CLARKSON COLLEGE OF TECHNOLOGY

DEPARTMENT OF ELECTRICAL ENGINEERING

A STUDY OF THE ELECTRONIC PROCESSES IN EXTRINSIC GERMANIUM AS EXHIBITED BY THE HALL AND MAGNETORESISTANCE EFFECTS

A SENIOR THESIS

by

PAUL M. GRANT

Submitted in partial fulfillment of the requirements

for the degree of

Bachelor of Electrical Engineering

January 20, 1960

Approved by Thesis Advisor — Date

Albert R. Martin — 26 Jan '60

George W. Reed — 26 Jan '60

Department Chairman

A STUDY OF THE ELECTRONIC PROCESSES IN EXTRINSIC GERMANIUM AS EXHIBITED BY THE HALL AND MAGNETORESISTANCE EFFECTS

A SENIOR THESIS

by

PAUL M. GRANT

TABLE OF CONTENTS

	Page
Biographical Sketch	2
Acknowledgements	4
Abstract	5
Introduction	6
The Hall and Magnetoresistance Effects	
I. Theory	10
II. Experiment	25
Appendix	
A. A Magnetoresistive Control Device	56
B. Non-destructive Sensing of Magnetic Storage in Digital Computers	72
C. Motion of an Electron in a One Dimensional Periodic Potential	78
D. Effective Mass and the Existence of Holes	91
E. Photographs	98
References	103

BIOGRAPHICAL SKETCH OF PAUL M. GRANT

The author was born on May 9, 1935 in Poughkeepsie, New York. He attended elementary school in Poughkeepsie and high school at the Oakwood School in Poughkeepsie and Wappingers Central School in Wappingers Falls, New York. Upon graduation in June, 1953, he entered the employ of the International Business Machines Corporation of Poughkeepsie, New York, as a mail clerk in the engineering laboratory. Because of a near life-long interest in amateur radio and skills learned in the practice thereof, he was shortly promoted to electronic technician. Some time later he attended the IBM Field Engineering School and upon graduation was assigned as a Field Engineer to the SAGE computer installation at the MIT Lincoln Laboratory, Lexington, Massachusetts. During his assignment at Lincoln Laboratory, Mr. Grant was promoted to Technical Specialist in programming.

The author was granted a leave of absence by IBM in September, 1956, in order that he might commence his matriculation as an Electrical Engineering Major at Clarkson College of Technology, Potsdam, New York. During the summer of his first year, he was located with IBM at Kingston, New York, performing the work of writing reliability and diagnostic programs for the maintenance of the SAGE computer. At the end of his second college year, the author was assigned to the IBM Kingston engineering laboratory where he participated in the development of thin film logical devices. In the summer of 1959,

Mr. Grant was involved in research on the electroluminescence phenomenon in thin films, working under the direction of Dr. Roland C. M. Beeh.

Mr. Grant is married to the former Miss Joan Geddes of Lexington, Massachusetts, and resides with his wife and son at 77½ Main Street, Potsdam, New York. The author is a member of the Institute of Radio Engineers, Eta Kappa Nu, and Tau Beta Pi.

ACKNOWLEDGEMENTS

The author wishes to take this opportunity to express his gratitude for the aid given him by the following people: Mr. Albert R. Martin, for his accepting the position of Thesis Advisor and for the guidance and criticism of this work throughout its performance; to Dr. Ralph G. Gentile, for the procuring of the magnetic circuit; to Dr. Robert C. Vernon, for many helpful discussions on the theory of the solid state and the Hall effect; to the International Business Machines Corporation, for supplying the semiconductor samples; and to the many associates, both academic and professional, who have enlightened and encouraged the author in his endeavors over the past few years. Finally, much of the credit for this paper must be given to the author's wife, Mrs. Joan R. Grant, without whose help the preparation of this report would have been impossible.

ABSTRACT

A study is made of the electronic processes in semiconductors as exhibited by the Hall and magnetoresistance effects. Measurements are made on the two carrier types of extrinsic germanium and the Hall coefficient and resistivity as a function of the power dissipated in the samples calculated. The hole and electron mobilities of each type are computed for room temperature conditions and the resultant values compared against those found in the current literature. Reasons for discrepancies are studied. A discussion of the theory of the Hall and magnetoresistance effects, along with the experimental apparatus, including photographs, is given. An analysis of two possible practical applications of the Hall and magnetoresistance effects and a discourse on some of the basic concepts of band theory are to be found in the appendices.

INTRODUCTION

The Hall and magnetoresistance effects have been well known for some time. The Hall effect was discovered by Professor Edwin Herbert Hall of John Hopkins University in 1879, whose name the phenomenon bears. Some of the earliest work on the magnetoresistance effect was done by Corbino after the turn of the century, and the geometry that produces the optimum measurements is named after him. The Hall effect, which is the production of a voltage at right angles to the plane formed by the directions of an electric current and a magnetic field, was first noticed and measured in thin sheets of the malleable metals -- gold,silver, and copper. It is not known when the magnetoresistance effect first came to light, but it has always been present to some degree in electrical machinery and it is probable that is was encountered early in the development of this art, as it involves the change of resistance of a conductor when the conductor is subjected to high magnetic fields at right angles to the direction of current flow.

Since both these effects are usually quite small, they remained scientific curiosities for many years. Some of the earliest patents involving the magnetoresistance effect are held by the Bell System and were awarded in the years 1900 to 1920. These patents pertained mostly to current modulation devices and, to the knowledge of the author, there is no indication that they were ever used in practice. One of the

first applications of the Hall effect was to the obvious task of measuring static magnetic fields and to this day a superior method has yet to be found. However, early attempts to use the two effects were usually crude and resulted in no accomplishments of any major technical importance.

With the advent of wave mechanics in the early 1930's, the Hall effect was explained satisfactorily for the first time. There had been certain materials which had given the "wrong" sign for the Hall voltage and the theoretical finding that, under certain conditions, the reversal of Hall potential could be explained is one of the very fruitful results of wave mechanics. Ever since this time, the Hall effect has risen steadily in importance. It has become a very effective tool in the study of semiconductor materials and provides the best means of determining the conduction processes in extrinsic semiconductors. It can also lead to a knowledge of the energy band structure in semiconducting compounds and to a knowledge of the electron scattering mechanism in the crystal. On the practical side, the development of some high resistivity, high mobility compounds has led to materials which produce a pronounced Hall effect. This fact, coupled with an increased use of static magnetic fields, presents some interesting possibilities.[1,2,3,4] For example, a Hall effect device can be used to perform the arithmetic function of multiplication in an analog computer. Also, this multiplication feature is made use of in a device which measures indirectly the torque of a DC machine. An application can be envisioned in which a time-

varying magnetic field is used to modulate a sinusoidal current of a different frequency. The Hall effect may eventually produce a means of non-destructively sensing the magnetic memory of a digital computer, as outlined in Appendix B. The old role of field measurement is making the Hall effect very useful in the study of some superconducting phenomena.[2]

The discovery of a new class of semiconducting materials, the intermetallic compounds, has given rise to a new interest in the magnetoresistance effect. Welker and Weiss[1] report that experiments on InSb and InAs have yielded mobilities five to ten times greater than germanium with resultant changes in resistivity of five to twenty times the normal specific resistivity. Such changes in resistance hold new hope for the application of the effect in an electronic amplifying device. The theory of such a device is presented in Appendix A and a model discussed which can be constructed through the art of thin film technology.

Because of the attention being received by the Hall and magnetoresistance effects today, this study was undertaken by the author primarily for his own education. In the preparation of this paper, he has been brought into contact with some of the more advanced concepts of modern physics and has in places attempted to clarify these concepts where they apply to the Hall and magnetoresistance effects. He has tried to include a rather detailed discussion of some of the aspects of band theory which at first baffled him and some others which still do. In doing so, the author feels that these attempts at

clarification of theory, as well as the performance of the experimental work, have helped him understand the electronic processes in semiconductors much better than he did previously.

I. THEORY

A. Introduction

The purpose of this section is to outline the Theory underlying the Hall and magnetoresistance effects. Because some understanding of modern band theory is necessary for the successful explanation of these effects, this topic is discussed briefly herein and again at length in appendices C and D. A simple vector model of each effect is used in order to derive the usual equations expressing the Hall and magnetoresistance phenomena. The MKS system of units is implied in all derivations and equations in this section; however, the experimental results are presented in the mixed MKS-CGS system found in most of the current literature. Transformation of any equation stated in the Theory to an equation in terms of the experimental results entails merely the application of certain well-known conversion factors.

B. Energy Bands in a Solid

When discussing particle or wave motion in continuous media, one is used to thinking of energy and momentum as being related to each other by continuous functions. The classical Hamiltonian form for such motion is expressed by

$$(1\text{-}1) \quad E = \frac{|\bar{p}|^2}{2m} + V$$

where E is the total energy, m the mass of the particle, $\bar{p}$ its momentum, and V its potential energy. It can be quickly seen that a simple parabolic relationship exists between the kinetic energy and the momentum, a relationship which is defined over all values of the variables. Such is not the case for electronic motion in the periodic potential encountered in crystalline solids. This is a most important fact because it is due to discontinuities in the energy - momentum relationship that all solid state phenomena occur.

These discontinuities arise from three very fundamental equations of modern physics, the first of which is the de Broglie hypothesis:

$$(1\text{-}2) \quad \lambda = \frac{h}{p}, \text{ or, } \bar{k} = \frac{\bar{p}}{\hbar}$$

where h is Planck's constant and $\hbar$ is Planck's constant divided by 2π. Equation (1-2) postulates that there exists a wavelength λ associated with the momentum p of any body. It is also convenient to define the "wave vector" k which is inversely proportional to the wavelength and directly proportional to the momentum. Equation (1-2) has been shown to be true through

the observance of electron diffraction patterns in crystals. Hence, if a particle has wave properties, then it would be reasonable to assume that there might exist a wave equation describing its motion. Indeed, this wave equation does exist and is known as the Schroedinger Equation. The Schroedinger Equation has many different forms; however, since most of the qualitative discussion of semiconductor phenomena is done using a one-dimensional model, it will suit our purposes to display only the non-relativistic, time independent, one-dimensional Schroedinger Equation:

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}\left[E - V(x)\right]\psi = 0 \tag{1-3}$$

where m is now the electronic mass. The quantity ψ can be given no valid physical meaning except that its magnitude squared expresses the probability distribution function for the position of the electron. E is again the total energy of the electron and V(x) describes the potential energy field throughout the lattice. If the potential energy field is periodic; that is, V(x) = V(x+a), where "a" is the lattice constant, then equation (1-3) has a general solution given by the Bloch Theorem.

$$\psi = e^{\pm ikx}\mu(x), \quad \mu(x) = \mu(x+a) \tag{1-4}$$

Equation (1-4) can be derived from another theorem in the theory of differential equations called Floquet's Theorem.

It can now be seen that the combination of equations (1-3) and (1-4) might yield a relationship between E and k or E and p. This relationship and the fact that it is discontinuous is derived in Appendix C.

The regions in momentum space, or k space, where the energy is continuous are called Brillouin zones. At the boundaries of the Brillouin zones, the function has what is mathematically termed a simple discontinuity in energy. This means that certain values of energy are forbidden to the electron during its motion throughout the crystal lattice. The result is the formation of the well-known energy bands within the solid. In this paper, we will be concerned only with the valence and conduction bands which are pictorially represented for an intrinsic (pure) semiconductor in Fig. 1-1(a). The meaning of the dotted line through the middle of the forbidden gap, called the Fermi level, is that at a temperature of absolute zero, electrons cannot occupy any allowed states above this level. At room temperature, however, there is a finite probability that electrons may occupy energy states beyond the Fermi level. Since the only allowed states over the Fermi level are in the conduction band, at room temperature there will be a finite number of electrons in an intrinsic semiconductor that will be available for conduction. The electrons which now are in the conduction band must have originated in the valence band, thus for every conduction electron, there is a vacancy or "hole" created in the valence band. Before continuing further, it will be convenient to

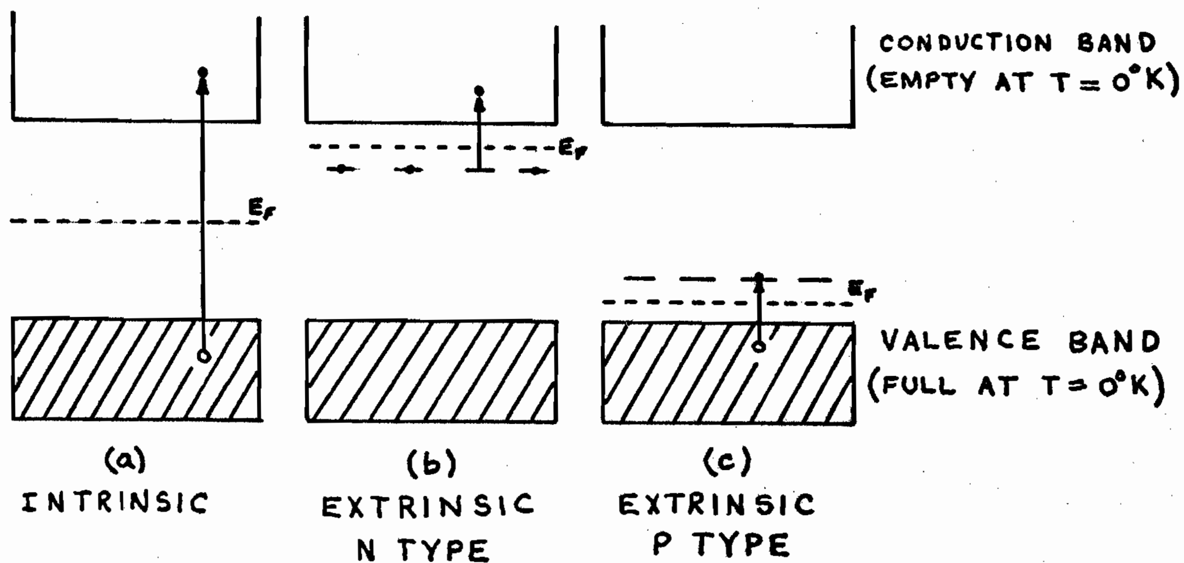
define, in conjuction with the hole concept, the effective mass, m^*, or "freeness" of an electron moving in an energy band.

$$(1\text{-}5) \quad m^* \equiv \frac{\hbar^2}{d^2E/dk^2}$$

One can see that the effective mass of an electron depends upon its location in the band. In Appendix D, equation (1-5) is derived and it is shown that the effective mass of an electron near the top of an energy band is negative or that it is a hole with positive mass and charge.

Let us now discuss the band structure of extrinsic, or doped, semiconductors. If one introduces minute amounts of group III or group V elements into a group IV semiconductor, such as germanium, the bond requirements of the chemical substitution are not satisfied, resulting in the creation of coulombic potential regions which spread over distances of several lattice constants from the impurity center. The solution of equation (1-3) in these regions leads, in the first approximation, to energy levels similar to those in the hydrogen atom. Thus, there are created in the forbidden gap certain allowed energy levels. The closeness of these levels to the valence band or the conduction band depends upon the type impurity substituted. If a group III element is used, the result will be an unoccupied level close to the valence band. Such a material is called P type because transitions may take place from the valence band to the unoccupied levels, thereby creating holes in the valence band. If the impurity is a group V element, the result will be occupied levels just

beneath the conduction band. This compound is called N type as transitions are possible from the occupied levels to the conduction band which produce conduction electrons. In neither case do transitions initiate carriers of the opposite type, as would be true in an intrinsic semiconductor. The energy band structure for intrinsic and extrinsic germanium, along with typical energy level and forbidden gap values, is shown in Fig. 1-1.



ENERGY GAP BETWEEN VALENCE AND CONDUCTION BANDS: .785 ev

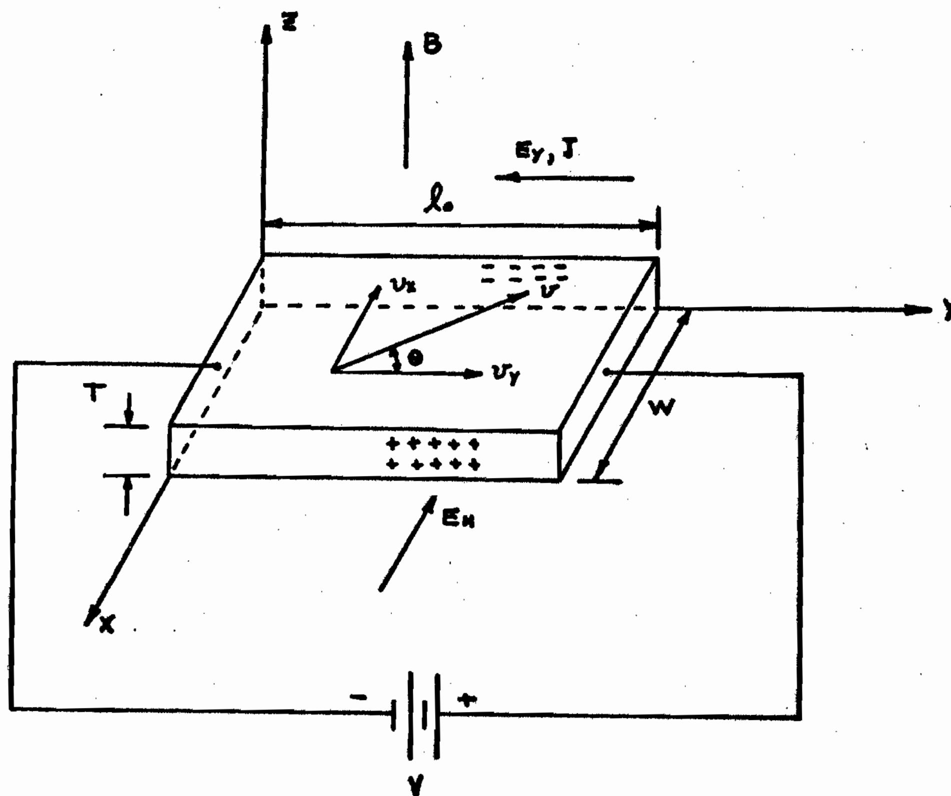
ENERGY OF IMPURITY LEVELS WITH RESPECT TO NEAREST BAND: .01 - .05 ev

ENERGY LEVEL STRUCTURE OF GERMANIUM

Fig. 1-1

C. Theory of the Hall Effect and Magnetoresistance

The Hall effect and magnetoresistance effect were largely unexplained until the advent of wave mechanics and the modern band theory of solids. However, it is best to consider these two effects from a classical standpoint first and then refine the concept qualitatively along the lines of band theory. We shall therefore take the vector model depicted in Fig. 1-2 as our starting point.



ELECTRONIC MOTION IN AN N TYPE SEMICONDUCTOR IN A MAGNETIC FIELD

Fig. 1-2

We shall consider electron motion in an N type semiconductor in

the directions shown by the velocity vectors. These vectors represent the direction of carrier flow upon application of the magnetic field, B. After a certain time has passed, an electric field, E_H, will be built up orthogonal to the applied field, E_y, and the direction of the average electron velocity given by Boltzmann statistics will be essentially in the y direction. Under these conditions, the Lorentz equation is,

$$(1\text{-}6) \quad \bar{F}_x = -e\left(-\bar{E}_H + \bar{v}_y \times \bar{B}\right) = 0$$

where $\bar{F}x$ is the net force on the electron, $\bar{v}_y$ is the average velocity of the electrons, and e is the electronic charge. Solving for $\bar{E}_H$,

$$(1\text{-}7) \quad \bar{E}_H = \bar{v}_y \times \bar{B}$$

The current density, $\bar{J}$, in the y direction is given by the well-known expression,

$$(1\text{-}8) \quad \bar{J} = \sigma \bar{E}_y$$

where σ is the conductivity tensor. We shall define the Hall mobility by the equation,

$$(1\text{-}9) \quad \bar{v}_y = \mu_H \bar{E}_y$$

Combining (1-9), (1-8), and (1-7), we conclude that

$$(1\text{-}10) \quad \bar{E}_H = \frac{\mu_H}{\sigma}\left(\bar{J} \times \bar{B}\right), \quad \text{with } R_H \equiv \frac{\mu_H}{\sigma}$$

R_H being defined as the Hall coefficient. The reason for defining a Hall mobility different from the normally encountered mobility μ, is that the carrier scattering process is altered in the presence of a magnetic field. It can be shown that,[5]

(1-11) $$\mu_H = \frac{3\pi}{8}\mu = \frac{3\pi/8\ \sigma}{ne}, \text{ from } \sigma = ne\mu$$

where n is the number of conduction electrons per unit volume. Thus (1-10) becomes

(1-12) $$\bar{E}_H = \frac{3\pi/8}{ne}(\bar{J}\times\bar{B}), \quad R_H = \frac{3\pi/8}{ne}$$

The following differential equation for the Hall voltage can be deduced from equation (1-12) and the definition of the electric intensity vector, $E_H = -\text{grad}\, V_H$;

(1-12a) $$\nabla V_H + \frac{3\pi/8}{ne}(\bar{J}\times\bar{B}) = 0$$

Let us now make some reasonable assumptions concerning our model shown in Fig. 1-2.

(1) The vectors $\bar{J}$ and $\bar{B}$ are at right angles with each other and the Hall voltage equipotential lines are essentially parallel in the region in which the pick-up probes are placed.

(2) Both $\bar{J}$ and $\bar{B}$ are constant.

(3) The sample has uniform thickness.

Accordingly, then, equation (1-12a) becomes:

(1-12b) $$\frac{dV_H}{dx} = -\frac{3\pi/8}{n e} J B = -R_H B \frac{dI_s}{dA} = -R_H B \frac{dI_s}{T\, dx}$$

where Is is the sample current, A is the sample cross-sectional area, and T is the sample thickness. Upon integrating (1-12b) and rearranging terms, we finally arrive at the Hall effect equation from which the Hall coefficient R_H can be determined experimentally.

(1-13) $$R_H = -\frac{V_H T}{I_s B}$$

Equation (1-13) indicates the proper sign for the Hall voltage and coefficient for N type semiconductors. The sign of the Hall voltage can also be found through the application of the right-hand rule. The first workers in the field found that some materials gave an anomalous behavior; that is, the sign of the Hall voltage was opposite to that expected if the current carriers were assumed to be electrons. For a long time this phenomenon was unexplained, and it was not until the coming of band theory, with its hole concept, that a plausible explanation was possible. The Hall effect is the strongest direct proof of the existence of two types of current carriers in semiconductor materials. It should be noted that the simple - minded theory of hole motion being merely electron motion in the opposite direction will not lead to a bi-polar Hall effect. This can be readily seen upon a moment's

contemplation of Fig. 1-2. It is for this reason that a knowledge of band theory is absolutely essential for even an elementary understanding of the Hall effect in semiconducting compounds. Holes have real existence in the energy bands and should be considered as entities entirely separate from electrons. As is shown in Appendix D, holes have mass and exhibit all the electrodynamic properties associated with a positively charged particle. Previously, it has been pointed out that an intrinsic semiconductor possesses equal concentrations of holes and electrons at thermal equilibrium. One would therefore expect the Hall effect in such a material to be quite small. Such is, in fact, the case; however, an intrinsic semiconductor usually exhibits a slight net Hall voltage of the N type. The reason is that for intrinsic materials, the conductivity for one type of carrier as expressed by

$$\sigma = ne\mu \tag{1-14}$$

must be modified. The conductivity is now due to both holes and electrons, so that we can write,

$$\sigma = (n_d \mu_d + n_h \mu_h) e \tag{1-15}$$

where the subscript d denotes donors or electrons, and the subscript h denotes holes. As the electron mobility is usually greater than the hole mobility, the resulting Hall voltage sign is of the N type polarity. In extrinsic

semiconductors, it has been shown that the thermal production of majority carriers does not simultaneously lead to the production of an equal number of minority carriers. Therefore, conduction in extrinsic materials is dominated by one type of carrier and the effect of thermal hole-electron pair creation can usually be neglected at normal operating temperatures. However, in P type specimens, one must be careful as the sign of the Hall voltage will reverse when the temperature increases to the point where hole-electron pair formation becomes appreciable and the electrons, due to their higher mobility, overcome the effect of the hole majority carriers. An analogous situation would not, of course, arise in N type extrinsic semiconductors.

