equipments which with vacuum tubes had not been considered feasible for military use. An example of the latter type of development would be small compact "transistorized" computers in which the use of several hundred vacuum tubes would have been impractical from a size and power consumption viewpoint. These programs are being conducted on a contractual as well as on an internal basis within government laboratories and will be greatly expanded during fiscal year 1953, as more transistors become available for the services to furnish to contractors working on application studies. Availability of transistors, however, is not the only problem; applications experience has confirmed the difficulties that circuitry engineers were expected to encounter in utilizing these new devices. There is a whole field of engineering circuitry data that is yet to be obtained for each of the general transistor types, and apparently the greatest results in the treatment of this problem can be obtained by the well-known methods of network analysis. Another typical problem associated with the application side of the program is how to make a small, economical battery source of constant current and long life. The large variation with temperature of the present type transistors' equivalent circuit parameters is also of particular difficulty, since so many military equipments require operation over extended temperature ranges.

Associated with the circuitry application programs, emphasis is being placed on the development of circuit packages wherein complete circuit functions are "potted" into plug-in units. Such items as oscillators, counting circuits, IF amplifiers, and so on, fall in this category. Investigations are also being conducted towards developing smaller passive components which will take advantage of the lower voltages utilized in transistor circuits to provide a minimum size component.

In more normal times the military services would embark on only a modest program of "transistorization" leaving the broad general problem of the maximum utilization of these devices to the ingenuity of our industry and research institutions. Now, however, in this period of international tension the services consider the possible benefits of transistors to military equipments as sufficient to warrant substantial programs in this field and to include concurrently not only research and development on the transistor itself and its application to equipments, but the planning and preparation of facilities for producing large quantities of these devices. The responsibility within the Department of Defense for planning these transistor production facilities has been delegated to the Army Signal Corps, with top policy guidance from the RDB Ad Hoc Group on Transistors. This program includes contracting for the establishment of factories, industrial mobilization study contracts, consideration of raw material availability, and the placement of supply orders. The concurrent establishment of production facilities and enlarged device research and development programs as well as increasing effort on applications faces one with the very real problem of what transistor types to prepare for in quantity.

The military are in fact now faced with the problem of preparing a filling feast without the unnecessary expense of a smorgasbord for very fussy guests while having only the slightest indication of their preference.

Fortunately, there is reason to believe that transistor factories can vary types of transistors produced without prohibitively costly modifications to capital equipment; however, there is still no solution to the use of stockpiles which could suddenly become outdated by the overnight discovery of a radical improvement.

It is hoped this provides at least some insight to the military's rather substantial interest in transistors. In present form, transistors will undoubtedly rather quickly fill a number of military requirements and will also provide devices for training of personnel engaged in circuitry studies. As they improve it is certain that additional applications will follow with increasing rapidity. With transistor pilot-production lines now starting to roll, the immediate need is to develop a new circuitry "know-how" which for electron tubes, took over a quarter century in its evolution toward current engineering practices. The military services will do everything within their means to compress the new circuitry and application studies into the smallest time interval possible, and in so doing will look largely to the ingenuity of the IRE membership for the assistance to realize the maximum utilization of new devices.

Transistor Electronics: Imperfections, Unipolar and Analog Transistors*

W. SHOCKLEY†

Summary—**The electronic mechanisms that are of chief interest in transistor electronics are discussed from the point of view of solid-state physics. The important concepts of holes, electrons, donors, acceptors, and deathnium (recombination center for holes and electrons) are treated from a unified viewpoint as imperfections in a nearly perfect crystal. The behavior of an excess electron as a negative particle moving with random thermal motion and drifting in an electric field is described in detail. A hole is similar to an electron in all regards save sign of charge. Some fundamental experiments have been performed with transistor techniques and exhibit clearly the behavior of holes and electrons. The interactions of holes, electrons, donors, acceptors, and deathnium give rise to the properties of *p-n* junctions, *p-n* junction transistors, and Zener diodes. Point-contact transistors are not understood as well from a fundamental viewpoint. A new class of *unipolar* transistors is discussed. Of these, the *analog* transistor is described in terms of analogy to a vacuum tube.**

1. Perfection and Imperfections in Crystals

TRANSISTOR ELECTRONICS exists because of the controlled presence of imperfections in otherwise nearly perfect crystals.[1] Many of the special words of the subject are the names of imperfections; hole, excess electron, and donor are examples.

In this article we shall describe these imperfections from the point of view of theory; again from the point of view of theory we shall show how the interplay of imperfections permits useful electronic devices to exist. We shall also indicate how the new experiments of transistor physics display the attributes of the imperfections and lead to measurements of their quantitative attributes.

Before discussing imperfections in a germanium crystal we must discuss perfection. Perfection is represented by an ideal crystal built on the plan of Fig. 1. Each atom has four neighbors, all at the same distance from it and all at equal distances from each other.[2] This arrangement satisfies perfectly the chemical affinity of the germanium atom.

For the purpose of transistor physics the germanium atom may be considered to consist of four valence electrons surrounding an inert core having a charge of plus four electronic units. The atom as a whole is thus neutral. Silicon and carbon atoms have this same structure and their crystals form in the same pattern as Fig. 1.