The transverse magnetoresistance effect is the apparent change in the specific resistivity of a conductor when it is placed in an orthogonal magnetic field. The effect is quite pronounced in certain semiconductors and often occurs along with the Hall effect, although, as we shall see, the best magnetoresistance effect occurs under conditions least favorable to the Hall effect. The theoretical approach to be used here is not the most rigorous; however, it has intuitive value and leads to much the same result as the more sophisticated theoretical and empirical methods. Our model will again be the vector model shown in Fig. 1-2. We will assume that our semiconductor sample is intrinsic with equal electron and hole mobilities. In this way, the application of a magnetic field will deflect both holes and electrons in the same

direction with resultant recombination at the edges of the sample. Therefore, no Hall field will be built up to balance the Lorentz force of the magnetic field and the electrons and holes will continue to be deflected. As the effect is usually small, we will assume the following two conditions to be true:

(1) The angle θ between the direction of the field and the direction of current flow is small.

(2) The path of the carrier in the material is a straight line.

The force exerted on the carriers in the x direction is,

$$(1\text{-}16) \quad |\bar{F}_x| = e\, v_y B$$

and the magnitude of the force in the y direction is,

$$(1\text{-}17) \quad |\bar{F}_y| = e E_y$$

Thus there is seen to exist an angle θ which expresses the deflection of the current carriers in the magnetic field and, if slight, is given by,

$$(1\text{-}18) \quad \theta \approx \tan\theta = \frac{v_y B}{E_y} = \mu_H B$$

We shall now take the following expression for the conductivity tensor,

$$(1\text{-}19) \quad \sigma = \frac{\bar{J}\cdot\bar{E}}{E^2}$$

which for our model is,

(1-20) $$\sigma = \frac{J \cos\theta}{E_y}$$

If we assume J/E_y as the definition of the zero field isotropic conductivity, σ_0 , then (1-20) resolves to,

(1-20a) $$\sigma = \sigma_0 \cos\theta \text{, OR, } \rho = \rho_0 \sec\theta$$

Since the angle θ is small, we may use the first order approximation to the secant function.

(1-20b) $$\sec\theta = 1 + \frac{\theta^2}{2} + \frac{5\theta^4}{24} + \frac{61\theta^6}{720} + \cdots \cong 1 + \frac{\theta^2}{2}$$

Combining (1-20a), (1-20b), (1-18), and (1-11), and rearranging terms, we have for our final equation for the transverse magnetoresistance effect,

(1-21) $$\frac{\Delta\rho}{\rho_0} = \frac{1}{2}\left(\frac{3\pi\mu}{8}\right)^2 B^2$$

where $\Delta\rho = \rho - \rho_0$.

Equation (1-21) agrees nearly with the equation given by Welker and Weiss as:[1]

(1-22) $$\frac{\Delta\rho}{\rho_0} = \frac{4-\pi}{\pi}\left(\frac{3\pi\mu}{8}\right)^2 B^2$$

Unfortunately, the derivation of this equation is not given, so that the discrepancy in the constants of proportionality can

not at this time be explained. It may well be that equation (1-22) expresses the transverse magnetoresistance of a Corbino disc, because it is this configuration, and not the rectangular shape, that is most commonly used. The magnitude of the transverse magnetoresistance effect depends largely upon the degree to which a Hall field is present. The greater the Hall field, the more effective will be the balancing of the Lorentz force. There will always be some magnetoresistance in even the highly doped semiconductors because of hole-electron recombination which causes a small current to flow in the x direction and because the Lorentz force holds in equilibrium only those carriers with the average Boltzmann velocity and all others in the normal distribution will be deflected. In passing, it is interesting to note that a result exactly equal to (1-21) can be derived by considering the increase in path length that the current must take, this increase being $l = l_0 \sec \theta$.

II. EXPERIMENT

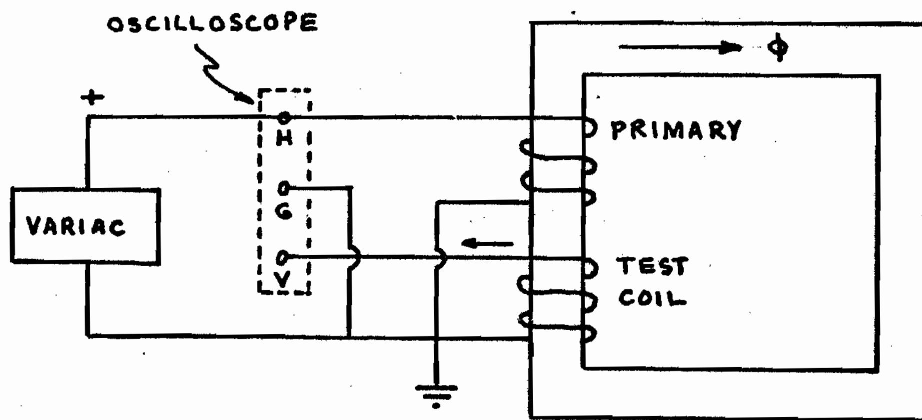
A. Apparatus and Equipment

In order to perform any experimental work involving the magnetoresistance and the Hall effect, a uniform, high flux density magnetic field source is absolutely essential. Consequently, obtaining such a source was of prime importance in order to permit the performance of the specified task. A search of the facilities of the Physics and Electrial Engineering laboratories indicated that no equipment already suitable for this purpose was on hand. However, a war surplus filament transformer was found which gave promise of being adapted to the production of the desired magnetic field. This transformer can be seen in detail in the photographs of the experimental apparatus in Appendix E. The name plate listed the characteristics as a General Electric model number 7470650 with a primary rating of 100 volts at 60 cycles and a secondary rating of 2.5 volts with an overall capacity of .100 KVA. No information was given concerning the direction of the primary winding nor its number of turns. Also, some idea as to the magnetic properties of the transformer iron had to be obtained as well. As a result, several experiments were planned to determine these various unknown quantities.

As it was intended that the primary be used as the source of excitation, the secondary winding was not needed and was removed from the core. It was noticed that this winding contained 10 turns, thus leading one to suspect that, from the

name plate data, the number of effective primary turns must be in the neighborhood of 400. By winding a test coil with a given number of turns upon this core, it was found that the number of effective turns depended upon the applied test voltage, thus indicating that a non-linear component of leakage flux was developed in the primary. At low voltages, the number of effective turns was about 480, at rated voltage it nearer the expected 400 value, and at higher than rated voltages, the effective turns decreased to nearly 300.

The test coil was comprised of 29 turns wound in a predetermined direction. It was now necessary to know the direction of flux produced in the core when a voltage of given polarity was applied to a given primary terminal. The circuit shown in Fig. 2-1 was constructed and the following procedure used.



FLUX DIRECTION TEST CIRCUIT

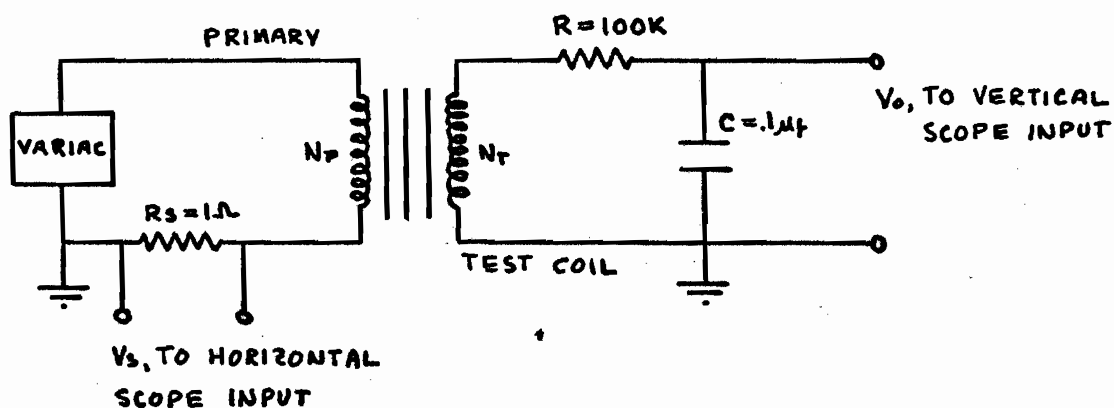
Fig. 2-1

(1) One terminal of the primary was designated positive and the other terminal grounded.

(2) If the applied voltage and the test voltage were in phase as indicated by the Lissajous pattern, then a positive voltage applied to the positive terminal of the primary would have produced a flux in the same direction as a positive voltage applied to the ungrounded end of the test coil.

(3) If the voltages were out of phase, then a positive voltage applied to the positive terminal of the primary would have produced a flux in the same direction as a negative voltage applied to the ungrounded end of the test coil.

After the determination of the flux direction, it was desired to gain some feeling for the magnetic properties of the iron composing the magnetic circuit. This objective could be most easily fulfilled by obtaining the AC hysteresis characteristics. Consequently, the circuit shown in Fig. 2-2 was built for this purpose, consisting primarily of an integrating circuit placed between the test coil output and the vertical input of an oscilloscope.



HYSTERESIS TEST CIRCUIT

Fig. 2-2

The following equation for the flux density B as a function of the output voltage Vo can be easily shown if one assumes the RC product to be much less than the angular frequency of the applied voltage.

$$(2\text{-}1) \quad B = \frac{10^4 RC V_o}{A N_T} \text{ gauss}$$

Vo is in volts, N_T is the number of test coil turns, and A is the area of the core in square meters. Similarly, an equation for the field intensity H in terms of the voltage drop across the primary series resistor can be derived from simple magnetic circuit considerations.

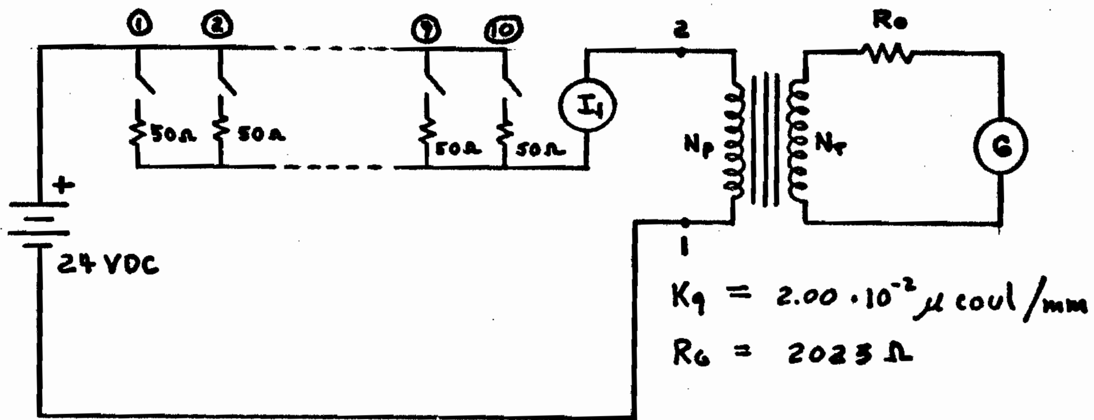
$$(2\text{-}2) \quad H = 1.26 \cdot 10^{-2} \frac{N_P V_S}{l_P R_S} \text{ oersted}$$

Vs is in volts, Np is the number of effective turns on the primary, Rs is the resistance of the primary series resistor in ohms and lp is the mean length of the magnetic path in meters. The resulting hysteresis pattern is shown in photograph E-1 and graph 2-1. From the linear part of pattern, the average permeability is estimated to be in the neighborhood of 7000.

To facilitate the performance of the Hall effect and magnetoresistance effect experiments, it was necessary that a gap be cut in the core in which the germanium samples could be placed. Several groups were approached with this problem. The work was finally accomplished by the Kendall - Lamar

Corporation of Potsdam, New York, where a 3/32" slot was machined through the core. The magnetic characteristics now needed to be examined again, but this time in more detail than before. The AC hysteresis test was first repeated and resulted in the pattern shown in photograph E-2. As can be seen, the air gap characteristic is for all practical purposes single valued and quite linear in portions along with having low retentivity at zero applied field. It was felt, however, that although the hysteresis characteristic gave a good qualitative picture of the general properties of the magnetic circuit, the presence of non-linear flux leakage and the fact that this was an AC characteristic voided its use as a calibration curve giving the flux density as a function of the primary current. It was therefore decided to calibrate the magnetic circuit through the use of the standard ballistic galvanometer test.

Because the particular galvanometer to be used had been inactive for a long period of time, it was dismantled and completely readjusted and recalibrated. The circuit shown in Fig. 2-3 was then constructed for the purpose of obtaining the DC magnetic characteristics.



DC MAGNETIZATION CURVE CIRCUIT

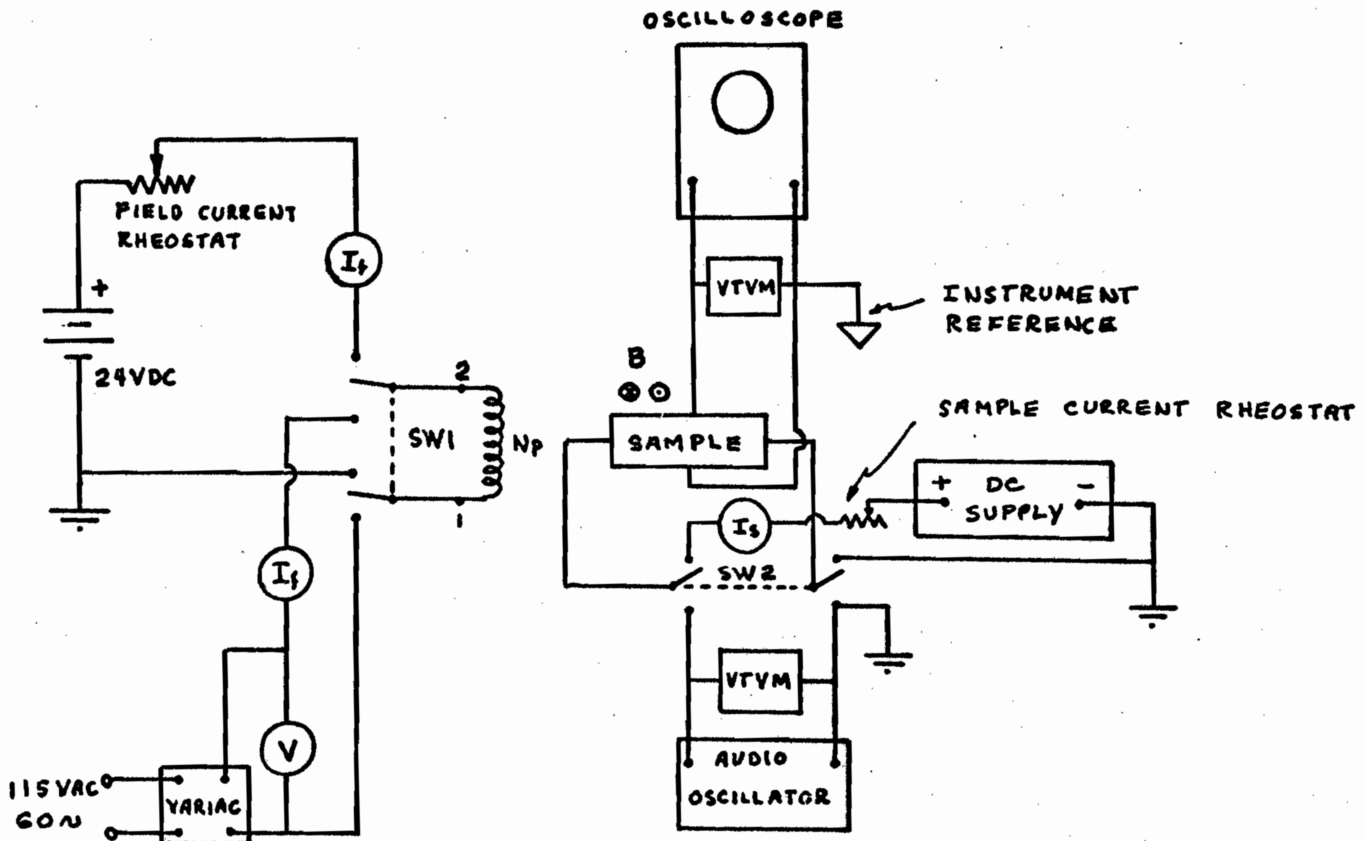
Fig. 2-3

The galvanometer constants shown in the diagram were obtained during its calibration. It is a simple matter to show that the standard equation expressing the total flux density in the circuit as a function of the total galvanometer deflections is

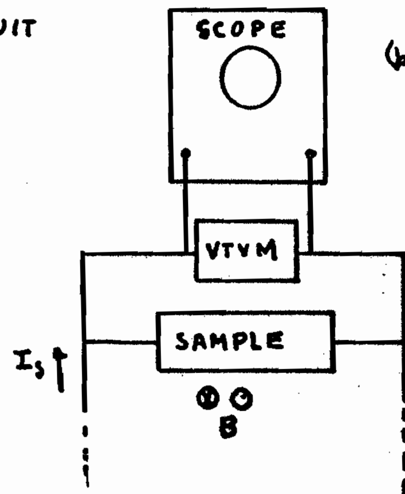
$$(2\text{-}3)\quad B_n = \frac{10^{-2} R K_q}{N_T A} \sum_{i=1}^{N} D_i \quad , \quad R = R_0 + R_G$$

If R is in ohms, A in square meters, Kq in μcoul/mm and Di in mm, then Bn, the total flux density linking the test coil N_T, will be expressed in gauss. The theory leading to this equation can be found in any good reference on electrical measurements.* The resulting calibration curve is displayed by graph 2-2 which was used to compute the air gap flux density from the field current, If, which was a quantity measured during the Hall effect and magnetoresistance effect experiments.

The complete circuit diagram for the measurement of the Hall and magnetoresistance effects is shown in Fig. 2-4.



(a) HALL EFFECT MEASUREMENT CIRCUIT



(b) MODIFICATION FOR MAGNETORESISTANCE EFFECT MEASUREMENTS

EXPERIMENTAL CIRCUIT DIAGRAMS

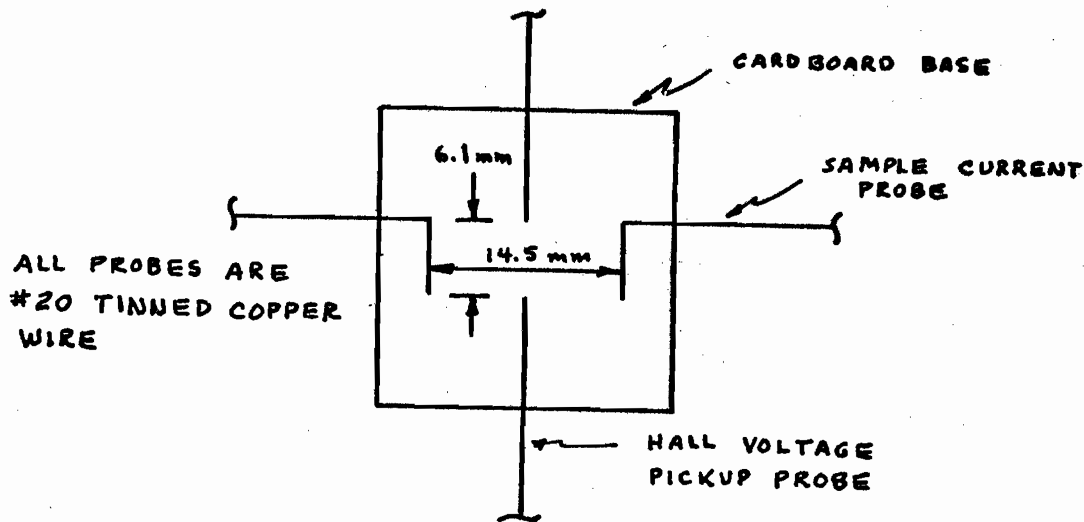
Fig. 2-4

The DC field current necessary for the production of the air gap flux was drawn from a source consisting of two 12 volt storage batteries in series. The amount of field current desired was determined through the setting of a motor speed control rheostat placed in series with the batteries. In this manner, up to approximately 5 amperes could be made to flow through the primary winding. For the purpose of demagnetizing the iron, or for providing a sinusoidal air gap flux, another circuit consisting of simply a variac could be switched into the primary circuit. The DC sample current, Is, was obtained from a standard regulated bench power supply and controlled through a slide wire rheostat. For the purpose of providing an AC sample current, a Hewlett - Packard audio oscillator was obtained which could be put into the sample circuit via SW2.

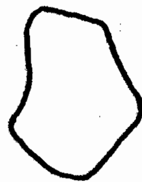
Voltages as low as those encountered in the Hall effect are usually measured with high precision potentiometers; however, an instrument of this type was not readily available and because of the thinness of the samples, the Hall voltage was great enough to be measured with an oscilloscope and a VTVM. The two instruments were used together in order to obtain a check on the consistency of the readings of each other. This method of measurement provided a convenient way to balance out the zero field voltages via the vertical position control on the oscilloscope and the zero position control on the VTVM. The accuracy of all equipment used was checked before commencing experimental operations.

The circuit and equipment for the performance of the Hall and magnetoresistance effect measurements can be seen in photograph E-6. In general, the equipment pertinent to the production of the magnetic field is on the left of the picture while that apparatus used to supply the sample current and measure the resulting phenomena are on the right.

The semiconductor samples used in this work were supplied through the courtesy of the International Business Machines Corporation of Poughkeepsie, New York. The samples were thin cross-sections of a larger crystal and their surfaces were etched. Various attempts were made to solder to these surfaces, but no satisfactory results were obtained. Also, the thin crystallites had irregularly shaped surface areas and, since diamond abrasive machinery was not available, some means had to be found to lend some semblence of regular geometry to the sample. Both the problem of connection and the problem of geometry were partially solved by the construction of a somewhat crude sample holder shown in Fig. 2-5(a).



(a) SAMPLE HOLDER ASSEMBLY (TOP VIEW, NOT TO SCALE)



(b) SAMPLE NO. 1 - N TYPE
(TOP VIEW, FULL SCALE)
THICKNESS - .37 mm

(c) SAMPLE NO. 2 - P TYPE
(TOP VIEW, FULL SCALE)
THICKNESS - .17 mm

SAMPLES AND ASSEMBLY
Fig. 2-5

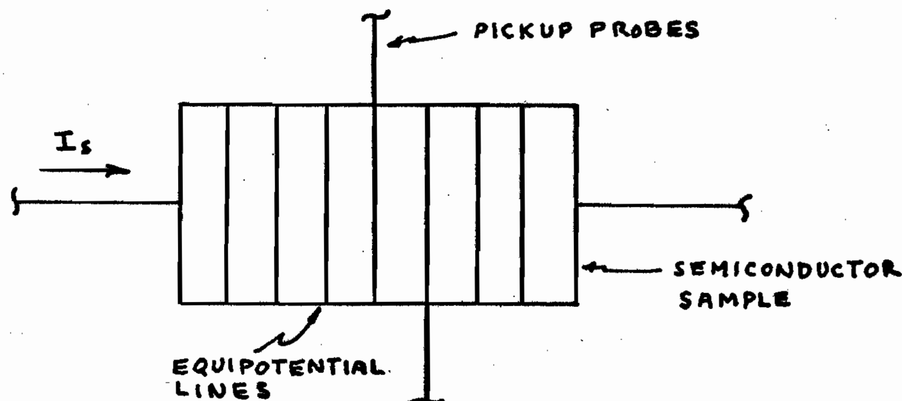
The probes were formed from #20 tinned copper wire and placed on a matchbook cover card board base with Scotch tape. One of the samples shown in Figures 2-5(b) and 2-5(c) was then placed on top of the probes and taped to hold it in position.

A second matchbook cover was laid on top of the sample and the entire assembly then fitted snugly into the air gap of the magnetic circuit. It was necessarily assumed that fringing between the two sample current probes was negligible and all subsequent calculations were carried out on this basis. Since nothing was known about the electronic parameters of the samples, except their carrier type (which later was proved to be wrong), the determination of these values became the object of the experimental work.

A close-up of the magnetic circuit with the sample holder in position can be seen in photographs E-8 and E-9.

B. Experimental Procedure and Results

Associated with the voltage produced at the pickup probes on the semiconductor sample are voltages due to several other effects as well as the Hall effect. The most important and largest of these voltages is the so-called IR drop. This voltage arises from the fact that it is difficult to allign the pickup probes so that they rest on an equipotential line. This effect is shown in Fig. 2-6, and, while it cannot be completely eliminated, a small amount of manipulation of the pickup probes can minimize the IR drop.



SAMPLE POTENTIAL DISTRIBUTION IN ABSENCE OF MAGNETIC FIELD

Fig. 2-6

The other sources of potential besides the Hall effect are the thermomagnetic effects. These are the Ettinghausen effect, the Nernst effect, and the Righi - Leduc effect. They will now be discussed separately.

The Ettinghausen effect, producing a voltage V_E, is similar to the potential difference associated with a thermocouple. The sign of V_E, as in the Hall effect, depends upon the direction of the current in the sample and the direction of the applied magnetic field. The Nernst effect and the Righi - Leduc effect are caused by the combined efforts of a temperature gradient in the sample and the presence of an electric current and a magnetic field. The voltages produced by these effects will be denoted V_N and V_{RL}, respectively. The signs of both voltages depend only on the direction of the magnetic field. Thus, for each possible combination of current and field in the sample, the voltage appearing at the pickup probes becomes,[7]

$$
\begin{aligned}
(2\text{-}4)\quad &(+B,+I) \quad E_1 = V_H + V_E + V_N + V_{RL} + V_{IR} \\
&(+B,-I) \quad E_2 = -V_H - V_E + V_N + V_{RL} - V_{IR} \\
&(-B,-I) \quad E_3 = V_H + V_E - V_N - V_{RL} - V_{IR} \\
&(-B,+I) \quad E_4 = -V_H - V_E - V_N - V_{RL} + V_{IR}
\end{aligned}
$$

Through proper algebraic manipulation of (2-4), we obtain

$$
(2\text{-}5)\quad V_H + V_E = \frac{E_1 - E_2 + E_3 - E_4}{4}
$$

Therefore, by taking readings at all possible combinations of field and current, we may eliminate all the effects except the Ettinghausen effect. Since the Ettinghausen effect, the Nernst effect, and the Righi - Leduc effect are usually very small when compared to the Hall effect, and since we were blessed in this experiment with a quite large Hall voltage, the above

procedure was followed only occasionally, and, when performed, did indicate that the other effects were truly negligible. There are ways, however, of performing the experiment under AC conditions where equation (2-5) expresses the average value of the output voltage.[7] The only effect, for our purposes, worth considering is the IR drop. However, this factor was easily overcome through adjustment of the balance controls on the instruments at the beginning of each experimental run. Equations (2-4) and (2-5) give no indication of the proper sign of the Hall voltage; however, the usual convention is to take the sign of the Hall voltage produced by an N type semiconductor as negative and by a P type as positive.

As the measurements were begun, it was soon found that the sign of the Hall voltage was reversed from that expected using the sign determining method given in Fig. 1-2 of the Theory. The semiconductor materials had been received with their carrier types indicated on their containers; however, those marked P type gave N type responses and those described as N type yielded P type responses. After carefully checking the direction of the sample current and the magnetic field, it was concluded that the samples had been received with incorrect identification.

Curves of Hall voltage versus flux density at constant sample current for both the N and P type samples are displayed in graphs 2-3, 2-4, and 2-5. Graph 2-3 depicts the characteristics for the N type sample. For the most part, this family of curves follows the functional relationship given by equation

(1-13); that is, at constant sample current, the relationship between the Hall voltage and the magnetic field flux density should be linear. On the other hand, the characteristics of the P type sample show a slight curve. At present, the reason for this non-linearity is unknown; however, it has been noticed before by other observers. Knapton, in a Master's Thesis submitted at Williams College[5], displays a straight line characteristic for his N type sample and a curved characteristic, similar to graphs 2-4 and 2-5, for his P type sample. His explanation is that temperature changes during the taking of measurements caused the parameters of the sample to drift. In our case, any temperature changes would have occured equally during the testing of both samples with the result that the characteristics of both N and P types would be slightly non-linear. Therefore, it is doubtful that the proposed temperature change could account fully for the curves in the P type sample characteristics. The only reason that comes to the mind of the author is that the orientation of the crystal axes in the electric and magnetic fields for a P type material may be more critical than for an N type because of the stronger effect of the lattice potential on the shape of the top of the valence band than on the shape of the bottom of the conduction band. Graph 2-5 shows that as the P type sample was thermally agitated into becoming intrinsic, the resulting characteristic became linear. At any rate, the curvature of the P type sample characteristics was not so great as to forbid an approximation of the Hall coefficient over the more linear portions.

As each sample was tested, data were obtained to compute the specific resistivity. As there were no thermocouples available, the same data were used to calculate the power input to the sample so that an indication of the Joule heating effect of the sample current could be obtained. The resistivity was then calculated assuming a solid rectangular geometry so that the familiar formula $R = \frac{\rho l}{A}$ could be applied. The results are shown in graph 2-6 which is a plot of sample resistivity versus power input. It is easily seen that that the behavior is typical of a doped semiconductor. As the temperature of the sample is increased, more and more impurity levels become emptied and the resistivity decreases due to the additional carriers. The case is slightly different for each type of extrinsic semiconductor so let us consider each one in turn. In the N type, transitions begin to occur directly from the valence band with the resultant production of hole-electron pairs. As the temperature increases further, the contribution by hole-electron pairs becomes negligible and the semiconductor reverts to its intrinsic state. In the P type, electron transitions normally take place between the valence band and an acceptor level. However, at higher temperatures, hole-electron pair creation becomes predominant and the material becomes intrinsic. So far there has been no difference between the thermal properties of N and P type semiconductors. But, in the P material, as transitions from the valence to conduction band become probable, transitions from acceptor levels to the conduction band become even more probable, with

the result that the net amount of holes over electrons decreases. This fact, coupled with the inherently higher mobility of electrons, indicates that at some temperature, a P type material, as well as becoming intrinsic, will become slightly N type. This point is indicated on the resistivity characteristic for the P type sample shown in graph 2-6. The regions to the left and right of graph 2-6 are regions of extrinsic and intrinsic conduction, respectively. The regions of rapidly decreasing resistivity are regions where hole-electron pair production is overcoming the effect of the impurity carriers and the transition from extrinsic to intrinsic is not quite complete.

Graph 2-7 is designed to give a qualitative insight into the thermal behavior of the Hall coefficient, found by equation (1-13) and the slopes of the curves on graphs 2-3, 2-4, and 2-5. The magnitude of the Hall coefficient is defined here as being positive if the Hall voltage is of the normal polarity for the sample in question. The reason for the general decrease in the Hall coefficient for both impurity types can be easily explained. From equation (1-12) one can see that R_H is inversely proportional to the number of carriers in the extrinsic region, so that as the number of impurity carriers available increases, the Hall coefficient decreases. In the intrinsic region, the relation between R_H and the number of carriers becomes slightly more complicated. In fact it is[8]

$$(2\text{-}6) \quad R_H = \frac{3\pi}{8e} \frac{n_h \mu_h^2 - n_e \mu_e^2}{(n_h \mu_h + n_e \mu_e)^2}$$

However, it can be seen that as the number of electron-hole pairs increases, the Hall coefficient will still decrease. The fact that a P type semiconductor will become slightly N type when thermally excited into the intrinsic region has been mentioned before. This situation is clearly depicted by the P sample characteristic on graph 2-7. The transition to N type is shown by the reversal in sign of R_H. Consecutive readings taken in this region were not consistent; however, the zero crossing occurred with 50 to 60 milliamperes of sample current flowing through the semiconductor.

The room temperature ($P_s \approx 0$) Hall coefficients of the N and P samples were 12000 cm^3/coulomb and 8060 cm^3/coulomb, respectively. At higher sample currents, R_H for the P type sample became slightly greater, but it is believed that this is due to difficulties encountered with the curvature of the characteristic. The room temperature resistivities were 14.35 ohm·cm for the N type and approximately 15 ohm·cm for the P type. The values for R_H compared favorable with those ranges found in the literature, being between 1000 and 20000 cm^3/coulomb, with R_H for the N type sample being greater than R_H for the P sample.[5,7,9] Through the use of the following equation derived from equations (1-10) and (1-11), we can find the mobility assuming only carriers due to impurity centers exist at room temperature.

$$(2\text{-}7) \quad \mu = \frac{8}{3\pi} \frac{R_H}{\rho}$$

Application of equation (2-7) yields values of 770 cm^2/volt·sec for the electron mobility μ_e in the N type sample, and 495 cm^2/volt·sec for the hole mobility μ_h in the P sample. There values disagree with the commonly accepted electron and hole mobilities in germanium of 3600 cm^2/volt·sec and 1700 cm^2/volt·sec, respectively.[7] The discrepancy is believed to lie in the values obtained for the resistivities. Surface contact connections to semiconductors are known to exhibit a wide range of resistances which are functions of pressure, surface state, rectification, and other variables. In view of these facts, it would not be too surprising if the recorded resistivity data were greater by a factor of two or three than the true resistivity. Certainly the values of R_H are reasonable enough so that it is quite probable that the errors in the calculated mobilities have their origin in poor resistivity measurements. The assumption of rectangular geometry could also have been in error if fringing effects were not negligible or if the sample current probes were not making uniform contact with the semiconductor. Some better means, such as the four probe method, should have been used to determine the resistivity than that of measuring only the voltage drop between the sample current probes. If a large piece of the crystal from which the samples were cut had been available, the four probe method of finding resistivity could have been utilized.

By assuming that all impurity levels are ionized at room temperature and knowing the number of atoms per cubic centimeter in the germanium crystal, one can use the Hall coefficient to

determine the concentration of impurity atoms in the material. From equation (1-12), it can be shown that

$$(2\text{-}7a) \quad K = \frac{3\pi/8}{R_H N_{Ge} e}$$

where K is the number of impurity atoms per germanium atom, R_H is the Hall coefficient, N_{Ge} is the concentration of germanium atoms, and e is the electronic charge. Taking N_{Ge} as $4 \cdot 10^{22}$ atoms/cm^3,[7] we have K equal to $1.53 \cdot 10^{-8}$ for the N sample or there is one donor impurity for $6.55 \cdot 10^{7}$ atoms of germanium at room temperature. Similarly, for the P type sample, we find that there is one acceptor impurity for every $4.35 \cdot 10^{7}$ atoms of germanium. These figures agree in magmitude with typical impurity concentrations found in commercial extrinsic germanium.[7]

Attempts to verify equation (1-21), which expresses the magnetoresistance effect, generally ended in failure. The major reason for failure was the unavailability of an intrinsic sample with the optimum geometry for the effect to occur. The best data that could be obtained were taken on the N type sample and the results are presented in graph 2-8. The change in resistance was measured by holding the sample current, Is, constant and reading the voltage drop across the sample current probes as the magnetic field was gradually increased. The relative increase in resistance or resistivity could then be expressed as:

$$(2\text{-}8) \quad \frac{\Delta R}{R_o} = \frac{\Delta V_{mr}}{V_o} = \frac{V_{mr} - V_o}{V_o}$$

where Vo is the zero field voltage drop in the direction of the sample current, and Vmr is the total voltage drop in the presence of the magnetic field. The effect did not seem to be present at all when using sample currents above one milliampere. At sample currents of 1 ma. and .5 ma., however, the effect was quite noticeable in the lower ranges of magnetic field flux density, indicating generally the usual parabolic relationship between the change in resistivity and the value of the magnetic field. On the other hand, at fields above 5000 gauss the behavior of the effect became completely anomalous, indicating a decreasing resistivity with increasing magnetic field. The reason for this reversal may arise from two factors. The first is that there is always going to be a Hall effect associated with an extrinsic semiconductor which does not have the optimum geometry for the magnetoresistance effect. However, in our case, because of the low sample current involved, the Hall field may not have been built up to sufficient strength at low flux densities to prevent some deflection of the sample current vector, thus producing a magnetoresistance effect. As the magnetic field increased, the Hall field may have increased to a value where any attempt of the sample current vector to deflect was fully compensated by the Hall field with a resultant tendency toward normal resistivity. The other factor that may have influenced events was the possibility of a junction

existing between the tinned sample current probe and the surface of the semiconductor which was sensitive to a magnetic field. It is not known whether such junctions are possible, but the question cannot be overlooked.

As one of the applications of the Hall effect is in multiplying or modulating devices, it was decided to try to modulate a 2000 cycle sample current "carrier" with a 60 cycle magnetic field. The results are shown in photographs E-3, E-4, and E-5. Before discussing these photographs, let us try to analyze what form the resultant Hall voltage will take. We can substitute sinusoidal expressions for the sample current and magnetic field, neglecting for the moment the IR drop, into equation (1-13) and obtain for the Hall voltage,

$$(2\text{-}9) \quad V_H = \frac{R_H I_s \cos\omega_s t \; B_o \cos\omega_f t}{T}$$

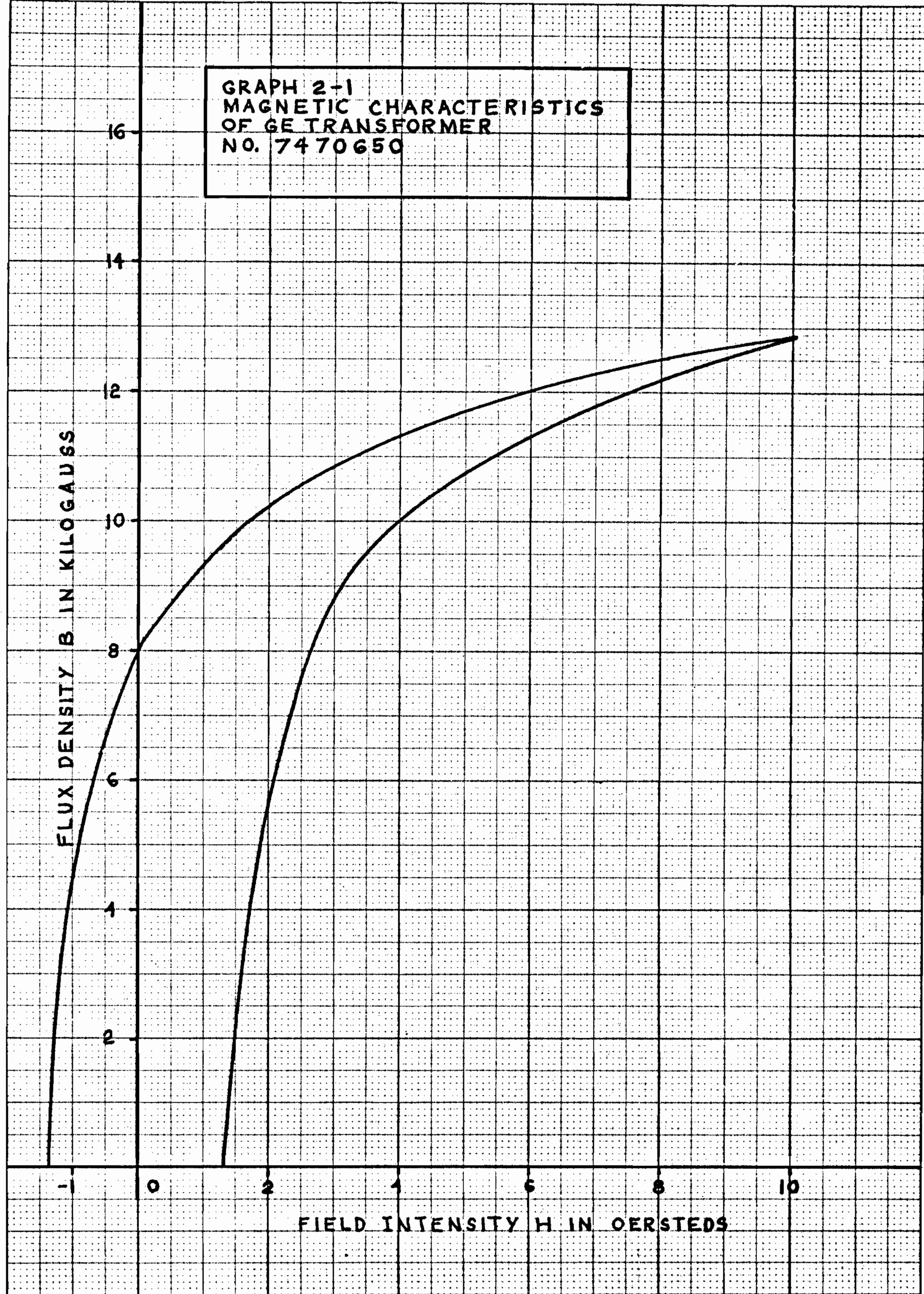
$$= \frac{R_H I_s B_o}{2T}\left[\cos(\omega_s+\omega_f)t + \cos(\omega_s-\omega_f)t\right], \quad \begin{array}{l}\omega_s - \text{carrier frequency}\\ \omega_f - \text{field frequency}\end{array}$$

where Is is the peak value of the sample current, Bo is the peak value of the magnetic flux density, ω_f is the angular frequency of the magnetic field, and the other symbols have their usual meaning. Thus it is seen that the output will be that of a double side band, suppressed carrier signal. In reality, however, there are three signals received at the pickup probes: the Hall voltage, the IR drop, and, when using alternating magnetic fields, an induced voltage, $V_o = -K\frac{dB}{dt}$. Under these conditions we have for (2-9),

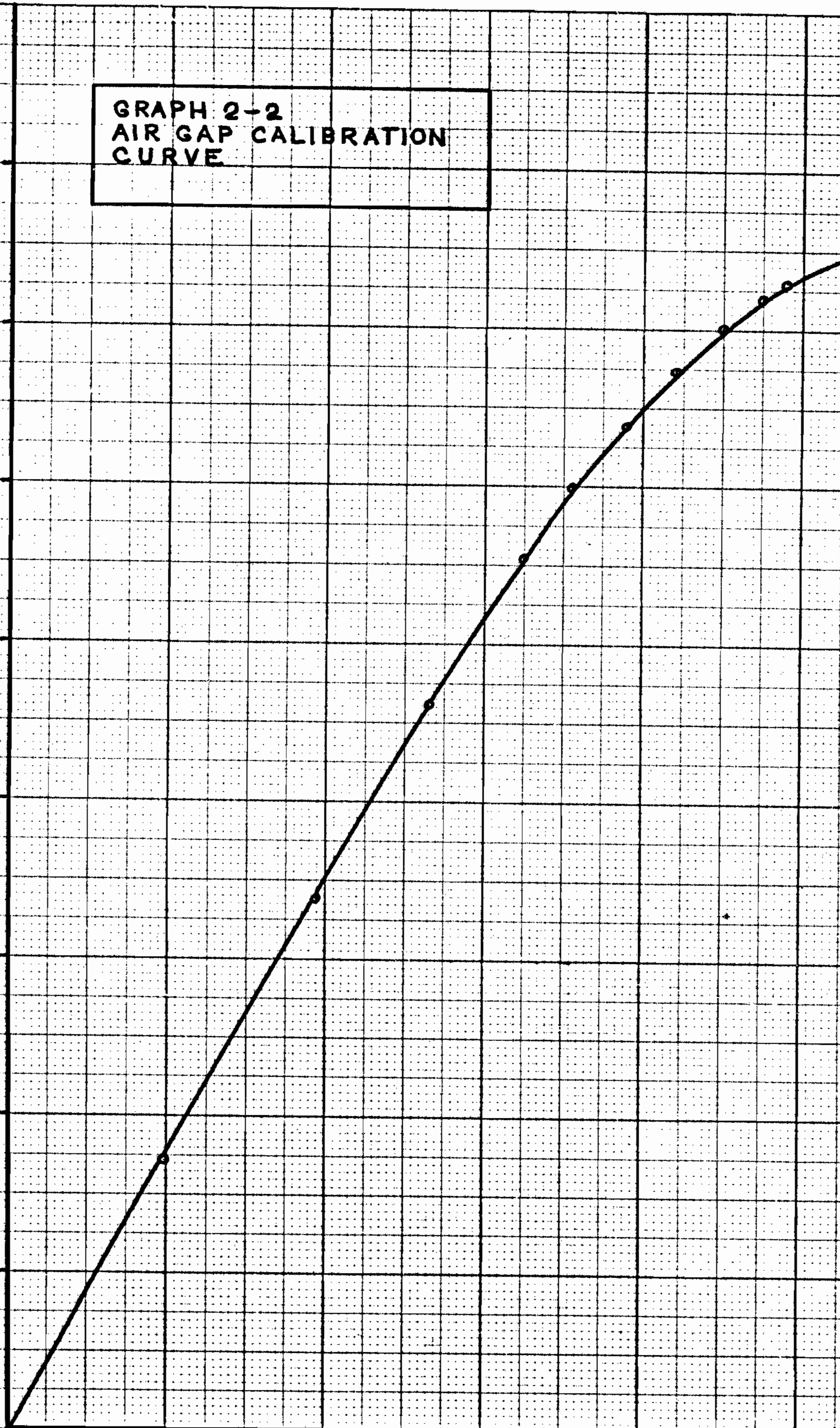
$$(2\text{-}10)\quad E_o = \frac{R_H I_s B_o}{2T} \cos \omega_s t \cos \omega_f t + I_s R \cos \omega_s t + K B_o \sin \omega_f t$$

$$= \frac{R_H I_s B_o}{2T} \left[\cos (\omega_s + \omega_f) t + \cos (\omega_s - \omega_f) t \right] + I_s R \cos \omega_s t + K B_o \sin \omega_f t$$