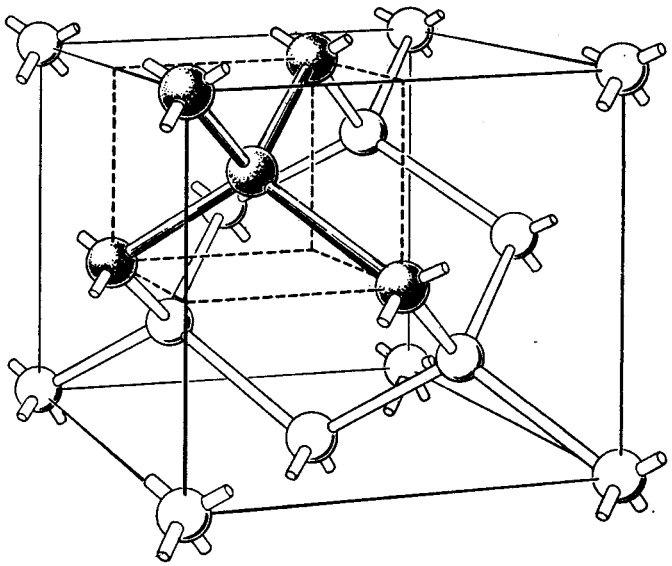


Fig. 1—Diamond structure.

The core of the germanium atom consists of the nucleus, whose charge is plus 32 electron units and 28 tightly bound electrons. These latter are closely held to the nucleus, as represented in Fig. 2, and do not enter into chemical reactions or transistor effects.

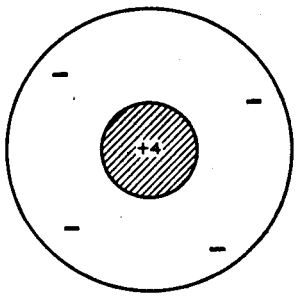


Fig. 2—The four valence electrons and positive core of a silicon or germanium atom.

The outer valence electrons are responsible for chemical binding. One of the principal building blocks of chemistry is the *electron-pair bond* or *covalent bond*. Such a bond may form when two atoms approach each other. In order to form the bond a pair of electrons, one from each atom, co-ordinate their motions so as to produce an accumulation of negative charge between the cores. This attracts the cores towards each other. Equilibrium is achieved when the attraction is balanced by electrostatic repulsion of the cores for each other.

The laws of physics which explain the behavior of the electron-pair bond are *quantum* laws and employ *wave-mechanics* to describe the motions of the electrons. These

* Decimal classification: R282.12. Original manuscript received by the Institute, August 26, 1952.

† Bell Telephone Laboratories, Inc., Murray Hill, N. J.

[1] The subject of imperfections has recently been reviewed in a book, "Imperfections in Nearly Perfect Crystals," Editorial Committee, W. Shockley, Chairman; J. H. Hollomon, R. Maurer, F. Seitz; pub. by John Wiley and Sons, New York, N. Y.; 1952.

[2] The reader with an instinct for spatial visualization may enjoy thinking about the problem of surrounding one atom with several symmetrically placed neighbors. He will find that four is the largest number of neighbors that can surround one atom at equal distances and be equidistant from each other.

laws explain why two electrons, rather than one or three, unite in forming the most stable bonds. We shall need to refer to this feature of the stability subsequently.

The germanium crystal gives each atom an ideal opportunity to form electron-pair bonds. Each atom is surrounded by four symmetrically placed neighbors. It forms one covalent bond with each of these. The electrons in each bond are shared equally by the two atoms at its ends, one electron coming from each atom.

Fig. 3 symbolically represents the bond structure of the crystal. In it the atoms lie in a plane rather than in space, but each atom is correctly surrounded by four neighbors. A crystal in which this bond structure is complete and every atom is at rest in its proper place, with no atoms or electrons missing or squeezed into wrong places, is in the *perfect* or reference state.

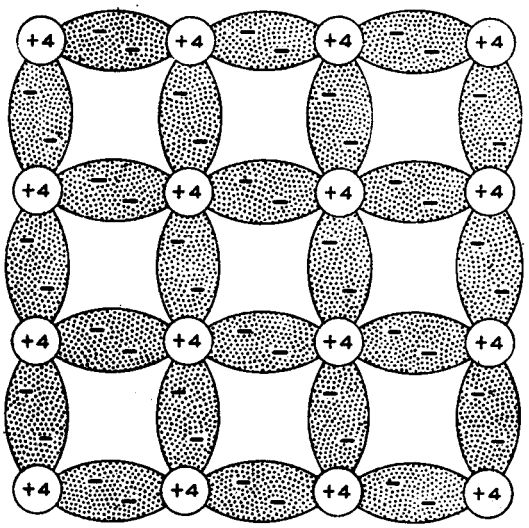


Fig. 3—Representation of electron-pair bond structure in silicon or germanium.

The perfect state is relatively uninteresting from the transistor point of view. A germanium crystal in its perfect state would be an insulator. Its dielectric constant, however, would have the unusually high value of 16. The effect of an electric field in producing polarization is represented in Fig. 4. Since the electric field exerts forces in one direction upon the cores and in the opposite direction on the bonding electrons, the structure is distorted. There is evidently a relative shift to the right of positive charge in respect to negative charge. This displacement is linear and elastic in the electric field and disappears when the field is removed. (Fig. 4 shows an exaggerated case corresponding to a field of about 10^7 volts/cm, much more than can be applied without breakdown effects.)

On the basis of perfection as represented by Figs. 1 and 3, the imperfections of importance in transistor electronics can now be described. In Table I these imperfections are listed under three classifications.

As is implied by their names, the energetic imperfections involve specialized concepts from quantum theory.