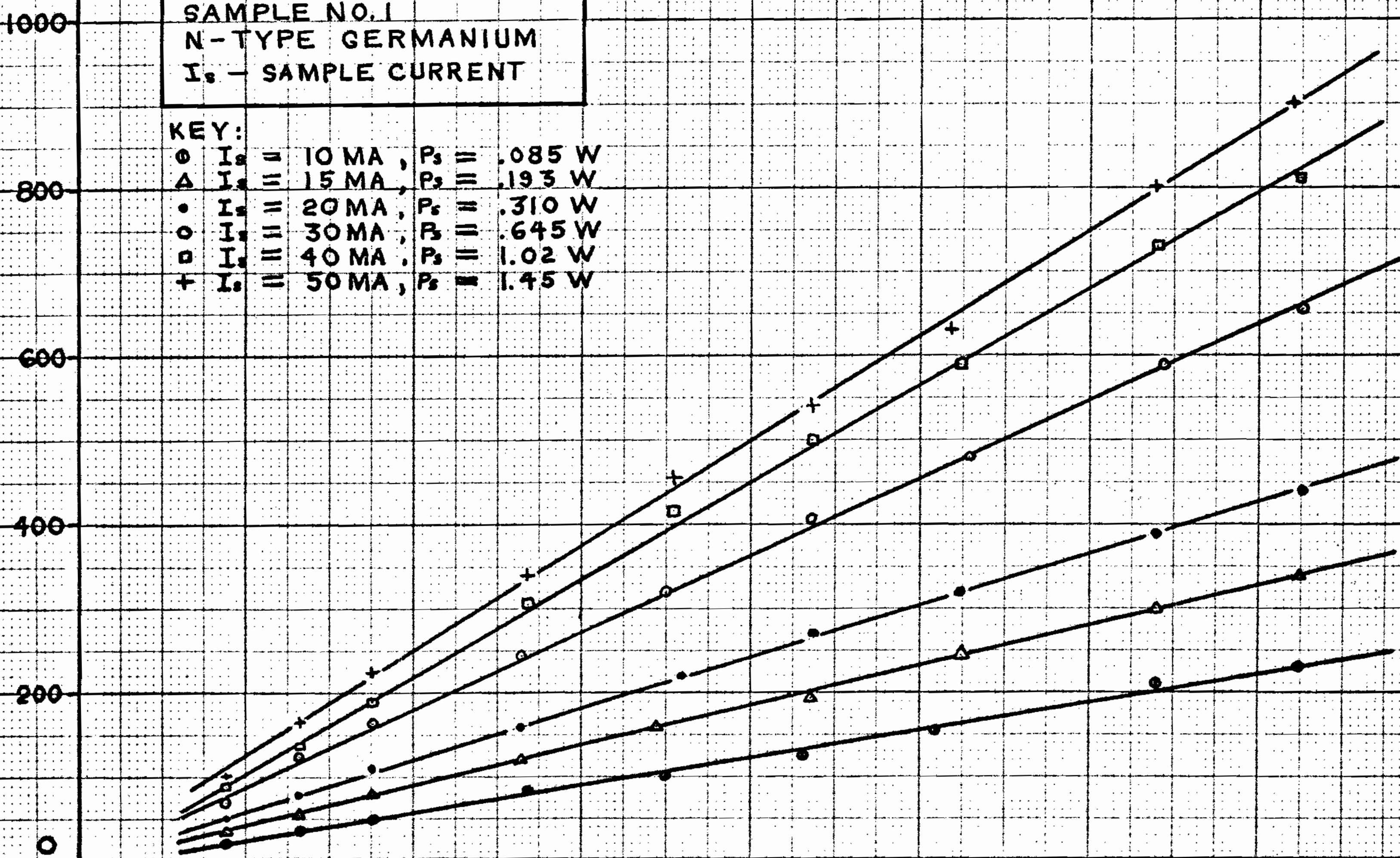
Equation (2-10) describes the result seen in photograph E-5. The signal output is similar to the familiar amplitude modulated carrier waveshape except for the phase shift caused by the last term in (2-10). Photograph E-3 depicts the induced voltage due to this term at zero sample current, and photograph E-4 displays the IR drop "carrier" at zero magnetic field or the quantity described by the second term in equation (2-10).



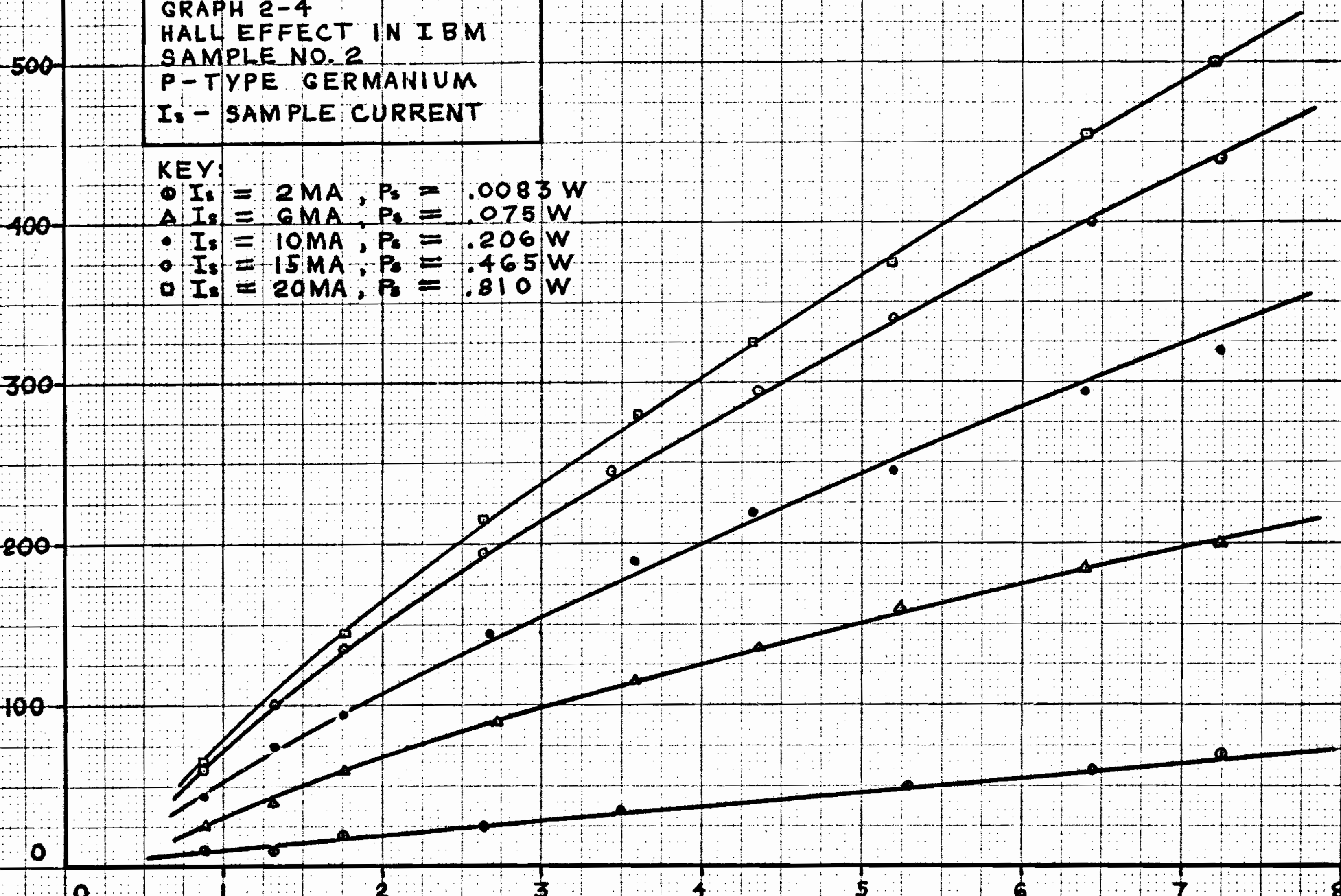
20 X 20 PER INCH

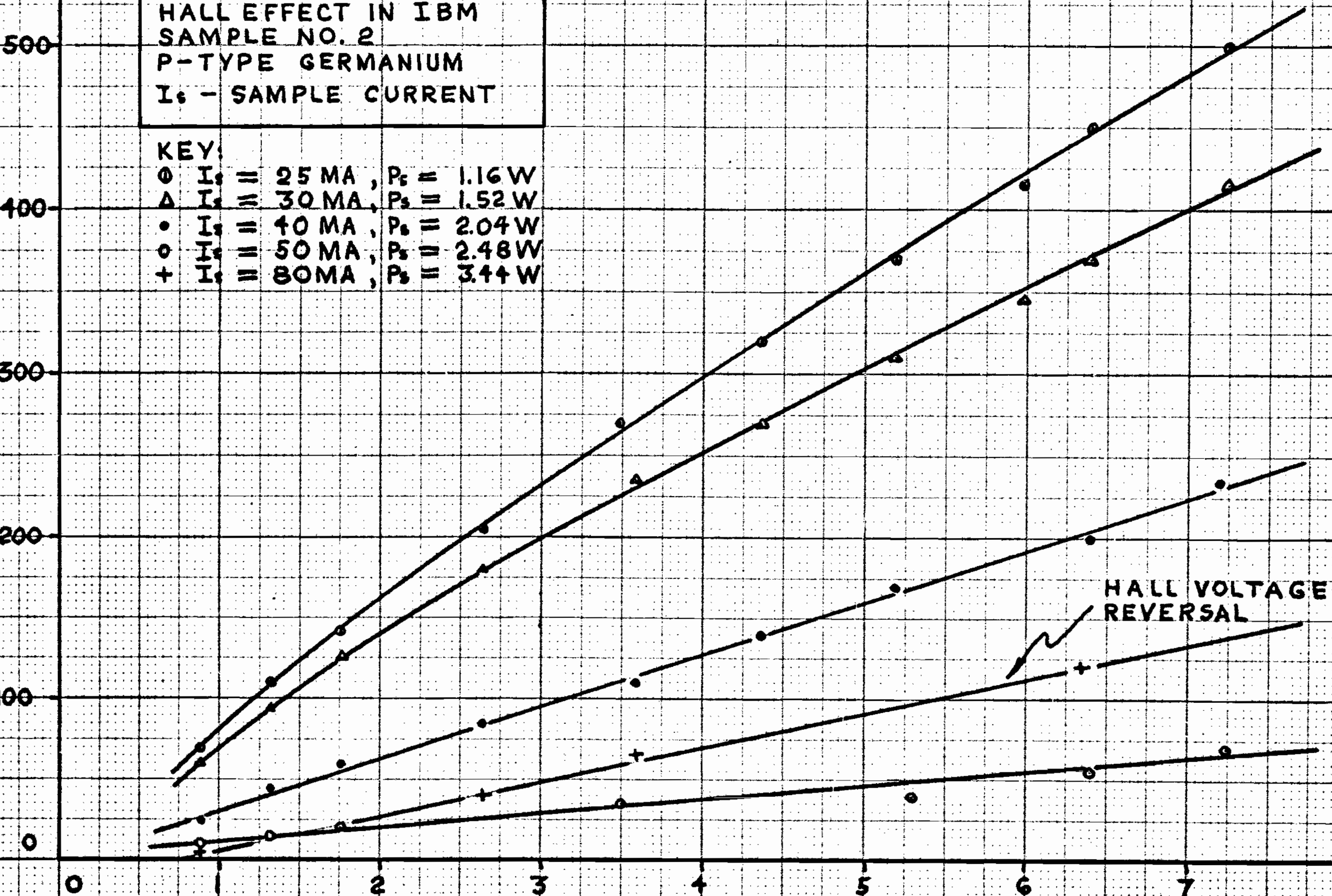


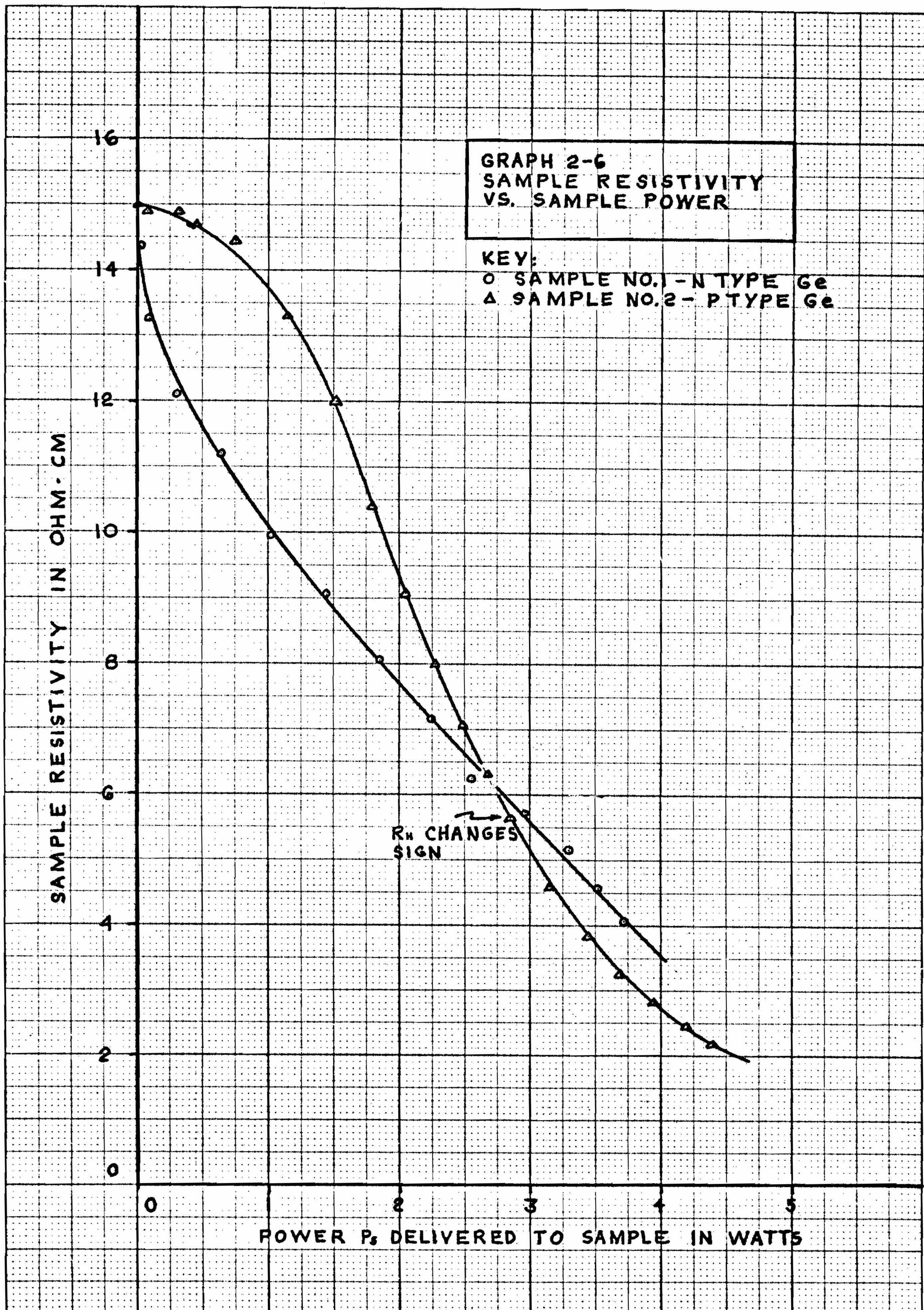
EUGENE DIETZGEN CO.
MADE IN U. S. A.
NO. 340 -20 DIETZGEN GRAPH PAPER
20 X 20 PER INCH



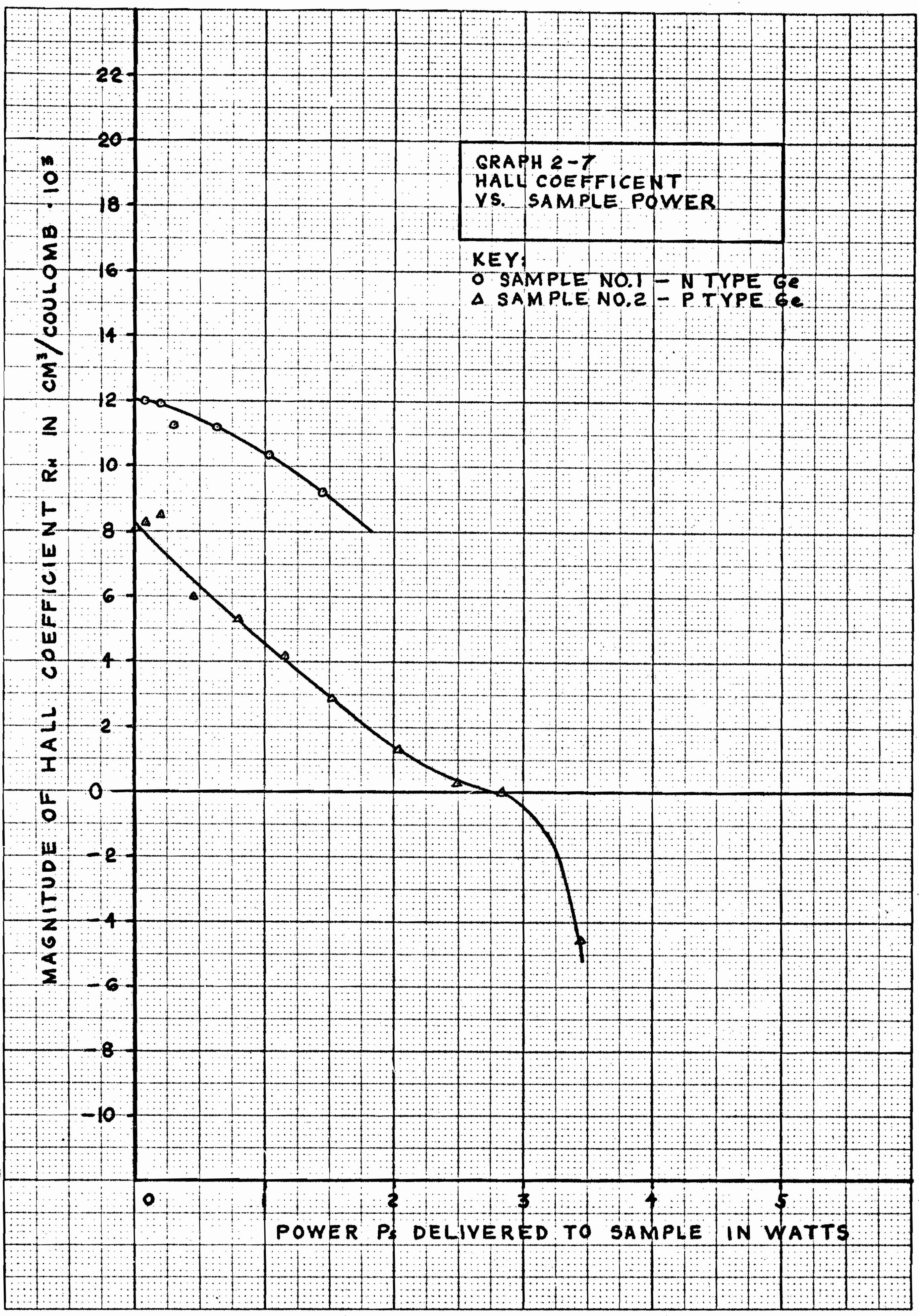
GRAPH 2-3
HALL EFFECT IN IBM
SAMPLE NO. I
N-TYPE GERMANIUM
Is — SAMPLE CURRENT
KEY:
Is = 10 MA, Ps = .085 W
Is = 15 MA, Ps = .193 W
Is = 20 MA, Ps = .310 W
Is = 30 MA, Ps = .645 W
Is = 40 MA, Ps = 1.02 W
Is = 50 MA, Ps = 1.45 W
HALL VOLTAGE VH IN MILLIVOLTS
1000
800
600
400
200
0
AIR GAP FLUX DENSITY B IN KILOGAUSS
0
1
2
3
4
5
6
7
8


GRAPH 2-4
HALL EFFECT IN IBM
SAMPLE NO. 2
P-TYPE GERMANIUM
Is - SAMPLE CURRENT
KEY:
Is = 2MA, Ps = .0083 W
Is = 6MA, Ps = .075 W
Is = 10MA, Ps = .206 W
Is = 15MA, Ps = .465 W
Is = 20MA, Ps = .810 W
HALL VOLTAGE VH IN MILLIVOLTS
0
100
200
300
400
500
0
1
2
3
4
5
6
7
8
AIR GAP FLUX DENSITY B IN KILOGAUSS


GRAPH 2-5
HALL EFFECT IN IBM
SAMPLE NO. 2
P-TYPE GERMANIUM
Is - SAMPLE CURRENT
KEY:
Is = 25 MA, Ps = 1.16 W
Is = 30 MA, Ps = 1.52 W
Is = 40 MA, Ps = 2.04 W
Is = 50 MA, Ps = 2.48 W
Is = 80 MA, Ps = 3.44 W
HALL VOLTAGE
REVERSAL
HALL VOLTAGE Vx IN MILLIVOLTS
AIR GAP FLUX DENSITY B IN KILOGAUSS
0
100
200
300
400
500
1
2
3
4
5
6
7
8
20 X 20 PER INCH
MADE IN U. S. A.



EUGENE DIETZGEN CO.
MADE IN U. S. A.
NO. 340-20 DIETZGEN GRAPH PAPER
20 X 20 PER INCH

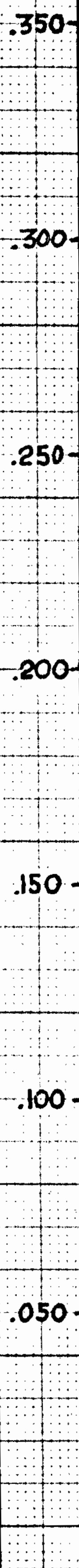


20 X 20 PER INCH
MADE IN U. S. A.

EUGENE DIETZGEN CO.
MADE IN U. S. A.

NO. 340R-20 DIETZGEN GRAPH PAPER
20 X 20 PER INCH

GRAPH 2-8



APPENDIX A

A MAGNETORESISTIVE CONTROL DEVICE

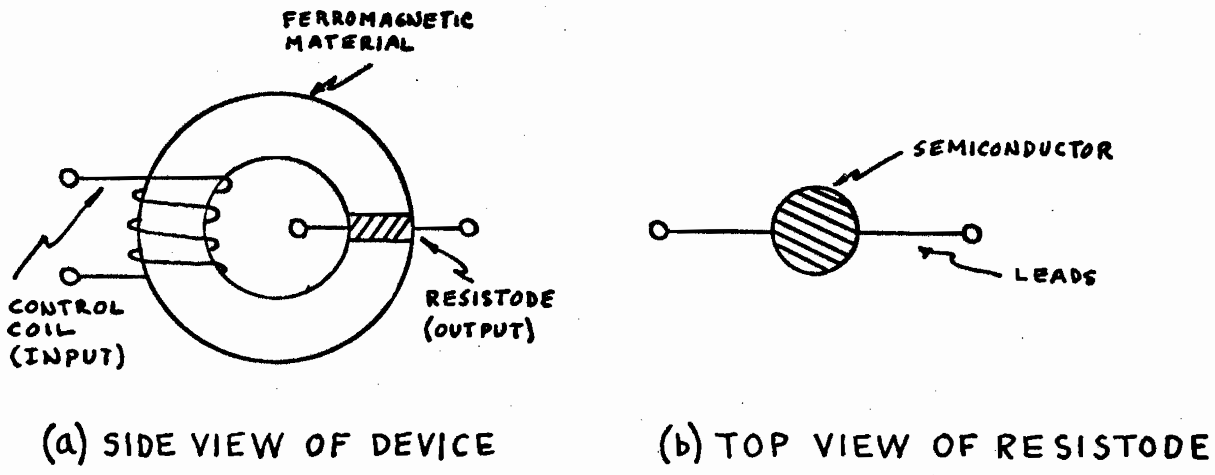
I. Theoretical Considerations of a Magnetoresistive Control Device (MCD)

The purpose of this appendix will be to outline the possibilities of a magnetoresistive control device. The magnetoresistance effect is associated with the same phenomenon that causes the well-known Hall effect. When a magnetic field is directed transverse to a current-carrying conductor, an increase in resistance, ΔR, proportional to the square of the magnetic flux density B will be observed[1]. This relation is given as

$$\frac{\Delta R}{R_o} = \frac{R - R_o}{R_o} = \alpha B^2 \qquad \text{(A-1)}$$

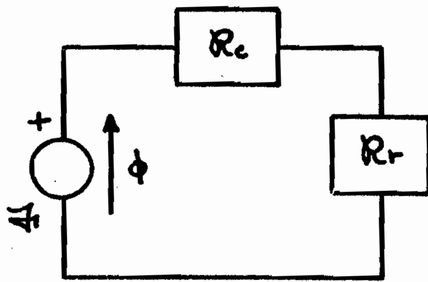
where α is a constant of proportionality which can be determined either experimentally or theoretically. In most conductors α is very small; however, recent investigations have shown that the increase in resistance over Ro, the "static" resistance, can be as great as 25:1 in InSb, a group III-V semiconductor[1]. This fact creates new possibilities for the practical application of this effect.

Let us now envision the following simplified design of an MCD.



HYPOTHETICAL MCD
Fig. A-1

The device depicted in Fig. A-1(a) consists essentially of a toroidal ring composed of some ferromagnetic material with a gap in which a semiconducting element or compound is placed. An input coil is wound on the toroid as shown for the purpose of creating the desired magnetic field which will pass through the semiconductor. It is evident that this configuration nearly eliminates mutual coupling between the input and output circuit. Also, it will be assumed that the flux linkage between the input and output is so small that Faraday induced output voltages due to time varying input currents can be neglected. For the sake of nomenclature, the element in which the resistance change takes place will be called the resistode and the input coil will be called the controller. The following magnetic circuit for the device can now be drawn.



MAGNETIC CIRCUIT

N - CONTROLLER TURNS
$\mathcal{F}$ - MAGNETOMOTIVE FORCE
ϕ - FLUX
$\mathcal{R}_c, \mathcal{R}_r$ - RELUCTANCE OF CORE, RESISTODE
l_c, l_r - PATH LENGTH OF CORE, RESISTODE
A_c, A_r - AREA OF CORE, RESISTODE
μ_c, μ_r - PERMEABILITY OF CORE, RESISTODE

Fig. A-2

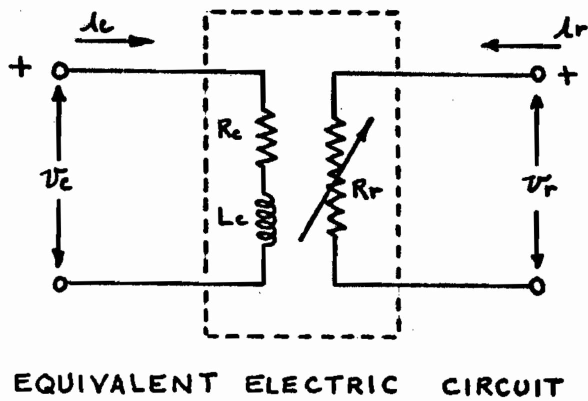
$$\text{(A-2)} \quad \mathcal{F} = \phi(\mathcal{R}_c + \mathcal{R}_r) = \phi\left(\frac{l_c}{\mu_c A_c} + \frac{l_r}{\mu_r A_r}\right)$$

Let us assume that the core length is short and that the magnetic material has high permeability and low loss (such as Supermalloy); then $\mathcal{R}_r \gg \mathcal{R}_c$ and

$$\text{(A-3)} \quad \mathcal{F} = \phi \frac{l_r}{\mu_r A_r}, \text{ or } N i_c = \frac{B l_r}{\mu_r}; \quad B = \frac{\mu_r N i_c}{l_r}$$

Thus we have a relationship between the flux density produced in the resistode and the instantaneous current flowing in the controller.

We shall now consider the following intuitive equivalent electric circuit of the MCD.



EQUIVALENT ELECTRIC CIRCUIT

v_r, i_r – INSTANTANEOUS RESISTODE VOLTAGE, CURRENT

v_c, i_c – INSTANTANEOUS CONTROLLER VOLTAGE, CURRENT

I_r, V_r – DC RESISTODE CURRENT, VOLTAGE

I_c, V_c – DC CONTROLLER CURRENT, VOLTAGE

L_c, R_c – CONTROLLER INDUCTANCE, RESISTANCE

R_r – RESISTODE RESISTANCE

Fig. A-3

Upon substitution of A-3 into A-1,

$$(A\text{-}4)\quad \frac{R_r - R_o}{R_o} = \frac{\alpha \mu_r^2 N^2 i_c^2}{l_r^2} = \gamma i_c^2 , \quad \gamma \equiv \frac{\alpha \mu_r^2 N^2}{l_r^2}$$

$$(A\text{-}5)\quad R_r = R_o \left(1 + \gamma i_c^2\right)$$

$$(A\text{-}6)\quad v_r = i_r R_o \left(1 + \gamma i_c^2\right) = f\left(i_c, i_r\right)$$

Equation (A-6) describes the resistode characteristics of the device. For the controller characteristics, it is obvious that

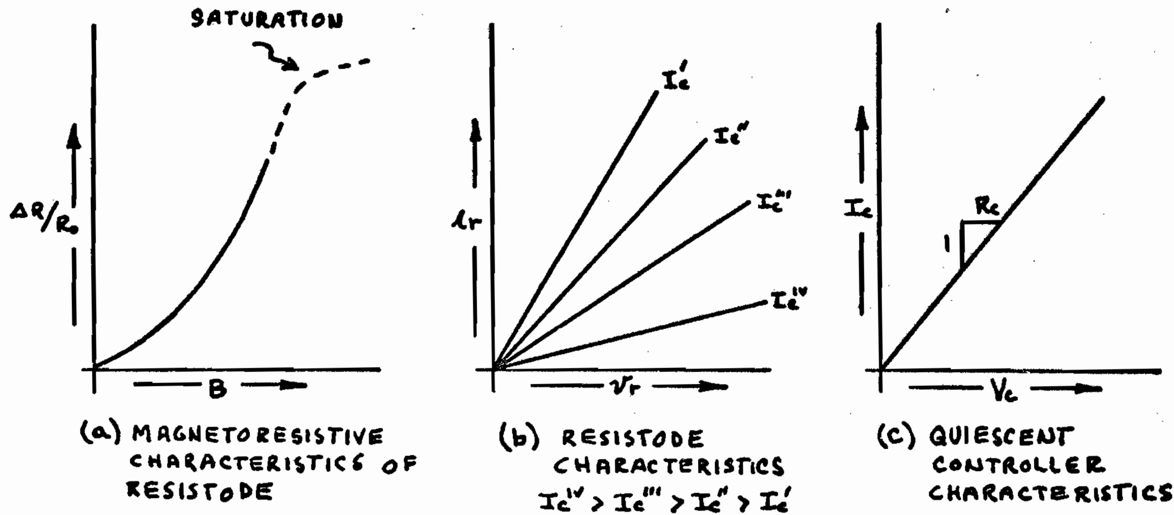
$$(A\text{-}7)\quad v_c = i_c R_c + L \frac{d i_c}{dt} = f\left(i_c, i_r\right)$$

Under quiescent conditions, (A-6) and (A-7) reduce to

$$(A\text{-}8)\quad \begin{aligned} V_r &= I_r R_o \left(1 + \gamma I_c^2\right) \\ V_c &= I_c R_c \end{aligned}$$

Directly from these equations and the equivalent circuit in

Fig. A-3, a load line analysis could be made or a set of characteristic curves could be constructed.



(a) MAGNETORESISTIVE CHARACTERISTICS OF RESISTODE

(b) RESISTODE CHARACTERISTICS $I_c^{iv} > I_c''' > I_c'' > I_c'$

(c) QUIESCENT CONTROLLER CHARACTERISTICS

MCD CHARACTERISTIC CURVES

Fig. A-4

With the help of equations (A-6), (A-7), and (A-8), we are now ready to commence a small signal analysis of the MCD. We will follow the usual procedure of setting up a four terminal network with the consideration that variations about the quiescent point are linear.

$$\text{(A-9)} \qquad \Delta v_c = \frac{\partial v_c}{\partial i_c} \Delta i_c + \frac{\partial v_c}{\partial i_r} \Delta i_r$$

$$\Delta v_r = \frac{\partial v_r}{\partial i_c} \Delta i_c + \frac{\partial v_r}{\partial i_r} \Delta i_r$$

The solutions to the partial derivatives are

$$\text{(A-10)} \quad \frac{\partial v_c}{\partial i_c} = Z_i = R_c + L\frac{d}{dt} \quad \text{(INPUT IMPEDANCE)}$$

$$\frac{\partial v_c}{\partial i_r} = Z_r = 0 \quad \text{(REVERSE IMPEDANCE)}$$

$$\frac{\partial v_r}{\partial i_c} = Z_f = 2\gamma R_o I_r I_c \quad \text{(FORWARD IMPEDANCE)}$$

$$\frac{\partial v_r}{\partial i_r} = Z_o = R_o(1+\gamma I_c^2) \quad \text{(OUTPUT IMPEDANCE)}$$

Let us define the transforms of the incremental voltages and currents as $\mathcal{L}(\Delta v_c) \equiv \mathcal{V}_c$, etc., using script notation to denote transformation. Also, the impedances shall be transformed. Equation (A-9) becomes

$$\text{(A-11)} \quad \mathcal{V}_c = \mathcal{I}_c \mathcal{Z}_i + \mathcal{I}_r \mathcal{Z}_r$$

$$\mathcal{V}_r = \mathcal{I}_c \mathcal{Z}_f + \mathcal{I}_r \mathcal{Z}_o$$

These are the equations of the following T networks.

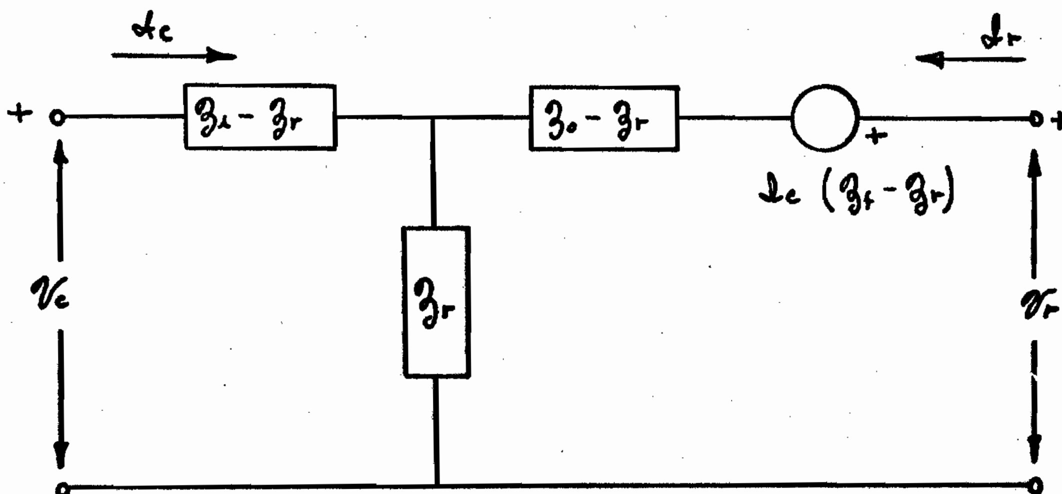
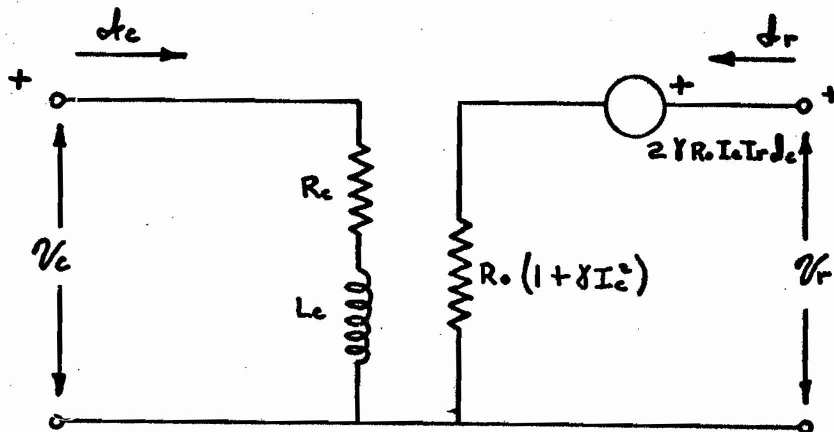


Fig. A-5(a) T NETWORK GIVEN BY EQUATIONS (A-11)



(b) EQUIVALENT CIRCUIT OF MCD RESULTING UPON SUBSTITUTION OF THE PARTIAL DERIVATIVES

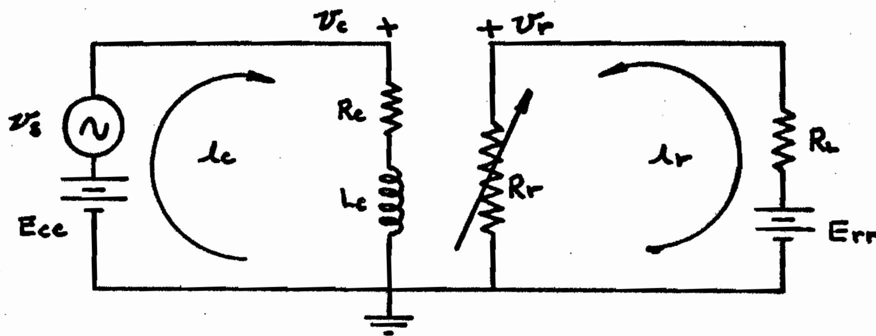
SMALL SIGNAL EQUIVALENT CIRCUITS

Fig. A-5

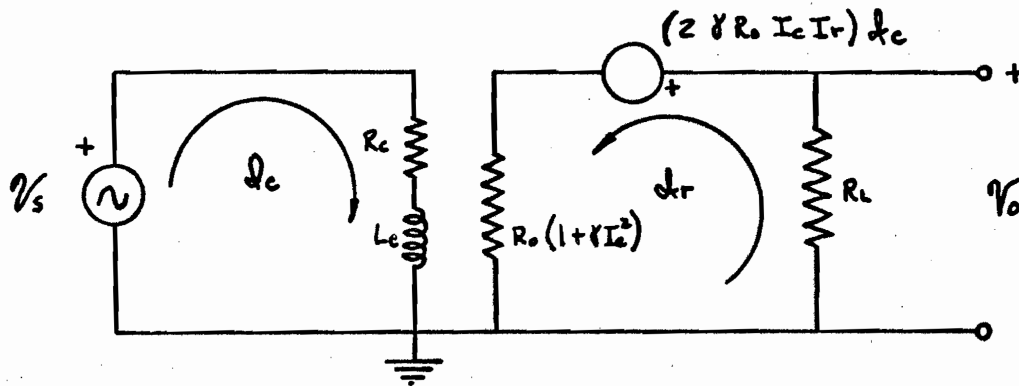
The equivalent circuit shown in Fig. A-5(b) can be used to analyze designs utilizing this device when small signals are used. It is interesting to note that the input circuit is completely independent of the output circuit. There is a distinct disadvantage, however, in having inductance in the input circuit as this limits the frequency response of the device to high frequency signals. On the other hand, it is an aid in that it helps limit the small signal current in the input circuit.

We are now in a position to determine the response characteristics of the MCD. Let us choose a simple application -- that of supplying small signal power to a resistive load.

The appropriate circuit diagrams are shown below.



(a) OVERALL EQUIVALENT CIRCUIT



(b) SMALL SIGNAL EQUIVALENT CIRCUIT

RESISTANCE LOADED MCD AMPLIFIER

Fig. A-6

The quiescent conditions are, from Fig. A-6(a),

$$I_c = E_{cc}/R_c \qquad \text{(A-12)}$$

$$I_r = \frac{E_{rr}}{R_L + R_o\left(1 + \gamma \frac{E_{cc}^2}{R_c^2}\right)}$$

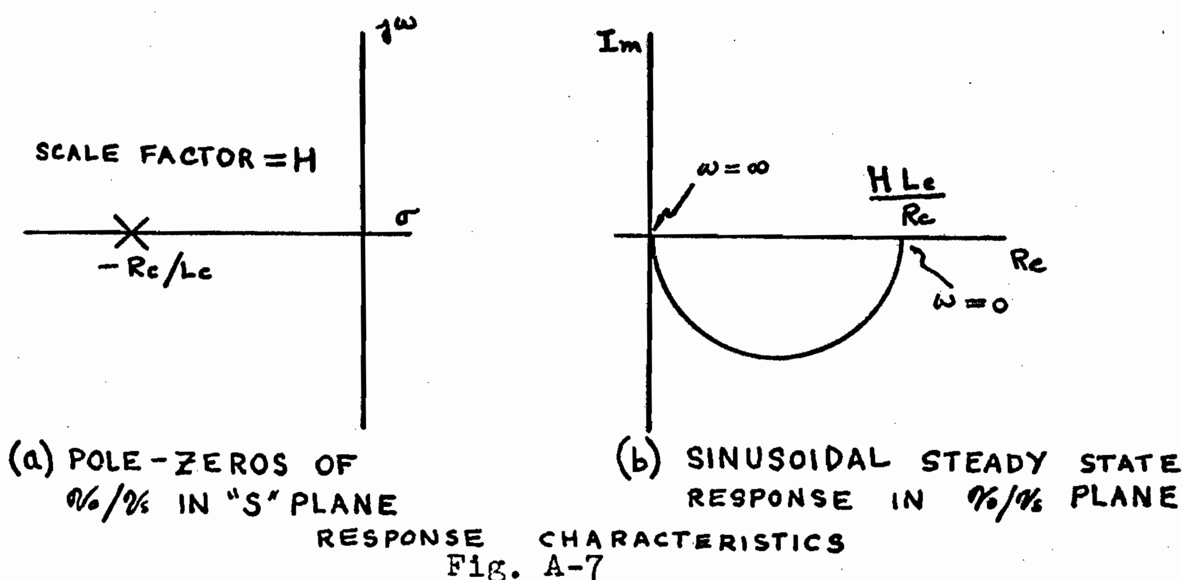
By applying straightforward linear circuit analysis to the small signal equivalent circuit shown in Fig. A-6(b), it can be easily demonstrated that the following voltage transfer characteristic results.

$$\text{(A-13)} \quad \frac{v_o}{v_s} = \left\{ \frac{2\gamma R_o R_L I_c I_r}{L_c\left[R_L + R_o\left(1+\gamma I_c^2\right)\right]} \right\} \frac{1}{s + R_c/L_c}$$

$$= \left\{ \frac{2\gamma R_o R_L E_{cc} \; E_{rr}}{L_c R_c \left[R_L + R_o\left(1+\gamma \frac{E_{cc}^2}{R_c^2}\right)\right]^2} \right\} \frac{1}{s + R_c/L_c} = \frac{H}{s + R_c/L_c}$$

Equation (A-13) can be seen to give a frequency response similar to a simple series R-L voltage divider circuit with the exception that H, the scale factor, will permit the gain to be greater than unity within a certain low frequency range. The criteria that any gain occurs at all is easily seen to be $H > R_c/L_c$ and that gain occur in the linear response range, that is, below the half power point, would then be $H > \sqrt{2}\, R_c/L_c$. From (A-13), it can be seen that this criteria finally resolves to

$$\text{(A-14)} \quad \frac{2\gamma R_o R_L I_c I_r}{R_L + R_o\left(1+\gamma I_c^2\right)} > \sqrt{2}\, R_c$$



(a) POLE-ZEROS OF V_o/V_s IN "S" PLANE

(b) SINUSOIDAL STEADY STATE RESPONSE IN V_o/V_s PLANE

RESPONSE CHARACTERISTICS

Fig. A-7

Voltage and current gain can be provided by such passive devices as transformers and resonant circuits; therefore, if a device is to be useful in most electronic applications, it must provide power gain as well. In order to arrive at a useful expression for the small signal power gain, we shall first make the following definitions.

(A-15) (a) $\text{POWER GAIN} = PG \equiv P_{out}/P_{in}$

(b) $P_{in} \equiv i_e i_e^* R_e$, $P_{out} \equiv i_r i_r^* R_L$

(c) $s \equiv \sigma + j\omega$

The first two definitions are obvious. The last definition follows from the definition of the Laplacian operator. Using these definitions along with the usual algebraic manipulations on the circuit equations, it can be shown that the following expressions result.

(A-16)
$$P_{in} = \frac{\mathcal{V}_s \mathcal{V}_s^* R_c}{(R_c + \sigma L_c)^2 + (\omega L_c)^2}$$

$$P_{out} = \frac{(2\gamma R_o I_c I_r)^2 \mathcal{V}_s \mathcal{V}_s^* R_L}{\left[R_o(1+\gamma I_c^2) + R_L\right]^2 \left[(R_c + \sigma L_c)^2 + (\omega L_c)^2\right]}$$

(A-17)
$$PG = \frac{(2\gamma R_o I_c I_r)^2 R_L}{\left[R_o(1+\gamma I_c^2) + R_L\right]^2 R_c}$$

Under conditions of maximum power transfer, when
$R_L = R_o(1+\gamma I_c^2)$,

(A-18)
$$(PG)_{max} = \frac{(\gamma I_c I_r)^2 R_o}{(1+\gamma I_c^2) R_L} = \frac{\gamma^2 E_{cc}^2 E_{rr}^2 R_c^2}{4(R_c^2 + \gamma E_{cc}^2)^3 R_o}$$

Equation (A-17) demonstrates a most unusual property of the MCD; that is, under the conditions of our initial assumptions, the power gain is independent of the controller coil inductance and the frequency and waveshape of the input signal.

Equations (A-14) and (A-17) would be the basis of design considerations for an MCD to be used to supply power to a resistive load. Such a device would have several distinct advantages over another common semiconductor device -- the transistor. Vacuum tube type isolation is provided between the input and output circuits, a feature which the transistor does

not possess. Also, the manufacture of the resistode semiconductor is less complex than the growth of multiple impurity crystals needed for transistor fabrication. The main disadvantage of an MCD amplifier lies in the use of a magnetic circuit. There will ultimately be iron losses as the input frequency increases and fringe losses will always be present in varying degrees. However, the advent of microminaturization techniques may help in reducing these losses encountered whenever magnetic circuits are involved.

II. Proposed Operating Model of an MCD

There exist three major considerations in the eventual success of any magnetoresistive control device.