TABLE I

CLASSES OF IMPERFECTIONS

Energetic	photons (light), phonons (heat)
Electronic	(excess) electrons, holes
Atomic	disordered atoms (vacancies and interstitial atoms)
	chemical impurities (donors and acceptors)
	"deathnium"
	"traps"

The photon is the name for a quantum of light energy. This energy comes in packets of size

$$\text{energy} = h\nu, \tag{1}$$

where h is Planck's constant and ν is the frequency of the light. When light is absorbed, the energy is delivered in entire quanta.

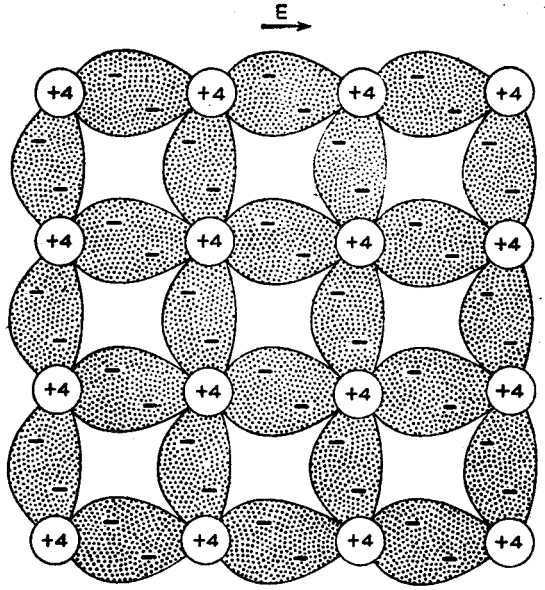


Fig. 4—Distortion produced by an electric field resulting in dielectric polarization.

Less generally familiar is the fact that mechanical vibrations of solids are also quantized. A beam of sound waves transmitted through a solid is composed of phonons just as a light beam is of photons. Just as a cavity is full of *black body radiation* under thermal equilibrium, so is a crystal full of a thermal equilibrium distribution of phonons. The phonons thus represent thermal vibrations and account for the specific heat of the solid. We shall be concerned with the phonons because of the way in which they interfere with electronic motion in the crystal. This is discussed more fully in the next section.

The other classes of imperfections involve particles. In the electronic class there are two possibilities, corresponding to adding and subtracting an electron. These are discussed in Section 2.

In the atomic class, it is also possible to add and subtract atoms of germanium from a germanium crystal. There are also other ways of disordering the crystal without introducing atoms of different kinds. We shall not deal with these here. The foreign atoms of chief interest are donors and acceptors, which come from columns of the periodic table adjacent to the germanium column.

Two classes of imperfection are shown described by their properties rather than by their nature. It is known that both the action of *deathnium* and that of *traps* can be accomplished by disorder. The same effects may be accomplished by impurities. These atomic imperfections are discussed in Section 3.

One important class of imperfection that has been omitted from Table I is the *grain boundary*. A grain boundary occurs when two crystals, each having a structure like Fig. 1, are tilted in respect to each other so that their axes are not parallel. Where these two crystal *grains* join, there is a region of atomic misfit called a "grain boundary." Grain boundaries have pronounced electrical effects in germanium, and are known in metallurgy to be loci for the precipitation of impurities.[3]

Grain boundaries are not considered in this article because they can be eliminated. As a result of stimulation from the desires of the transistor development group for highly uniform bulk material Teal and Little undertook a program of growing large single crystals of germanium.[4] These single crystals, weighing up to several hundred grams, have the same orientation throughout and no grain boundaries. For the last few years, practically all advances at Bell Telephone Laboratories in transistor electronics and transistor physics have been based on the availability of single-crystal material.

The importance of artificially grown single crystals in transistor work is not unique. The same situation holds true in all branches of solid-state physics; magnetism, plasticity of metals, piezoelectricity, and the study of order-disorder transformations in alloys are examples. In these cases, the availability of artificial single crystals has had a very marked stimulus upon the development of the science and the technology.

In this article emphasis is placed chiefly upon germanium with silicon also receiving some attention. This is appropriate in view of the current status of development. It should be mentioned, however, that these elements are not unique. Transistors have also been made using lead sulfide (galena)[5] and lead selenide.[6]

2. The Behavior of Holes and Electrons

A. *Introduction*

In this section we shall be concerned chiefly with the behavior of electronic imperfections. These imperfections play the essential role of current carriers in transistor electronics. A large part of this section is devoted to the interactions between electronic imperfections and the energetic imperfections mentioned in Section 1. It turns out to be convenient to include, at the same time, a discussion of deathnium, although it is not an electronic imperfection.

As most readers of this article probably know, the two important electronic imperfections are the *excess electron* and the *hole*. These are produced by adding and subtracting, respectively, one electron from the perfect situation. The behavior of the *excess electron*, or simply *electron* for short, can be understood relatively simply from an intuitive approach. This is not true for the hole. In fact, intuition leads to quite erroneous conclusions regarding the hole. In this section, the electron is, therefore, discussed first. Once the behavior of the electron has been described, that of the hole can be readily explained, even though the reasons for the behavior cannot. It is the author's opinion, based on several years' experience with these expositional problems, that the theoretical basis for the behavior of the hole cannot be simplified and presented on an intuitional level. Fortunately, the behavior itself is relatively simple and can be verified by direct experiment. The engineer who wishes to think in terms of holes may thus have confidence in the concepts as experimentally established engineering factors.

B. *The Behavior of an Electron*[7]

The simplest electronic imperfection is produced by adding one extra electron to the otherwise perfect crystal. From the electrical point of view this extra electron represents one negative unit of charge. The electric field set up by this charge is, of course, screened by the dielectric constant of the germanium, which has a value of 16. Thus the force which the electron will exert upon another charge is reduced by the effect of the medium upon its electric field. If we suppose that the energetic disturbances in the crystal are present in negligible proportions, then the behavior of the added electron is very simple. If it is added to the crystal with zero kinetic energy, it will stay substantially in the same place indefinitely. It is not bound to one place, however, as is an electron in a valence bond. The high stability of an electron-pair bond is restricted to two electrons. A third electron cannot fit into this structure and, consequently, the excess electron is free to move. If an electric field is applied, the excess electron will be accelerated by the resulting force. The acceleration is affected by the fluctuating electric fields due to the positively charged atomic cores and the negative valence bonds. For most purposes the effect of this potential, which varies periodically in space, can be lumped into one quantity, the *effective mass* of the electron. Under the influence of a given force the electron in the crystal will have an acceleration different from that which it would have in free space under the influence of the same force.

[3] G. L. Pearson, "Electrical properties of crystal grain boundaries in germanium," *Phys. Rev.*, vol. 76, p. 459; 1949.
N. H. Odell and H. Y. Fan, "Impedance characteristics of grain boundaries in high resistivity *n*-type germanium," and W. E. Taylor and H. Y. Fan, "D.C. Characteristics of high resistance barriers at crystal boundaries in germanium," *Phys. Rev.*, vol. 78, p. 354; 1950. See, also, footnote reference 1 for a review of the theory of grain boundaries.

[4] G. K. Teal and J. B. Little, "Growth of germanium single crystals," *Phys. Rev.*, vol. 78, p. 647; 1950.
G. K. Teal, M. Sparks, and E. Buehler, Proc. I.R.E., vol. 40, pp. 906–909; August, 1952.

[5] H. A. Gebbie, P. C. Banbury, and C. A. Hogarth, "Crystal diode and triode action in lead sulphide," *Proc. Phys. Soc.* (London), vol. B63, p. 371; 1950.

[6] A. F. Gibson, "Absorption spectra of single crystals of lead sulphide, selenide and telluride," *Proc. Phys. Soc.* (London), vol. 653, pp. 378; 1952.
P. C. Banbury, "Double surface lead sulphide transistor," p. 236.

[7] For a more complete treatment and references, see W. Shockley, "Electrons and Holes in Semiconductors," D Van Nostrand Co., Inc., New York, N. Y.; 1950.

The result can be interpreted by saying that in the crystal the mass of the electron is in effect different from what it would be in free space. Experiments indicate that the effective mass of an electron in a crystal is about as large as the mass in free space and may be less by a factor of four in the case of semiconductors of interest to transistor electronics.[8]

The reader unfamiliar with the background of the modern energy-band theory of solids probably finds it surprising that the electron can move through the strong fluctuating electric fields in a crystal as though it were in free space. The reason that it can do this is that the motion of the electron is governed by wave mechanical laws. It is a property of wave propagation that waves can proceed unattenuated through lossless structures which have perfect periodicity. Examples of iterated electrical filters and loaded lines will remind the reader of this fact. A more striking example from electromagnetic theory is that of the artificially loaded dielectric which is used in lens systems for microwave antennas.[9] For the case of electrons in crystals, similar mathematical developments apply. It is found that the electron wave can cause the electron to propagate indefinitely through the crystal provided that the electrical potential in which it moves is perfectly periodic, as it will be if the crystal is in the perfect reference condition.

If energetic imperfections in the form of phonons are present, the perfect periodicity of a lattice will be destroyed. As a result, the electron once set in motion will not move indefinitely in one direction but will instead be scattered so that after a time its motion is changed to some other direction. The laws of scattering of electron waves by interference with phonons can be derived on the basis of quantum-mechanical theory. When the derivation is carried out, it is found that the effect of the phonons can be very accurately simulated by introducing a set of *phantom particles*.[10] These phantom particles are to be thought of as uncharged hard, spherical, ideally elastic masses. Phonons in a germanium crystal at room temperature correspond to phantom masses about 170 times that of the electron mass. Consequently, in any collision the electron does not lose much energy but bounces nearly elastically from the phantom particle. The choice of a spherical shape is dictated by the fact that scattering by phonons is isotropic, that is, after being scattered an electron is equally likely to proceed in any direction. The scattering of a light spherical mass by a much heavier spherical mass also has this isotropic property. The phantom particles should be thought of as moving with random thermal energies. The number of them and their cross sections are such that as an electron moves through the crystal it will, on the average, move for a distance of about 10^{-5} cms between collisions with the phantom particles. We shall refer to this *mean free path* as l; we thus have

$$l = 10^{-5} \text{ cm}. \tag{2}$$

This distance is approximately 1,000 times the interatomic distance between germanium atoms in the crystal.

If the electron is placed at rest in a germanium crystal at room temperature, then it will be jostled by the phantom particles which represent the phonons. As a result of this jostling, the electron will acquire thermal energy. A general theorem from statistical mechanics applies to this situation. This theorem, known as the principle of equipartition of energy, states that the average kinetic energy of a free particle at temperature T is given by the formula

$$\text{thermal kinetic energy} = (3/2)kT, \tag{3}$$

where T is the absolute temperature in degrees K and k is Boltzmann's constant. Since the electron has only about 1/170 the effective mass of a phantom particle, its speed will be 13 times greater. If we take the electron's mass as equal to the mass of a free electron and define a thermal velocity by the equation

$$mv_T^2/2 = kT, \tag{4a}$$

then we find that at room temperature the thermal velocity of an electron is

$$v_T \doteq 10^7 \text{ cm/sec}. \tag{4b}$$

In terms of the mean free path l and the thermal velocity v_T, we can define a *mean free time* τ;

$$\tau = l/v_T \doteq 10^{-12} \text{ sec}.$$

This is the average time during which an electron moves between collisions.