(1) The resistode material must exhibit both a marked magnetoresistance effect and a relatively high resistivity.

(2) The geometry of the resistode should be that which yields an optimum magnetoresistance effect.

(3) There must exist an efficient mechanism for producing a relatively high magnetic field in the resistode for a given amount of controller current.

The first requirement is met by InSb, although its resistivity is not as high as that of some other semiconductors such as Ge and Si. The optimum geometry of the resistode is a configuration known as a Corbino disc.[1,4] However, in our model, we shall use the next best which is an element of rectangular shape. The third condition is usually the most difficult to

meet. The obvious solution is to use a ferromagnetic path of low reluctance and a small gap in which to place the resistode. It is seen from equation (A-3) that because of the reciprocal relationship between the flux density and the gap width, the gap becomes a very important factor in determining the flux density developed in the resistode. Therefore, it would behoove us to find a method in which the gap could be made extremely small. Such a method is provided by the art of thin film technology. By vacuum depositing the device in layers, resistode gap widths can be controlled to a great degree. The proposed device will be constructed by this technique. Reference should now be made to Fig. A-8.

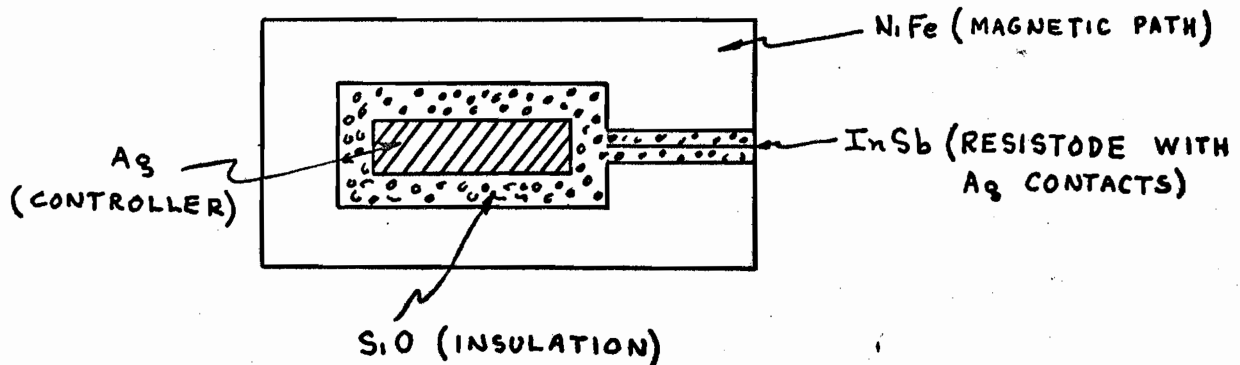
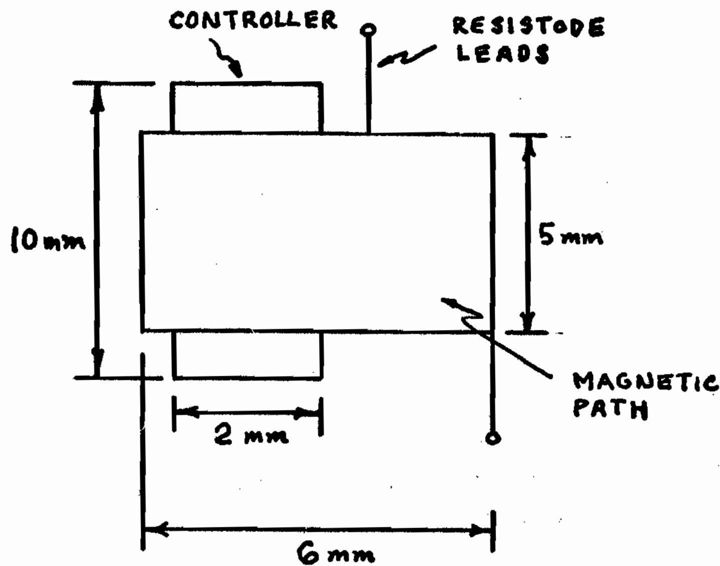
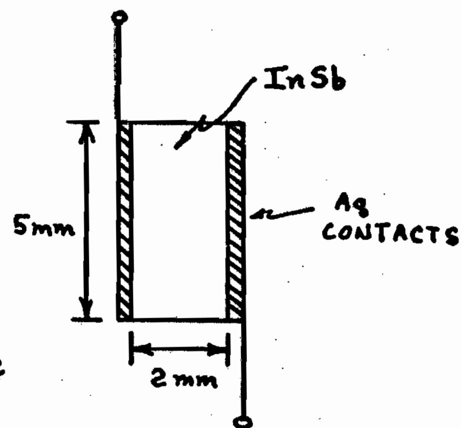


Fig. A-8(a) SIDE VIEW, THIN FILM MCD
(NOT TO SCALE)





(b) TOP VIEW, THIN FILM MCD

(c) TOP VIEW, RESISTODE DETAIL

PICTORIAL DIAGRAM OF PROPOSED MCD (NOT TO SCALE)

Fig. A-8

In the figure A-8(a), all of the vertical dimensions are extremely exaggerated. The thinnest part of the magnetic path will be 1 mm and the silver controller will be 10^{-3} mm or 10000 A in thickness. The resistode gap will be made 5000 A thick with the resistode itself at 3000 A, protected by layers of silicon monoxide insulation of thickness 100 A apiece.

Because of the thinness of the iron magnetic path, equation (A-3) cannot be assumed as the reluctance of certain parts of the path is appreciable. Instead, we must use equation (A-2) directly. After some manipulation, we can obtain the following expressions.

(A-19) $$B = \frac{\mu_r N I_c}{\beta l_r}, \quad \text{where } \beta = \frac{l_c A_r \mu_r}{l_r A_c \mu_c} + 1$$

(A-20) $$\gamma \equiv \alpha \left(\frac{\mu_r N}{\beta l_r} \right)^2$$

β may be thought of as a magnetic circuit quality factor which, when large, is also nonlinear. Assuming a permeability at the quiescent point of 100,000 and the stated dimensions, we obtain β equal to 1.22 as compared to nearly unity for a very good magnetic path. The value for the constant α will be found empirically by using values given by Welker and Weiss for the magnetoresistance of InSb. Using their values of $\frac{\Delta R}{R_o}$ = 15 at B = 8700 gausses along with equation (A-1), one obtains a value of 19.8 $(\text{weber/m}^2)^{-2}$ for α. Substitution of these values of α and β into (A-20) yields 84.5 amp^{-2} for γ.

Using $6 \cdot 10^2$ $(\text{ohm} \cdot \text{cm})^{-1}$[1] as a nominal value for the conductivity of InSb and the geometry shown in Fig. A-8(c), the resistode resistance Ro, neglecting lead resistance, can be determined to be 22.2 ohms. In a similar manner, using $1.63 \cdot 10^{-6}$ ohm · cm[10] as the resistivity of silver, the controller resistance Rc becomes 0.0815 ohm.

We are now in a position to compute the power and voltage gain of the device. Let us choose our quiescent point with Ir equal to 50 milliamperes and Ic equal to 100 milliamperes and a load resistance such that the optimum power transfer condition

exists. For voltage gain considerations, we will neglect the effects of controller inductance. The value for maximum power gain can be found from equation (A-18) and is determined as 26.3 or 14.2 db. The voltage gain, as obtained from equation (A-13), is 62.4 or 36 db. Therefore, this design is capable of both power and voltage amplification for small signals.

This example was not meant to be the best possible design, but merely to indicate the feasibility of producing practical magnetoresistance devices through the use of thin film technology. For instance, the resistode current in the model discussed herein is much too high to be used without externally cooling the device; however, improvements in the detail of the design would overcome this difficulty. A patent application concerning thin film magnetoresistive control devices has been filed by the International Business Machines Corporation on behalf of the author.

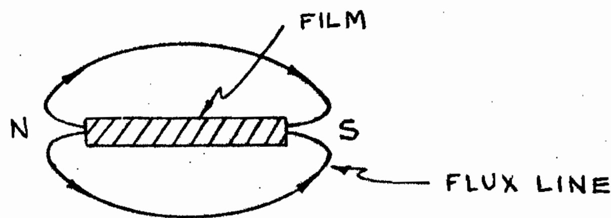
APPENDIX B

NON-DESTRUCTIVE SENSING OF MAGNETIC STORAGE IN DIGITAL COMPUTERS

In the past, the usual means of sampling the contents of magnetic memory elements in modern digital computers has been to attempt changing the magnetic polarization of the element through the mechanism of an externally applied unidirectional field. If the polarization of the element was in the same direction as the applied field, no change in its state would take place. If the polarization of the memory device happened to be antiparallel to the direction of the field, the element would "switch" thus producing a time variant flux which would in turn produce a small voltage in a sense winding linking the field of the device. In this manner, the binary intelligence stored in the element could be sampled; that is, the lack of a voltage induced in the sense winding would represent a binary zero and the presence of a sense voltage would indicate a binary one. The principle drawback to this method is the fact that the information stored in a memory register would be destroyed in the process. This required that the information be temporarily stored in an auxiliary register (usually flip-flops) and "re-written" back into the proper memory storage register at a later time in the instruction cycle. As a result, time is wasted by referring to the magnetic memory twice in an instruction cycle and error occurence is made more probable due to the increased frequency of memory disturbance.

Most of the digital computers produced today use ferrite cores as the basic high speed magnetic memory storage element. However, there has recently developed the possibility of using thin ferromagnetic films with induced uniaxial anisotropy as computer storage elements. These films are produced in a high vacuum environment by evaporation of a NiFe alloy onto a polished glass surface. While the evaporation process is under way, a large magnetic field, perhaps 100 oersteds, is applied orthogonal to the direction of travel of the NiFe molecules. In this way, the film crystallizes with the magnetic moment of its molecules oriented in a single direction. Also, because of the thinness of the film, thermodynamic considerations demand that the film be composed of nearly one entire domain. These films have extremely fast switching speeds (less than 20 millimicroseconds) plus the fact that the flux path travels outside the film on its closing loop, a fact that will be useful in attempting to devise a means of sensing its magnetic state.

Figure B-1 is an exaggerated side view of a magnetic film.

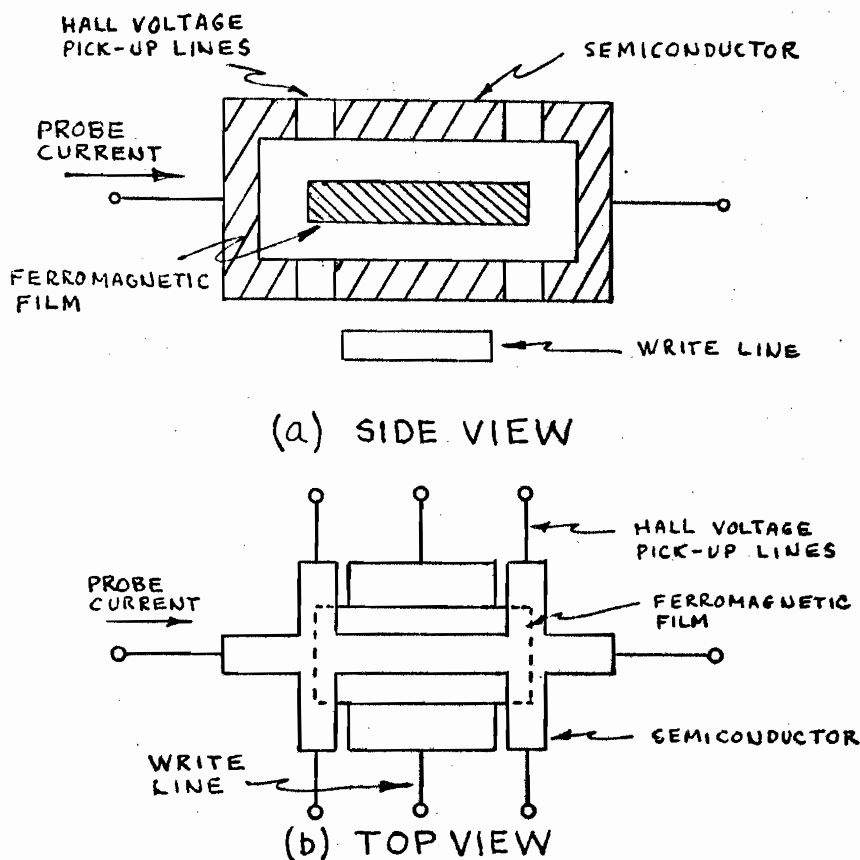


CROSS-SECTIONAL VIEW OF A THIN FERROMAGNETIC FILM (NOT TO SCALE)

Fig. B-1

It is seen that the film has roughly the appearance of a bar magnet along with similar flux paths. Thus a static magnetic field exists in the space outside the film whose direction indicates the binary state of the film.

It can now be quite easily seen that by properly placing a Hall probe in this field, one could effect a means of non-destructively sensing the state of the film through the sign of the Hall voltage produced. The only means of placing the Hall probe would be through vacuum deposition techniques; and, since it would be desirable to have the flux transverse to the plane of the probe to effect maximum output, the optimum position would be near one of the "poles" of the film.



SIMPLIFIED DIAGRAM OF A THIN FILM MEMORY CELL UTILIZING HALL EFFECT SENSING
Fig. B-2

Figure B-2 depicts a complete thin film memory cell, one of many thousand which would exist in the high speed memory of a digital computer. The structure of the cell would consist of multiple layers of vacuum deposited films. For the sake of simplicity, the insulation layers have not been shown in the diagrams. Let us now trace the function of each of the layers individually. The write line, usually a film of silver or another good conductor, is for the purpose of storing information in the NiFe film. Depending upon the direction of current flow, it can be seen that the film will be polarized accordingly. The semiconductor used to sense the state of the device completely envelops the ferromagnetic film. There exists a very good reason for this construction. The probe current, or read current, supplied to the Hall probe is of the same magnitude as the write current; therefore, if the probe were only on one side of the film, a strong field would be applied transverse to the direction of polarization of the NiFe film thus reducing the net flux passing through the area being sensed and thereby reducing the magnitude of the Hall voltage produced. However, the field produced by probe current flowing in the same direction in two layers equidistant from the NiFe film can be seen to nearly cancel out leaving a net transverse field close to zero. The two layer probe construction also has the added advantage of providing four output terminals which can be added to increase the output signal or used separately to feed other registers. As mentioned before, the sign of the Hall voltage on the four output terminals will indicate the

direction of polarization of the film and thus the intelligence stored within, the direction and magnitude of the probe current and the type impurity semiconductor being known and held constant. The static magnetic flux density near the surface of the film has been found to be of the order of 500 gausses.[3] Short duration probe currents of up to 3 amperes peak have been sent through bismuth films with no ill effects. Under these conditions, a Hall output of the order of 100 millivolts could be expected, more than enough to drive a transistor sense amplifier.

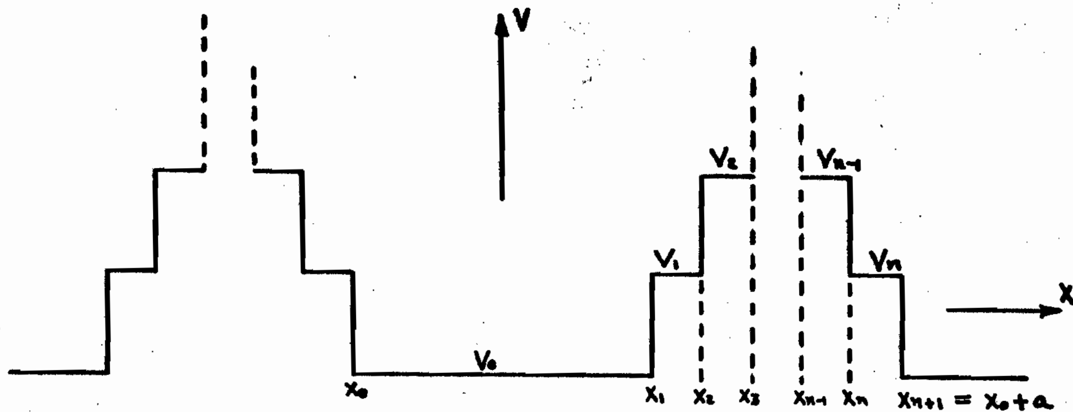
Let us quickly analyze the function of this device during a typical computer cycle. A register will consist of several of these elements with their probe current lines connected in series. The entire memory will be comprised of thousands of registers. The Hall voltage pick-up lines from each element in each register pass through "or" circuits to the same sense amplifiers, their number being the same as the number of elements in a memory register. It is assumed that the intelligence in each register was placed there by the write lines at some previous time. The desired register is selected by the addressing matrix and a "read" pulse is sent along the probe current line of the register. A pulse of Hall voltage, whose polarity indicates the state of each element, is sent from each element in the register to an appropriate sense amplifier and from there to an arithmetic register to be used. All reading operations have occured without destroying the stored information and no subsequent re-writing function is needed.

This method promises to make possible faster and more reliable digital computers, and will become a reality as the art of thin film technology progresses. The design proposed in this appendix is covered in a patent application submitted to the United States Patent Office by the International Business Machines Corporation on behalf of the author.

APPENDIX C

MOTION OF AN ELECTRON IN A ONE-DIMENSIONAL PERIODIC POTENTIAL

In this appendix, we shall consider the motion of an electron in a one dimensional periodic crystal lattice. In the vicinity of the lattice point, or atom, the electron will see a potential field created by the assumed fixed nucleus plus an average field produced by the charge distribution of all other electrons. Such a model is known as the one electron approximation and leads to results very close to the truth. The normal procedure for finding the motion of the electron would be to solve the Schroedinger equation (1-3) when the function $V(x)$ is known. Almost always, however, the functional relation for $V(x)$ is never known and if it is known, it usually leads to a non-linear Schroedinger equation which cannot be solved in closed form. On the other hand, if we can approximate the potential distribution by a staircase function, we will see that a set of linear, homogenous, Schroedinger equations result which can be handled by simple and well known methods. We shall therefore take for our model of the potential function the distribution shown in Fig. C-1.



POTENTIAL DISTRIBUTION IN A 1-DIMENSIONAL LATTICE

Fig. C-1

It is seen that this potential distribution is periodic in "a", the lattice constant; that is,

(C-1) $V(x) = V(x+a)$

The Schroedinger equation in one dimension is,

(C-2) $\psi'' + \frac{2m}{\hbar^2}\left[E - V(x)\right]\psi = 0$, where $\psi'' \equiv \frac{d^2\psi}{dx^2}$

The above model consists of a series of constant potential levels, the number of which, and the magnitude and width of each being selected to yield the closest fit to the actual potential function. The resultant wave motion of an electron moving in this distribution can be described by a set of linear, second degree Schroedinger equations with constant coefficients and right-hand side equal to zero, each holding over a particular domain of definition.

$$(C\text{-}3) \quad \psi_0'' + \frac{2m}{\hbar^2}(E - V_0)\psi_0 = 0 \qquad x_0 < x < x_1$$

$$\psi_1'' + \frac{2m}{\hbar^2}(E - V_1)\psi_1 = 0 \qquad x_1 < x < x_2$$

$$\psi_2'' + \frac{2m}{\hbar^2}(E - V_2)\psi_2 = 0 \qquad x_2 < x < x_3$$

$$\vdots \qquad\qquad\qquad \vdots$$

$$\psi_n'' + \frac{2m}{\hbar^2}(E - V_n)\psi_n = 0 \qquad x_n < x < x_{n+1}$$

$$\psi_{n+1}'' + \frac{2m}{\hbar^2}(E - V_{n+1})\psi_{n+1} = 0 \qquad x_{n+1} < x < x_{n+2}$$

To simplify the notation, we will introduce the following definition:

$$(C\text{-}4) \quad \alpha_j^2 \equiv \frac{2m}{\hbar^2}(E - V_j), \quad j = 0, 1, 2, 3, \cdots n, n+1$$

Because of the periodicity of the lattice, the solution in each region will be given by the Bloch Theorem; that is,

$$(C\text{-}5) \quad \psi_j = e^{ikx}\mu_j(x), \quad \mu_j(x) = \mu_j(x+a), \quad x_j < x < x_{j+1}$$

Taking the Schroedinger equation for the jth potential region as,

$$(C\text{-}6) \quad \psi_j'' + \alpha_j^2\psi_j = 0, \quad x_j < x < x_{j+1}$$

and substituting (C-5), we arrive at another set of linear, homogenous, differential equations.

$$(C\text{-}7) \quad \mu_j'' + 2ik\mu_j' + (\alpha_j^2 - k^2)\mu_j = 0, \quad x_j < x < x_{j+1}$$

The general solution for (C-7) by either Laplace transforms or the D-operator method is,

$$\text{(C-8)} \quad \mu_j = A_j \exp\left[i(\alpha_j - k)x\right] + B_j \exp\left[-i(\alpha_j + k)x\right], \quad x_j < x < x_{j+1}$$

It will again be convenient to simplify the notation by defining the following functions.

$$\text{(C-9)} \quad \gamma_j(x) \equiv \exp\left[i(\alpha_j - k)x\right], \quad \varphi_j(x) \equiv \exp\left[-i(\alpha_j + k)x\right], \quad x_j < x < x_{j+1}$$

Therefore (C-8) becomes,

$$\text{(C-10)} \quad \mu_j = A_j \gamma_j(x) + B_j \varphi_j(x), \quad x_j < x < x_{j+1}$$

Let us now reflect for a moment upon what has been done. Equation (C-10) represents the general solution for the function u in each constant potential region. The overall solution for the wave function ψ will be of the form:

$$\text{(C-11)} \quad \psi = \begin{cases} \psi_0, & x_0 < x < x_1 \\ \psi_1, & x_1 < x < x_2 \\ \vdots & \vdots \\ \psi_n, & x_n < x < x_{n+1} \\ \psi_{n+1}, & x_{n+1} < x < x_{n+2} \end{cases} = e^{ikx} \begin{cases} \mu_0, & x_0 < x < x_1 \\ \mu_1, & x_1 < x < x_2 \\ \vdots & \vdots \\ \mu_n, & x_n < x < x_{n+1} \\ \mu_{n+1}, & x_{n+1} < x < x_{n+2} \end{cases}$$

In other words, the overall wave function ψ will be made up of the segments ψ_j each defined over a constant potential region. However, if the wave function of an electron is to be similar to other wave functions in nature, and there is no reason why it should not be, then the overall wave function ψ must be

piecewise continuous. This means that at the boundary of each potential region, ψ and its derivative must be continuous. From the Bloch Theorem, it can be seen that u has the same continuity requirements as ψ. Therefore, at the jth boundary, we may write,

$$\text{(C-12)} \quad u_{j-1}(x_j) = u_j(x_j), \; j = 1,2,\cdots n, \; u_n(x_0+a) = u_0(x_0)$$
$$u'_{j-1}(x_j) = u'_j(x_j), \; j = 1,2,\cdots n, \; u'_n(x_0+a) = u'_0(x_0)$$

The equations on the right indicate the boundary conditions at the end of the period. Imposing these boundary conditions on (C-10), we may determine the constants A_j, B_j, which will satisfy the boundary conditions and make ψ piecewise smooth and continuous.