One of the convenient units of energy to use for describing electrons is the electron volt. This is the energy which an electron would acquire in falling through a potential difference of one volt. At room temperature T is about 300° on the Kelvin scale and

$$kT = 0.026 \text{ electron volts}. \tag{5}$$

We shall see below that the voltage which gives an electron thermal energy plays a critical role in the nonlinear behavior of p-n junctions and p-n junction transistors.

The model of electron collisions with phantom particles represents quite accurately the behavior of an electron interacting with the phonons. In particular it correctly represents the ability of electrons to interchange energy with the phonons. This interchange of energy is a very important process which occurs whenever an electrical field is applied to a semiconductor. Under these conditions electrons acquire energy from the electric field which does work upon them. In order for this energy to be converted to heat in the crystal it is necessary for it to be transferred to the thermal vibrations or phonons. From our model of the phantom particles it is evident that if the electrons acquire a greater velocity

[8] E. M. Conwell, "Properties of silicon and germanium," PROC. I.R.E., vol. 40, pp. 1327–1337; this issue.
[9] W. E. Kock, "Metallic delay lenses," *Bell Sys. Tech. Jour.*, vol. 27, pp. 58–82; 1948.
W. E. Kock and F. K. Harvey, "Refracting sound waves," *Jour. Acous. Soc. Amer.*, vol. 21, pp. 471–481; 1949.
[10] W. Shockley, "Hot electrons in Germanium and Ohm's law," *Bell Sys. Tech. Jour.*, vol. 30, p. 990; 1951.

than their normal thermal velocity, then upon collisions with the phantom particles they will lose their excess energy. A mathematical analysis of this process can be carried out in order to find out at what critical field Ohm's Law will fail. We shall return to this point briefly after discussing the mobility of the electrons below.

Diffusion: The behavior of an electron in the absence of an electric field is represented crudely in Fig. 5. Its path consists of a number of short straight line segments. On the average, each one of the segments has a length of l. At the end of each segment the electron has a collision

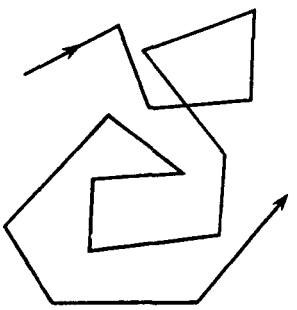


Fig. 5—The random diffusive motion of an excess electron.

with a phantom particle and starts off with approximately the same thermal velocity but in a completely random direction. As a result, it *diffuses* about in the crystal. The process of diffusion is a basic one in transistor electronics and deserves special discussion at this point. It is difficult, however, to explain the process of diffusion in terms of the motion of an individual electron. For this reason we shall consider a large number of electrons so that we can deal with their average statistical behavior. If a large number of electrons were actually present in a crystal which contained no other electronic or chemical imperfections, then they would strongly influence each other by their electrostatic repulsion. There are certain experimental conditions, which we shall discuss in Section 4, in which the effect of this electrostatic repulsion is suppressed by other factors. It would greatly prolong the argument to consider these other factors here. In order to proceed with the discussion of diffusion, therefore, we shall simply assume for the moment that the electrostatic interaction between the electrons is so small that we can neglect it in comparison with the diffusive motion.

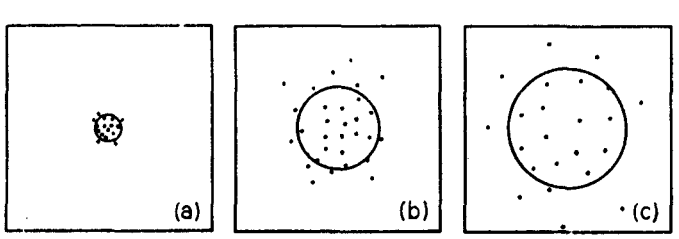


Fig. 6—The spreading of a group of electrons by diffusion.

The quantity which characterizes diffusion is the *diffusion constant* D. It has the dimensions of cm²/sec. We shall illustrate with two examples how diffusion processes are governed by the diffusion coefficient and shall also show how the diffusion coefficient is related to the microscopic quantities l and v_T. The first example is represented in Fig. 6. We imagine that a group of N electrons are all contained inside a very small sphere at $t=0$. Then at later times, the random diffusive motion causes them to spread out. Their distribution in space at time t is then given by the formula, which we introduce just for this example,

$$n(r, t) = N(\pi Dt)^{-3/2} \exp(-r^2/4Dt), \quad (6)$$

where r is the distance from the initial position of the electrons and $n(r, t)$ is the number of electrons per cm³ at r at time t. The most important term is the exponential. It shows that at any instant of time the density of the electrons is a *three-dimensional error function.* The density is a maximum at the origin, and decreases by a factor of $\exp(-1)=0.37$ at $r=2(Dt)^{1/2}$ and by a factor of $\exp(-4)=0.019$ at twice that distance. In effect, the electrons are contained in a sphere of radius slightly larger than $2(Dt)^{1/2}$. One way of describing the size of the distribution is to consider the average value of r^2 for the electrons. This is found to be

$$\text{average of } r^2 = 6Dt. \quad (7)$$

The second example is shown in Fig. 7. In this case the electron density is independent of the y and z coordinates and increases with increasing x. The rate of increase is called the gradient of the electron density

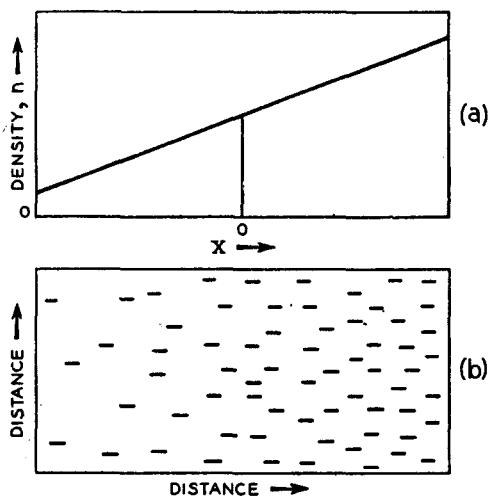


Fig. 7—A distribution of electrons with a density gradient.

and is the slope of the line in Fig. 7(a). Part (b) suggests the nature of the corresponding distribution of electrons in space. As these electrons diffuse, they tend to smooth out their distribution, with the result that a diffusion current of electrons will flow to the left in the figure. The strength of this current is

$$\text{electrons per unit area per unit time} = D \text{ times density gradient.} \quad (8)$$

We shall discuss below the way in which D may be derived from l and v_T on the basis of the situation represented in Fig. 7. There is, of course, only one diffusion constant, and it can be shown that either of the two equations (7) and (8) may be derived from the other.

In order to deal with the situation of Fig. 6 from a microscopic viewpoint we shall imagine that a group of N electrons are contained within a small sphere at time $t=0$. As a result of thermal velocities and collisions they spread out in a random fashion. In order to illustrate how this spreading results in (7) we shall consider what happens in one mean-free time. During this time each electron travels one mean-free path or, more accurately, parts of several mean-free paths. For purposes of this exposition, however, no important error will be made by assuming that all the electrons start and conclude a mean-free path simultaneously. Now consider the group of electrons whose radial distance from the origin is r when they start a mean-free path. After a time τ, about half of these will have moved outward and will be approximately at radius $r+l$, and half will have moved inward to $r-l$. The new average of r^2 for these electrons will thus be

$$\text{new av of } r^2 = (1/2)[(r+l)^2 + (r-l)^2] = r^2 + l^2. \quad (9)$$

This result is somewhat in error because the outward motion is not necessarily radial; when this effect is taken into account, it is found that the increase in average of r^2 for this group of electrons in time τ is

$$\text{increase in average of } r^2 \text{ is } l^2/3.$$

Since this result is independent of r, it will be obtained for all groups of electrons. In a time t, such increases will occur t/τ times so that

$$\text{increase in average } r^2 \text{ in time } t = l^2t/3\tau = (v_T l/3)t. \quad (10)$$

Thus we see, on the basis of this very crude analysis, that r^2 will increase linearly in time and that the coefficient is given in terms of the microscopic quantities l and v_T.

The spreading of a group of electrons is one manifestation of diffusion of interest in transistor electronics. Another is the diffusion current produced by a concentration gradient. For this purpose we suppose that the electrons are distributed nonuniformly, as in Fig. 7, so that the number per unit volume is n and

$$n = n_0 + ax. \quad (11)$$

The concentration gradient is then

$$dn/dx = a. \quad (12)$$

Since there are more electrons to the right than to the left of the plane at $x=0$, we would expect a diffusion current to flow to the left. The magnitude of this diffusion current may be estimated roughly as follows: In one mean-free time, half the electrons in the region from $x=0$ to $x=+l$ will cross the $x=0$ plane towards the left and half those in the region from $x=0$ to $x=-l$ will cross it to the right. Now half the number of electrons per unit area of the $x=0$ plane within a distance l to the right of $x=0$ is

$$(1/2)l \text{ times } (n_0 + al/2) \quad (13)$$

since the average density in this region is the density at $x=l/2$. A similar treatment for the electrons to the left leads to an expression with $-al/2$ in place of $al/2$. From this it follows that the net unbalance of flow causes

$$(l/2)(n_0 + al/2) - (l/2)(n_0 - al/2) = al^2/2 \quad (14)$$

electrons to cross the plane at $x=0$ moving towards the left in time τ. Hence the electron flow is

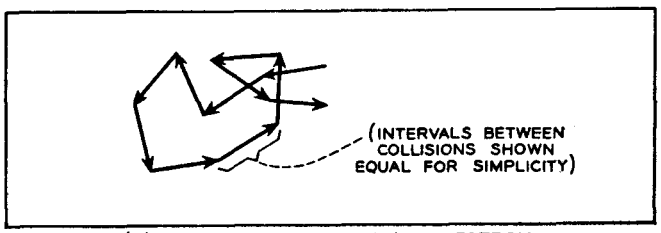
$$\text{electrons per unit area per unit time} = al^2/2\tau = (lv_T/2)a. \quad (15)$$

If correct averaging procedures are employed, it is found that the coefficient of a becomes identical with D found in connection with the increase in r^2. Hence the general result is

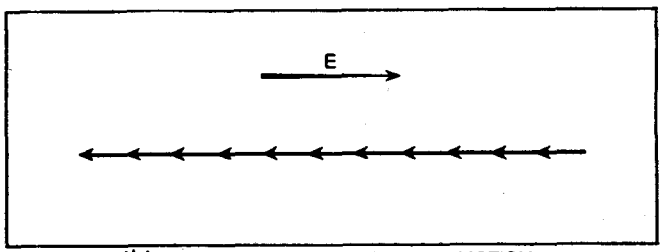
$$\text{particles per unit area per unit time} = D \text{ times density gradient.} \quad (16)$$

The direction of flow is in the direction of lower concentration.