$$\text{(C-13)} \quad A_{j-1}\gamma_{j-1}(x_j) + B_{j-1}\varphi_{j-1}(x_j) - A_j\gamma_j(x_j) - B_j\varphi_j(x_j) = 0$$
$$A_{j-1}\gamma'_{j-1}(x_j) + B_{j-1}\varphi'_{j-1}(x_j) - A_j\gamma'_j(x_j) - B_j\varphi'_j(x_j) = 0$$
$$j = 1,2,3,\cdots n$$
$$A_0\gamma_0(x_0) + B_0\varphi_0(x_0) - A_n\gamma_n(x_0+a) - B_n\varphi_n(x_0+a) = 0$$
$$A_0\gamma'_0(x_0) + B_0\varphi'_0(x_0) - A_n\gamma'_n(x_0+a) - B_n\varphi'_n(x_0+a) = 0$$

It can be seen that (C-13) leads to 2(n+1) linear algebraic equations in 2(n+1) unknowns. We can, therefore, determine each coefficient A_j, B_j, by familiar matrix operations and finally arrive at the overall ψ function." We may now recall that there exists a relation between the energy and momentum of a classical particle; viz.,

(C-14) $E = \frac{p^2}{2m}$, $V = 0$

From the de Broglie hypothesis (1-2), $k = \frac{p}{\hbar}$ we have for (C-14),

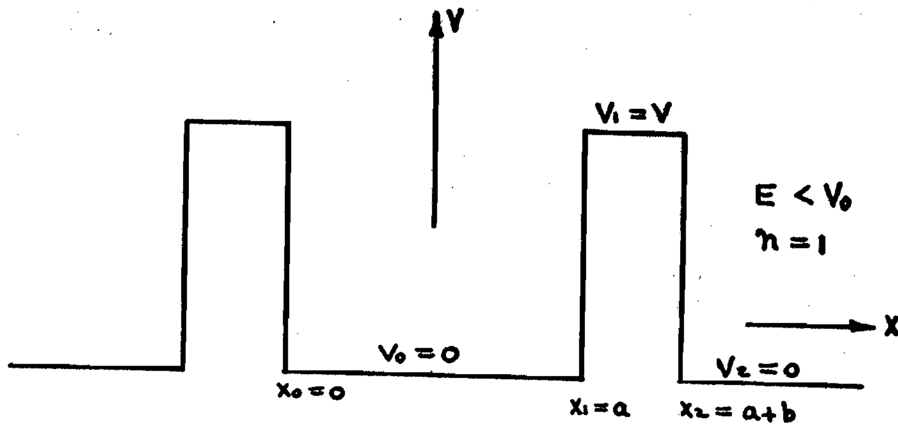
(C-15) $E = \frac{\hbar^2 k^2}{2m}$

We must now remember that our solution for ψ contains the constant α which is proportional to $\sqrt{E}$ and also the constant k, which is the wave vector. However, if we choose values of E and k that satisfy (C-15) and substitute them into our general solution for ψ, we will find that for certain values of energy, the solution for ψ will indicate that there is no probability at all that an electron will possess these values of E and k simultaneously. This means that there are certain energy values that would be forbidden to the electron. Therefore, instead of attempting to solve for all the coefficients A_j, B_j, which is a lengthy process, it would behoove us to find E as a function of k for the purpose of determining the electron energies that are allowed and forbidden. This relation can be found by setting the determinant of the coefficients A_j, B_j, equal to zero as must be done if A_j, B_j, themselves are to be non-zero. Instead of attempting to depict a determinant of order 2(n+1), we will coin a new terminology expressly for this purpose. Let the determinant of the coefficients equal to zero be written

$$
(C\text{-}16) \quad \underset{j=1}{\overset{n}{\mathrm{DET}}} \begin{Bmatrix} A_{j-1}\,\gamma_{j-1}(x_j) + B_{j-1}\,\varphi_{j-1}(x_j) - A_j\,\gamma_j(x_j) - B_j\,\varphi_j(x_j) \\ A_{j-1}\,\gamma'_{j-1}(x_j) + B_{j-1}\,\varphi'_{j-1}(x_j) - A_j\,\gamma'_j(x_j) - B_j\,\varphi'_j(x_j) \\ A_0\,\gamma_0(x_0) + B_0\,\varphi_0(x_0) - A_n\,\gamma_n(x_0+a) - B_n\,\varphi_n(x_0+a) \\ A_0\,\gamma'_0(x_0) + B_0\,\varphi'_0(x_0) - A_n\,\gamma'_n(x_0+a) - B_n\,\varphi'_n(x_0+a) \end{Bmatrix} = 0
$$

where n is one less than the number of constant potential regions. The result of (C-16) is the relationship between the energy of the particle and its momentum. This relationship also leads to the Brillouin zones, which are those regions of momentum space in which the energy is a continuous function of the momentum. Through the use of modern digital computer techniques, (C-16) can be solved for any reasonable number of approximating potential levels and the Brillouin zones computed to a corresponding degree of accuracy. This method should also reduce to closed form in the limit for approximations to potential functions which are known to yield an analytic form of Schroedinger's equation, for example, $V(x) = V_0 \sin x$, the limit being taken as $x_{j+1} - x_j \rightarrow 0$, $V_j \rightarrow V_0 \sin x_j$, $n \rightarrow \infty$.

In order to gain a better appreciation of the energy-momentum relationship, let us examine a special case of our model, the well known Kronig - Penney model shown in Fig. C-2.



KRONIG-PENNEY MODEL
Fig. C-2

Application of (C-16) and the boundary conditions (C-12) to the model shown in Fig. C-2 results in the following determinant.

(C-17)

$$\begin{array}{cccc} A_0 & B_0 & A_1 & B_1 \end{array}$$

$$\begin{vmatrix} e^{i(\alpha_0-k)a} & e^{-i(\alpha_0+k)a} & -e^{i(\alpha_1-k)a} & -e^{-i(\alpha_1+k)a} \\ 1 & 1 & e^{i(\alpha_1-k)(a+b)} & e^{-i(\alpha_1+k)(a+b)} \\ i(\alpha_0-k)e^{i(\alpha_0-k)a} & -i(\alpha_0+k)e^{-i(\alpha_0+k)a} & -i(\alpha_1-k)e^{i(\alpha_1-k)a} & i(\alpha_1+k)e^{-i(\alpha_1+k)a} \\ i(\alpha_0-k) & -i(\alpha_0+k) & -i(\alpha_1-k)e^{i(\alpha_1-k)(a+b)} & i(\alpha_1+k)e^{-i(\alpha_1+k)(a+b)} \end{vmatrix} = 0$$

After considerable manipulation, the solution to (C-17) can

be shown to be:

(C-18) $$-\frac{(\alpha_0^2+\alpha_1^2)}{2\alpha_0\alpha_1} \sin\alpha_1 b \sin\alpha_0 a + \cos\alpha_1 b \cos\alpha_0 a = \cos k(a+b)$$

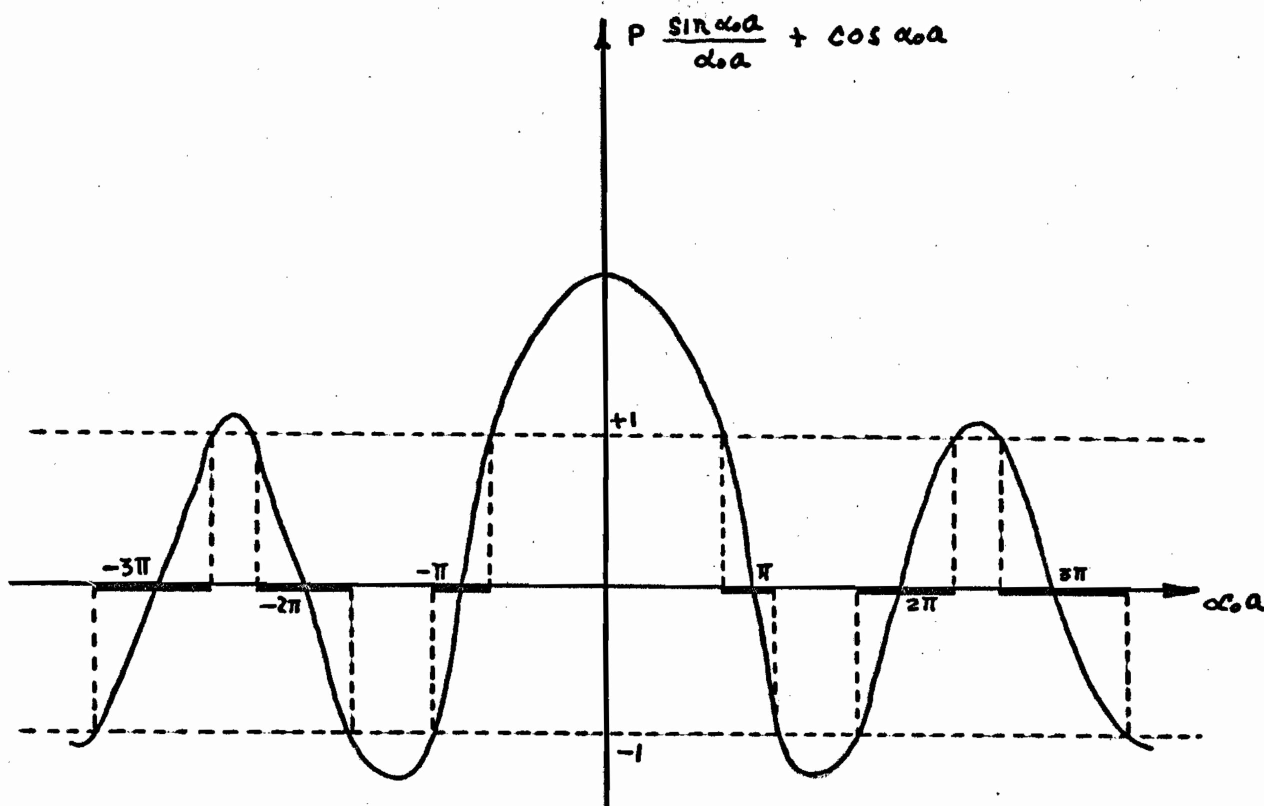
In order to arrive at a more workable form, we shall apply a Dirac limit to (C-18); that is, we will let $V \rightarrow \infty$ and $b \rightarrow 0$, while keeping the product Vb constant. It is well to remember that, as V is greater than E, α_1 is complex and the sinusoids become hyperbolic functions so that a limit is possible. Upon this action, (C-18) resolves to

(C-19) $$\frac{mVb}{\hbar^2\alpha_0} \sin\alpha_0 a + \cos\alpha_0 a = \cos ka$$

Equations (C-18) and (C-19) yield the desired relationship between E and k for the model depicted in Fig. C-2, since both α_0 and α_1 contain the energy E. In order to display the discontinuous and periodic nature of this relationship, let us make the following analysis. We can define the constant $P \equiv \frac{mVba}{\hbar^2}$, thus making (C-19)

(C-20) $$P\frac{\sin\alpha_0 a}{\alpha_0 a} + \cos\alpha_0 a = \cos ka$$

The plot of the left hand side of (C-20) is given below.



PLOT OF EQUATION (C-20) (LEFT HAND SIDE)

Fig. C-3

Since the ordinate is equal to cos ka, it is quite obvious that real solutions exist only between plus and minus one. Since α_0 is proportional to $\sqrt{E}$, the darkened portions of Fig. C-3 are the allowed energy regions. It is further seen that the energy range is continuous over all values of k except at $K = \pm\frac{n\pi}{a}$, $n = 1, 2, 3, \ldots$. From Fig. C-3, it is apparent that the allowed regions, or bands, become wider for increasing energy and thinner for increasing P, the constant which is a measure of the potential barrier area, Vb. Equation (C-20) can be solved numerically for the energy-wave vector

relationship. This relationship is illustrated by Fig. C-4. In Fig. C-4(a), it is clear that the quantum mechanical energy-momentum function approaches the classical Hamiltonian given by equation (C-15). Because the energy-momentum function of the electron is multivalued, as indicated by Fig. C-3, a plot of the relationship over only its principle values is given in Fig. C-4(b).

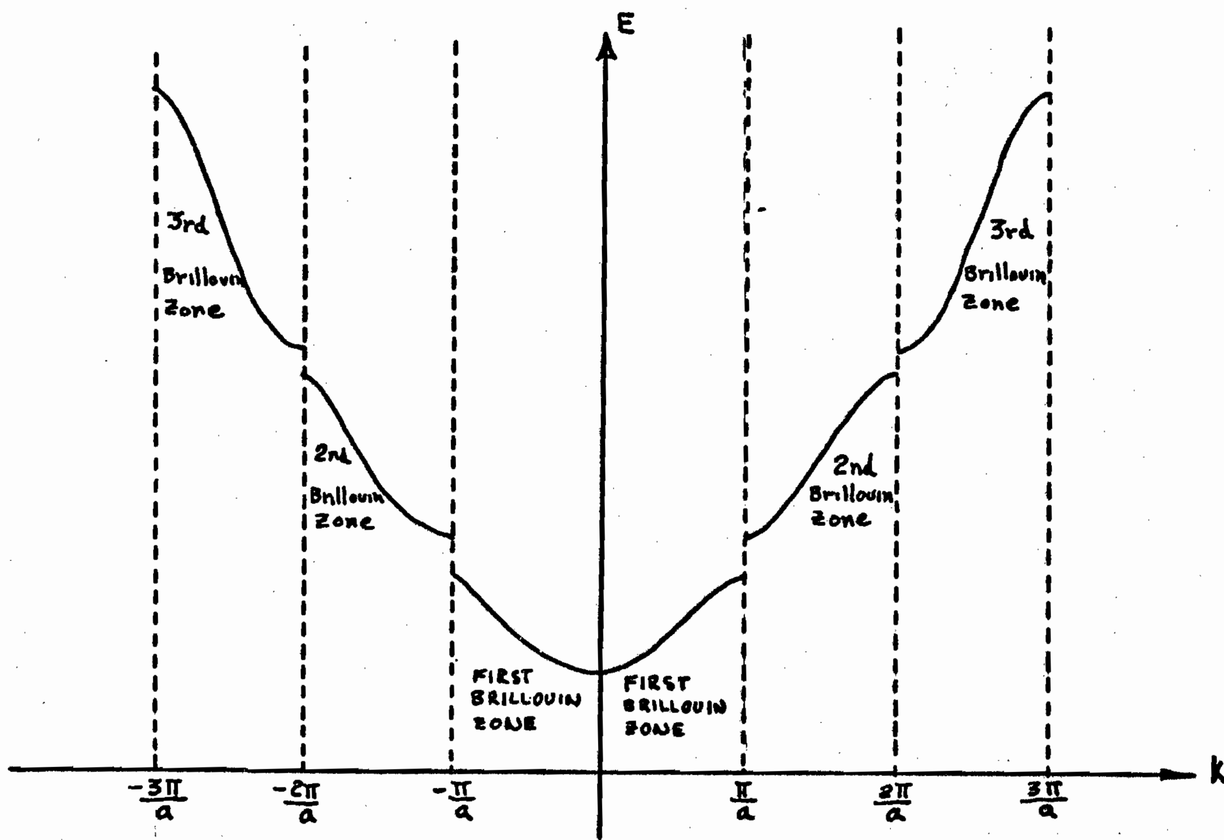
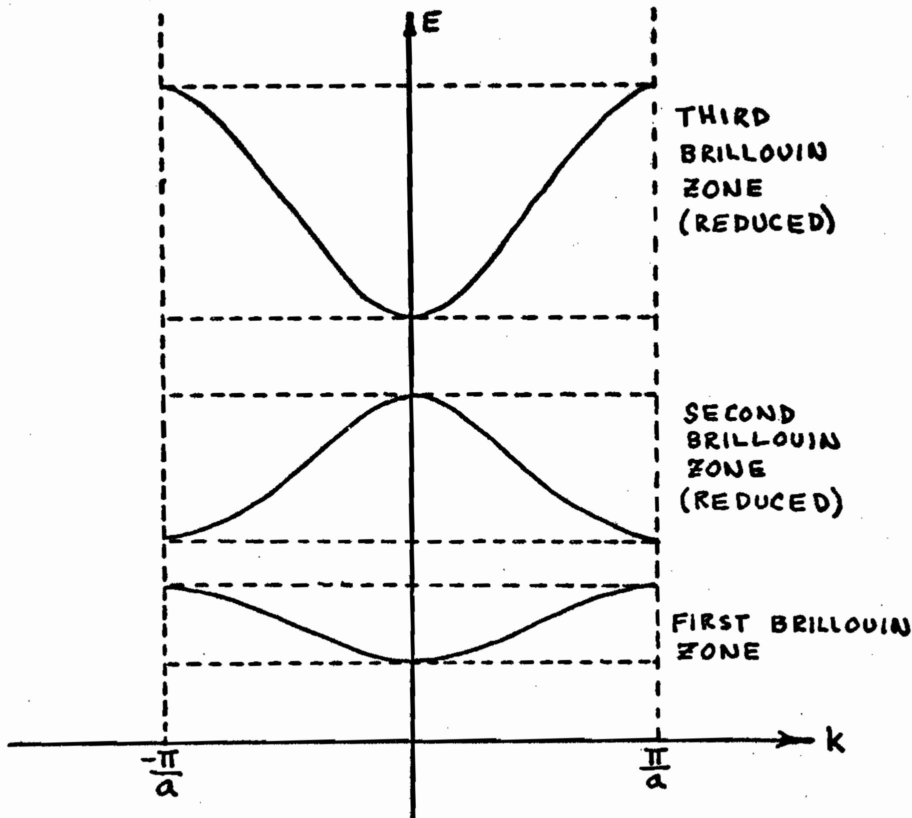


Fig. C-4(a)



E VS. K PLOTTED OVER PRINCIPLE VALUES OF K (REDUCED K-SPACE)

Fig. C-4(b)

Figure C-4 presents a very important relationship in solid state physics. The Brillouin zones, or allowed energy regions in k or momentum space, also define the energy bands. The curves represent the relationship between the energy and momentum of an electron moving in a periodic crystal lattice. In appendix D, several uses shall be made of Fig. C-3 in describing the subtle concept of effective mass and the existence of holes.

It should be pointed out that although the one dimensional

model predicts a great many qualitative electronic processes in solid materials, it does not reveal some of the more unusual phenomenon such as the overlapping of energy bands. To fully describe the behavior of electronic motion in periodic structures, one needs to use a three dimensional model and there are several excellent reviews on this subject in the literature.[12]

APPENDIX D

EFFECTIVE MASS AND THE EXISTENCE OF HOLES

Because of the new relationship between energy and momentum found in Appendix C, one could expect that we will have to change our concept of the mass of a particle moving in a Brillouin zone or an energy band. The derivation of this new concept, called the effective mass, will be useful as we can then show the existence of holes and retain, in a modified form, some of the operations of classical mechanics. For the sake of completeness, we shall derive the effective mass first for the three dimensional case and then adapt it to our one dinensional model.

In reality, the particle is essentially a wave packet moving with a group velocity v. The usual equation representing the group velocity of a wave is,

(D-1) $$\bar{v} = \text{grad}_k \, \omega$$

where grad_k is the gradient in k or momentum space and ω is the angular velocity of the de Broglie wave. It is well to remember at this point that the wave vector k is also inversely proportional to the wavelength of the de Broglie wave, so that an intuitive grasp of equation (D-1) is possible. We now recall that the two basic equations expressing the de Broglie hypothesis are,

(D-2) $$E = h\nu = \hbar \omega$$

where E is the energy of the particle and ν its frequency,

(D-3) $p = h/\lambda = \hbar k$

where p is the momentum of the particle and λ its wavelength. Upon substituting (D-2) into (D-1), we find,

(D-4) $\bar{v} = \frac{1}{\hbar} \text{grad}_K E$

Now, in an external force field, the electron will change its momentum according to Newton's Second Law,

(D-5) $\bar{F} = \frac{d\bar{p}}{dt} = \hbar \frac{d\bar{k}}{dt}$

To the definition of the directional derivative for an infinitesimal change in the energy in the direction of the change in momentum or wave vector; viz.,

(D-6) $dE = \text{grad}_K E \cdot d\bar{k}$

we apply the time derivative and arrive at

(D-7) $\frac{dE}{dt} = \text{grad}_K E \cdot \frac{d\bar{k}}{dt}$

From the definition of acceleration, we have,

(D-8) $\bar{a} = \frac{d\bar{v}}{dt} = \frac{1}{\hbar} \text{grad}_K \frac{dE}{dt}$

Combining (D-8) with (D-7) and (D-5), we finally have

(D-9) $\bar{a} = \frac{1}{\hbar^2} \text{grad}_K \left(\text{grad}_K E \cdot \bar{F} \right)$

We can now draw an analogy between (D-9) and the familiar equation of classical mechanics, $\bar{F} = m\bar{a}$, if we define the following operator as the reciprocal of the effective mass, m*.

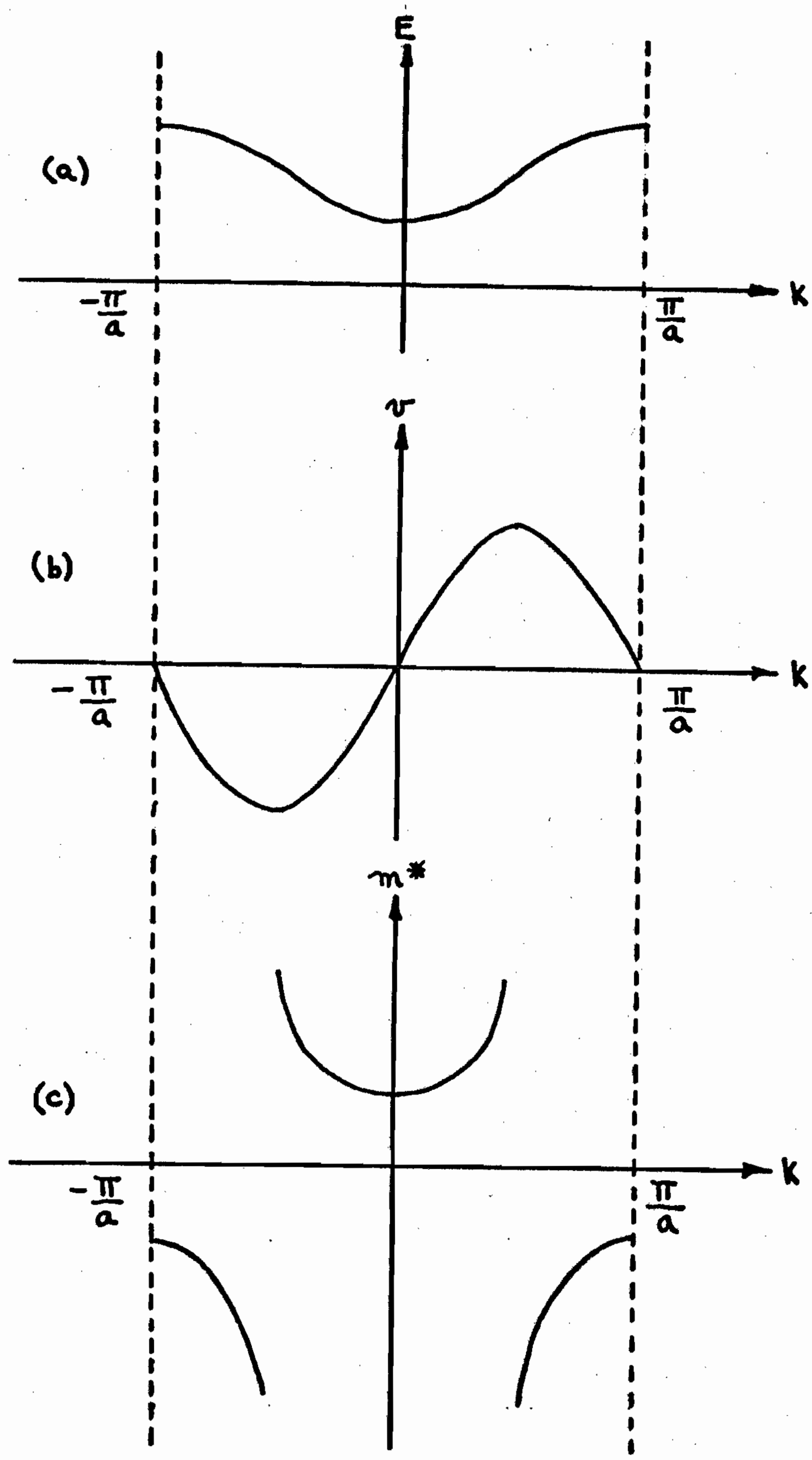
(D-10) $\frac{1}{m^*} = \frac{1}{\hbar^2} \text{grad}_K \text{grad}_K E$

Equation (D-10) is usually called the effective mass tensor because of its directional properties. In one dimensional k space, it is only a matter of quite simple manipulation of the gradient operator to show that

(D-11) $m^* = \hbar^2 / \frac{d^2E}{dK^2}$

In an actual cubic crystal, such as germanium, it is known that the equi-energy surfaces near the top and bottom of the Brillouin zones are spherical. Converting all the quantities to spherical co-ordinates in (D-9), it is again easy to show that the effective mass is of the form (D-11), but this time k is the radius vector in k space.