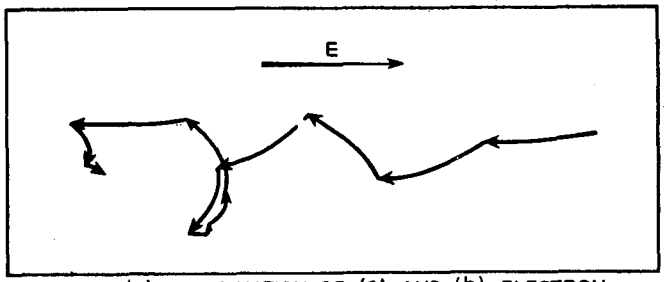
Mobility: The effect of an electric field upon a diffusing electron is indicated in Fig. 8. The electric field



(a) RANDOM MOTION OF AN ELECTRON IN A CRYSTAL



(b) COMPONENT OF ELECTRON MOTION ADDED BY ELECTRIC FIELD



(c) COMBINATION OF (a) AND (b) ELECTRON MOTION IN AN ELECTRIC FIELD

Fig. 8—The effect of an electric field in superimposing drift on random motion.

which is shown as being directed towards the right in the figure exerts a force to the left upon the electron because of its negative charge. This force accelerates the electron with the result that the segments of the electron's path between collisions become curved rather than straight lines. The end points of the paths are consequently offset, as is indicated in part (c) of the figure. The offset produced by the electric field in each mean-free path can

be calculated from the falling-body formula. This formula is usually written in the form

$$S = (1/2)at^2. \tag{17}$$

The use of this equation to calculate the effect of the electric field makes use of the fact that the effect of an acceleration can be superimposed upon an already established velocity. (This is the theorem used in answering the question, "Where should you aim if the monkey drops when he sees the flash?") After a collision, the electron starts with a perfectly random velocity which will carry it one mean-free path in a time τ. During its motion it acquires additional velocity because of the force, and the additional motion is given by (17).

In order to apply (17) to our case we must take the acceleration as being equal to the force divided by the mass and must use the expression qE for the force where q represents the magnitude of the charge on an electron. This leads to the substitutions

$$a = F/m = qE/m \tag{18}$$

$$t = \tau. \tag{19}$$

The cumulative effect of the displacements given by equation (17) is represented in part (b) of Fig. 8. This results in a steady drift being superimposed on the random motion. Since one displacement S occurs in each interval τ of time, the average *drift velocity* is evidently given by the formula

$$v_d = S/\tau = (qE/m)\tau/2 = (q\tau/2m)E. \tag{20}$$

If a more sophisticated method of averaging is used which takes into account the fact that not all mean-free paths have exactly the same length, then it is found that a better value for the drift velocity is given by

$$v_d = (q\tau/m)E = \mu E. \tag{21}$$

In the last term of this equation we have introduced the symbol μ which is known as the *mobility*. It is the proportionality factor between electric field and drift velocity:

$$\mu = v_d/E = q\tau/m. \tag{22}$$

The linear relationship between electric field and drift velocity is valid so long as the electrons are able to dissipate the energy they acquire from the field sufficiently rapidly to the phonons.

If the electric field is too high, the electrons become "hot" in the sense that the energy of their random motion is increased. As a result they have collisions more frequently and their mobility is decreased. This effect gives rise to a fundamental deviation from Ohm's law in certain experiments and the resistance of a specimen increases with increasing electric field.[8,11] The critical field at which the mobility deviates appreciably from its low electric field value is simply related to the behavior of the phantom particles that simulate the phonons. When the drift velocity of the electrons equals the thermal velocity of the phantom particles, then the linear relationship between drift velocity and electric field fails.[10]

Since both the diffusion constants and the mobility are determined by the microscopic quantities l, v_T and m, there should be a relationship between the two. This relationship can be exhibited in terms of the ratio

$$D/\mu = (1/3)lv_T/(q\tau/m) = mv_T^2/3q. \tag{23}$$

This equation is inaccurate because of the approximations made in evaluating D and μ. The correct equation is

$$D/\mu = mv_T^2/2q = kT/q = 0.026 \text{ volt at } 300^\circ K. \tag{24}$$

This relationship between diffusion constant and mobility is known as the Einstein relationship. It was derived by Einstein on the basis of much more general considerations which show that such a relationship must hold regardless of the nature of the mechanism involved in producing diffusion and mobility. In semiconductor work it has usually proved more effective and accurate to measure the mobility and then to use the Einstein relationship to evaluate the diffusion constant rather than to measure the diffusion constant directly. Recently, however, the Einstein relationship has been directly verified by measuring simultaneously mobility and diffusion of both electrons and holes in germanium. We shall discuss the experiments involved in Section 4.

The orders of magnitude of the diffusion constant and the mobility can be estimated from the values given earlier in this section. Evidently, the diffusion constant will be approximately

$$D = (1/3)lv_T = 33 \text{ cm}^2/\text{sec}. \tag{25}$$

The mobility may then be estimated from the diffusion constant with the aid of the Einstein relationship, and the result is

$$\mu = (q/kT)D = 1{,}300 \text{ cm}^2/\text{volt sec}$$
$$= 1{,}300 \text{ (cm/sec) per (volt/cm)}. \tag{26}$$

These values are typical of mobilities for electrons and holes in silicon and germanium. The actual values may differ by 3 or 4 from these and are of course functions of the temperature. The best estimates for values for germanium and silicon are given in Table II. The effects of impurities and temperature upon μ and D are reviewed in a recent article by Conwell.[8]

TABLE II

MOBILITIES IN CM²/VOLT SEC AND DIFFUSION CONSTANTS IN CM²/SEC

	Electrons		Holes	
	μ	D	μ	D
Si	1,200±120	30±3	250±50	65±12
Ge	3,600±180	93±5	1,700±90	43±2

The combined effects of drift and diffusion are shown in Fig. 9. Here it is supposed that the group of electrons start at $t=0$ at one point indicated by the dot. Their subsequent distributions are represented by the circles.