Let us now examine the properties of particle velocity and mass in a one dimensional crystal using the results of Appendix C.



E, v, m* VERSUS k (REDUCED)

Fig. D-1

One Brillouin zone is plotted in Figure D-1(a). From equations (D-4) and (D-11), Figures D-1(b) and D-1(c), respectively, are deduced. It is seen that the behavior of an electron in a solid is completely different than for a classical particle. For one thing, as the electron increases its energy or moves higher in the energy band, its velocity decreases -- an impossible occurence in classical mechanics. Moreover, and this is really remarkable, the effective mass of the electron becomes infinite at the middle of the band and negative near the top. Since the concept of negative mass is distasteful to most people, the explanation commonly given is that a negatively charged particle with negative mass behaves electrodynamically as a positively charged particle with positive mass. This can be demonstrated by a simple exercise. Consider the following classical equation for the acceleration of a charged particle in an electric field.[13]

$$\bar{a} = \frac{q}{m} \bar{\mathcal{E}} \tag{D-12}$$

For an electron in an energy band, (D-12) becomes modified in accordance with (D-9).

$$\bar{a} = \frac{(-e)\bar{\mathcal{E}}}{m^*} \tag{D-13}$$

(D-13) is obvious because a negatively charged particle is accelerated in a direction opposite to that of the applied field. Suppose the effective mass was negative, then

$$\text{(D-14)} \quad \bar{a} = \frac{(-e)\,\bar{\mathcal{E}}}{-m^*} = \frac{e\,\bar{\mathcal{E}}}{m^*}$$

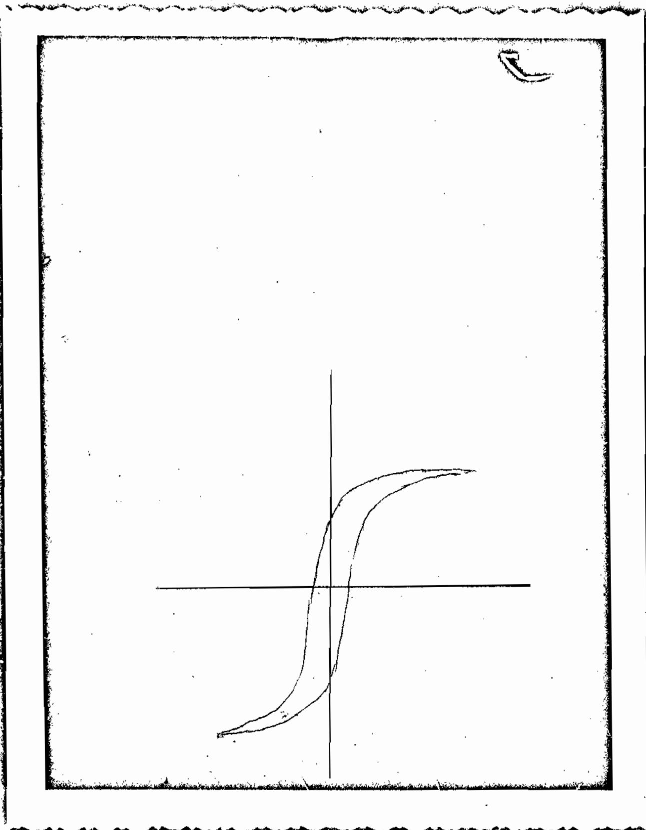
or an electron with negative mass behaves as a hole. Physically, the situation is this: an electron which leaves the upper part of an energy band creates a particle which, in all respects, behaves electrodynamically as a positive charge with positive mass. This fact therefore explains the anomalous Hall effect in which the number of holes available for conduction is greater than the number of electrons. Also it can now be seen why electron conduction takes place in the bottom of the conduction band and hole conduction takes place near the top of the valence band.

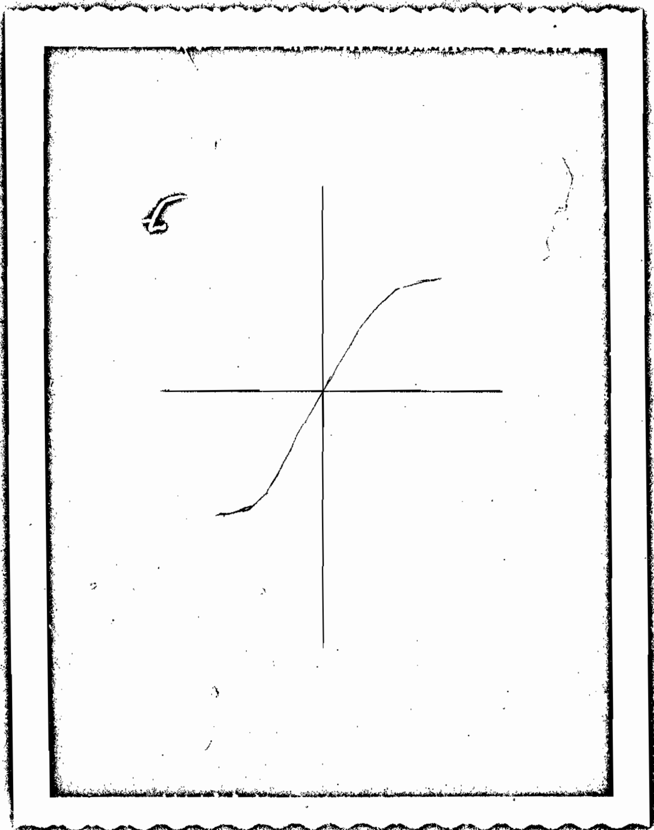
In conclusion, it may be well to answer a question that might arise in the reader's mind. If the effective mass of an electron approaches infinity near the middle of the band, how is it that electrons can escape from the conduction band and indeed from the surface of the solid? The answer is that the model considered in Appendix C was concerned only with electrons having energies less than the maximum height of the potential distribution. The bottom of the conduction band lies just below this level and electrons driven to energies just above the bottom of the band are therefore essentially free and begin to behave as classical particles. However, a hole will always be bound to some degree and its average effective mass will be greater than that of the electrons. Therefore, this is

a partial explanation of why holes are less mobile than electrons and why the Hall effect is usually less for P type materials with the same impurity concentration as the corresponding N type.

APPENDIX E

PHOTOGRAPHS



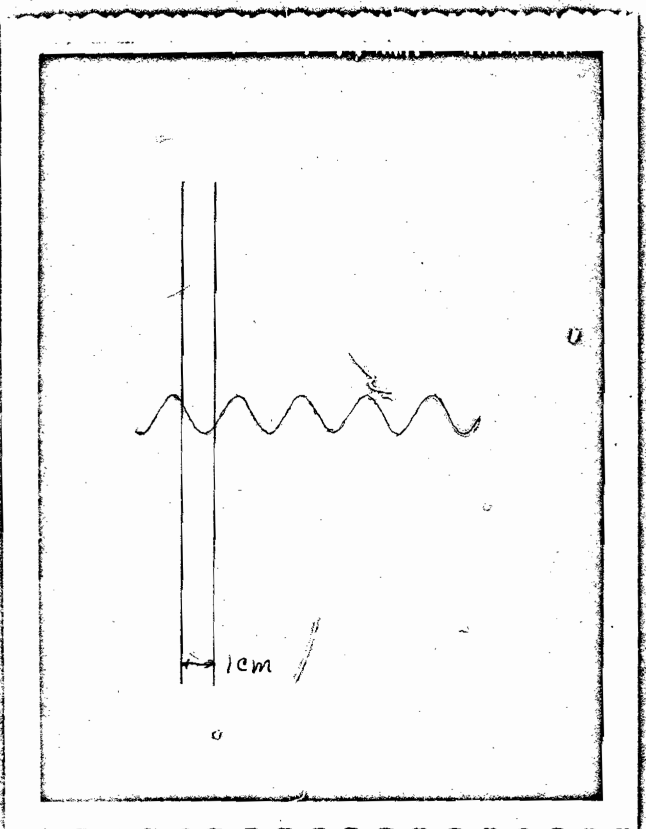


Photograph E-1

Transformer Core Hysteresis Pattern
Vertical Sensitivity = 3240 gauss/cm
Horizontal Sensitivity = 2.7 oersted/cm

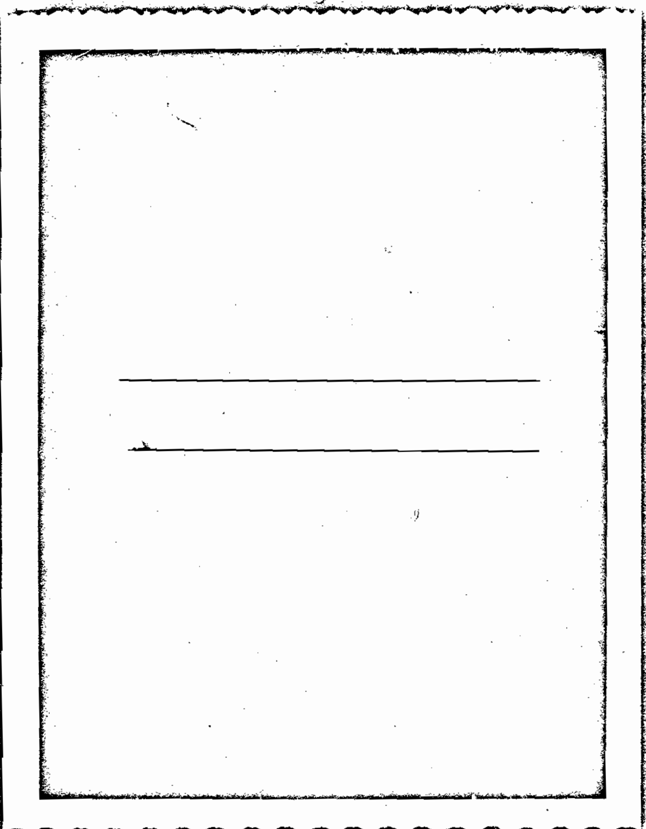
Photograph E-2

Air Gap Characteristic
Saturation Level (vertical) -- 8000 gausses
Saturation Current (horizontal) -- 6 amperes



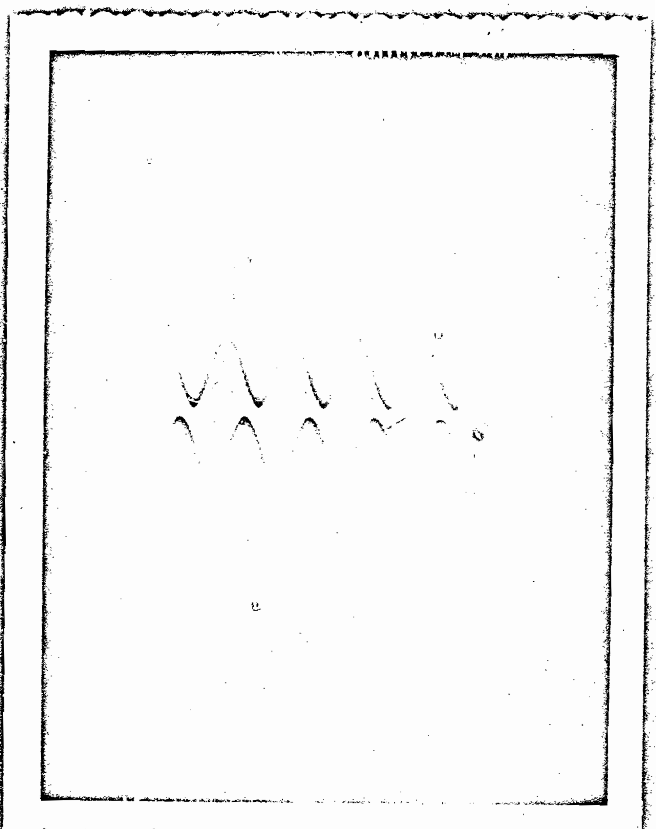
Photograph E-3

60 cycle induced voltage from magnetic field with zero sample current
Vertical Sensitivity -- 300 mv/cm



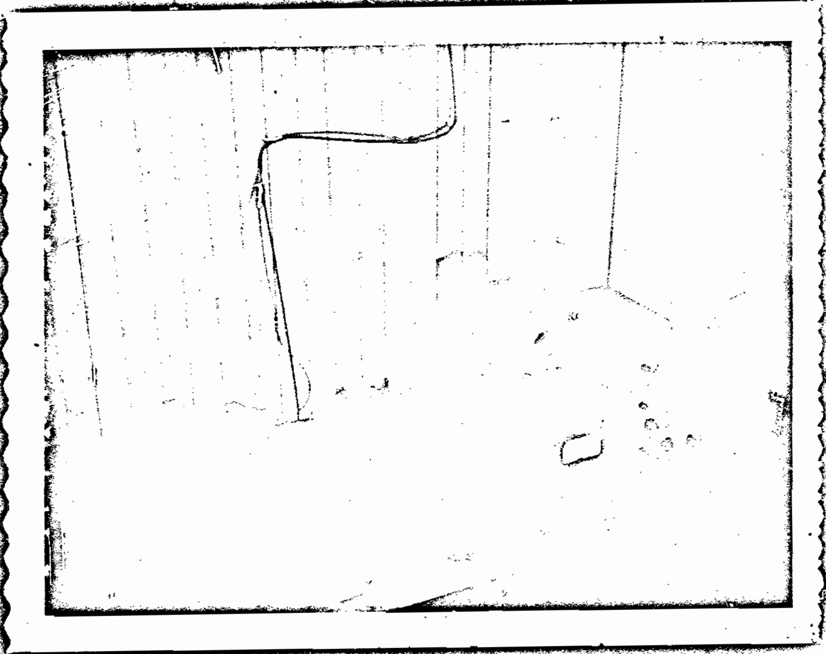
Photograph E-4

IR drop carrier at 2000 cycles.
zero magnetic field
Vertical Sensitivity -- 300 mv/cm



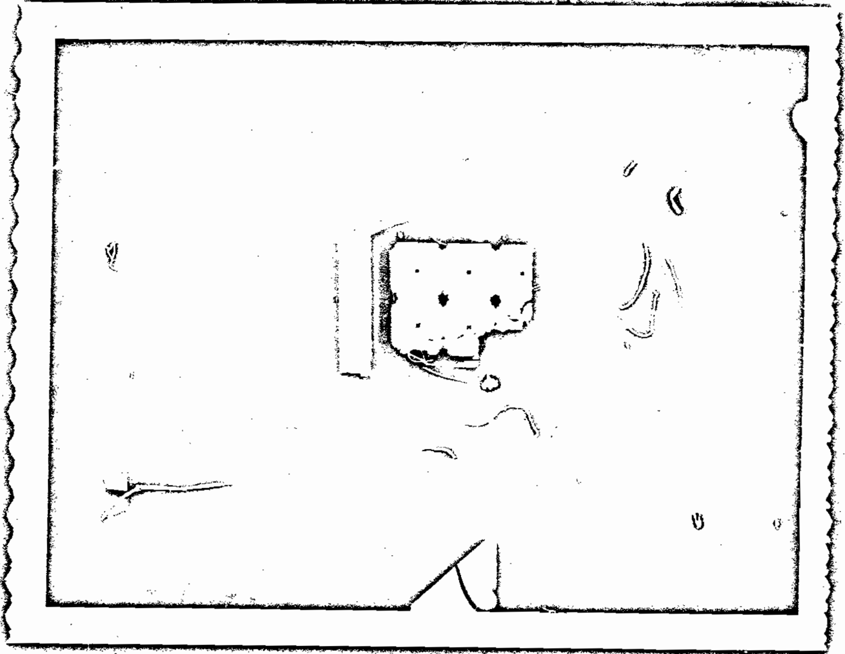
Photograph E-5

2000 cycle sample current modulated by 60 cycle magnetic field
Vertical Sensitivity -- 300 mv/cm



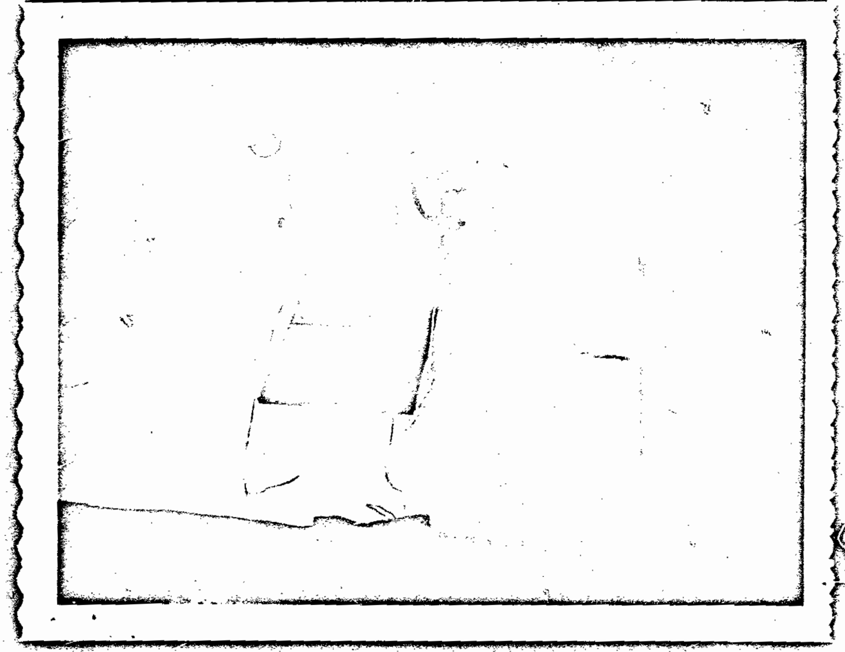
Photograph E-6

Experimental Apparatús for Measurement of the Hall Effect and Magnetoresistance Effect



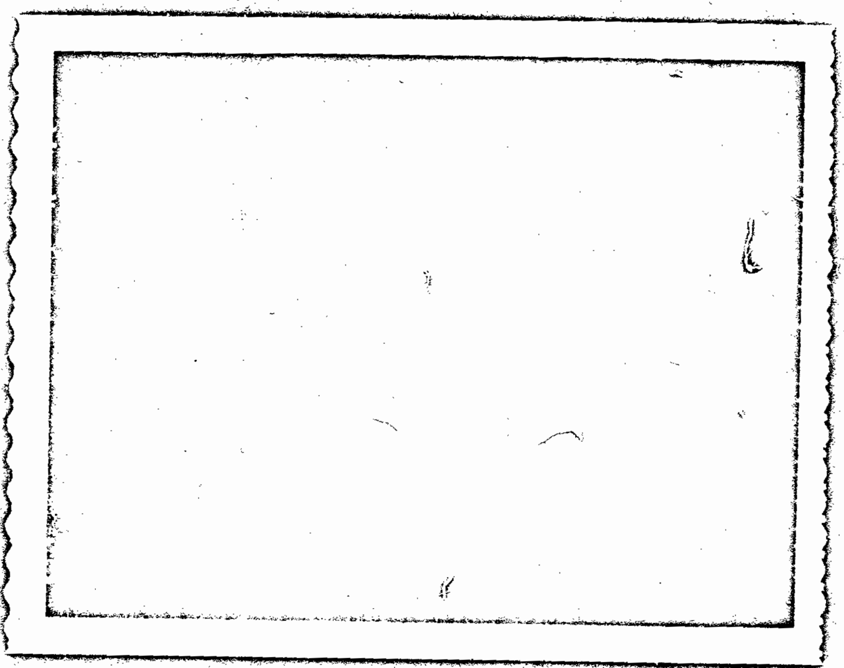
Photograph E-7

Experimental Apparatus for Calibration of Magnetic Circuit



Photograph E-8

Overall View of Magnetic Circuit



Photograph E-9

Close-up of sample assembly in position in air gap

REFERENCES

1. H. Welker and H. Weiss, "Group III - Group V Compounds," Solid State Physics, ed. Seitz and Turnbull, Vol. III (Academic Press, New York, 1956), pp 1-78.

2. H.E. Kronick, "Magnetic Field Plotter for Superconducting Films," IBM Journal of Research and Development, Vol. 2, No. 3, July, 1958, pp. 252-254.

3. IBM Report on Air Force Contract ECPX 0023

4. H. N. Putschi, "Magnetoresistive Amplifiers and Second Harmonic Modulators," Digest of Technical Papers, 1958 Transistor and Solid State Circuits Conference, (IRE-AIEE Publication, Philadelphia, 1958), pp. 60-62.

5. J. D. Knapton, "Electronic Processes in Semiconductors," Master's Thesis, Williams College, Williamstown, Massachusetts (Williamstown, 1956) pp. 19-51.

6. H. R. Reed, T. C. G. Wagner, G. F. Corcoran, Electrical Communications Experiments (John Wiley and Sons, New York, 1956), pp. 78-90.

7. O. Lindberg, "Hall Effect," Proceedings of the IRE, Vol. 40, No. 11, November, 1952, pp. 1414-1419.

8. A. J. Dekker, Solid State Physics (Prentice-Hall, Inc., Englewood Cliffs, 1959), pp. 275-348.

9. A. Coblenz and H. Owens, Transistors: Theory and Applications (McGraw Hill Book Company, Inc., New York, 1955), pp. 215-228.

10. Handbook of Chemistry and Physics (Chemical Rubber Publishing Co., Cleveland, 1958), pp. 2384-2391.

11. J. R. Britton and L. C. Snively, Algebra for College Students (Rinehart and Co., Inc., New York, 1953), pp. 446-448

12. J. R. Reitz, "Methods of the One Electron Theory of Solids," Solid State Physics, ed. Seitz and Turnbull, Vol. I (Academic Press, New York, 1955), pp. 2-92.

13. T. S. Gray, Applied Electronics (John Wiley and Sons, Inc., New York, 1957), pp. 8-11.