[11] E. J. Ryder and W. Shockley, "Mobilities of electrons in high electric fields," *Phys. Rev.*, vol. 81, p. 139; 1951.

The times since starting for the five circles are in the ratio 1:9:25:49:81. It is seen that the effect of diffusion is more powerful than the effect of drift at short times, and during the first interval of time, some electrons diffuse against the field. As the circle surrounding the electrons becomes larger, the rate at which it grows radially decreases and the dominant effect is that of drift in the electric field. If the circle is taken as the radius at which the electron density is 0.1 of its value at the center, then (7) leads to

$$R = (2.3 \times 4 \times Dt)^{1/2}, \tag{27a}$$

where t is the time of flight. The distance drifted in the same time is

$$L = \mu Et = \mu Vt/L, \tag{27b}$$

where V is the voltage drop in distance L. These equations lead to

$$\text{spread/drift} = R/L = (9.2kT/qV)^{1/2}. \tag{27c}$$

From this it is seen that the spreading is small compared to the drift when V is large compared to $9.2\ kT/q$ or about $\frac{1}{4}$ volt. The largest circle in Fig. 9 corresponds to about 4 volts.

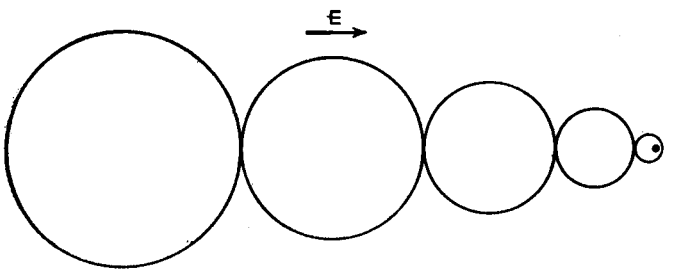


Fig. 9—The relative influence of diffusion and drift.

A magnetic field applied to the specimen will lead to an additional deflection of the electron paths and will produce a phenomenon known as the Hall effect. Although the Hall effect has proved to be of considerable interest in connection with semiconductors and transistor electronics, in general it does not appear to be of sufficient immediate concern with the functioning of transistors to warrant a complete discussion of it at this place. Consequently, we refer the reader to the literature for a discussion of the Hall and Suhl effects.[7]

C. The Behavior of a Hole

The word *hole* is used to describe the electronic imperfection produced by removing one electron from a valence bond. Such a disturbance obviously represents one positive charge equal to the electronic charge. Like the excess electron this charge will be shielded by the dielectric constant of the material. When the perfection of the crystal is disturbed in this way, electronic conduction can be thought of as taking place by a replacement process. An electron in an adjacent bond can jump into the hole in the incomplete bond, thus producing an electronic motion and a reciprocal motion of the hole. If an electric field is applied such as to move electrons to the left, the hole will move towards the right as a result.

On the basis of this picture, what would we expect the attributes of a hole to be? In the first place, its charge as stated above would be plus-one electronic unit. We would expect the mean-free path to be about one interatomic distance. It is hard to tell what effective mass or mean-free time should be assigned to the hole on the basis of the description given above. The results drawn from this simple reasoning are in disagreement, however, both with theory and with direct observation of the behavior of holes. The conclusions drawn from more complete reasoning and from experiment are at first surprising. When the theory is worked out in detail, it is found that the application of the wave equation to the behavior of the electrons when a hole is present leads to the conclusion that the effective mean-free path for hole motions is of the same order of magnitude as for electrons. The effective mass for the hole is also found to be of the same order of magnitude.[12] Qualitatively then, the hole behaves just as does an electron except for the fact that it acts as though it were a positive charge. Quantitatively, D and μ are somewhat less for a hole than for an electron.

There is no simple way of seeing how the electron replacement process can lead to these long mean-free paths for holes. The analytical reasoning required to reach this conclusion inevitably seems to be complicated.

From the experimental point of view, on the other hand, the behavior of the hole may be regarded as an established fact. The mobility and diffusion constants for holes in germanium have been directly measured, as have those for electrons. It is found that the hole is approximately $\frac{1}{2}$ as mobile as an electron in this case. If the mean-free path were really as short as one interatomic distance, the ratio would have to be 1:1,000 instead of 1:2. Thus the important attributes of a hole may be regarded as determined by direct experimental observations. So we are justified in using these attributes of the behavior of holes in design theory and in the explanation of the way in which transistor devices function.

Although the hole has acquired a very substantial reality as a result of new experiments in transistor physics, its true nature should not be forgotten. The hole is after all simply a convenient way of describing the behavior of an incomplete assemblage of electrons. Attributing to the hole a positive mass, a positive charge, and a mean free path of about 10^{-5} cm leads to a correct description of the way in which this imperfection in the incomplete assemblage diffuses and drifts under the influence of electric and magnetic fields. These are the processes of chief importance in transistor electronics, and for them no error will be made in considering the hole to be a real particle. There are pitfalls, however, in the blind acceptance of this model, and there are circumstances in which the true

[12] An extensive treatment of this topic is given in footnote reference 1.