THE BELL SYSTEM

Technical Journal

DEVOTED TO THE SCIENTIFIC AND ENGINEERING ASPECTS OF ELECTRICAL COMMUNICATION

VOLUME XXXIII JANUARY 1954 NUMBER 1

DESIGN OF RELAYS

THE BELL SYSTEM TECHNICAL JOURNAL

THE BELL SYSTEM TECHNICAL JOURNAL is published six times a year by the American Telephone and Telegraph Company, 195 Broadway, New York 7, N. Y. Cleo F. Craig, President; S. Whitney Landon, Secretary; John J. Scanlon, Treasurer. Subscriptions are accepted at $3.00 per year. Single copies are 75 cents each. The foreign postage is 65 cents per year or 11 cents per copy. Printed in U. S. A.

THE BELL SYSTEM
TECHNICAL JOURNAL

VOLUME XXXIII JANUARY 1954 NUMBER 1

Design of Relays

Introduction

(Manuscript received September 28, 1953)

The electromechanical relay is the basic building block of modern dial switching systems and also of various automatic control systems and computers.

All of these systems depend on the action of the relay which is simple in its functions, but has to meet complex requirements placed on it by the systems in which it is used. Perhaps the best illustration of this apparent conflict is to note that a relay has simply to close or open electrical contacts when its coil is energized or deenergized. However, these simple functions must be performed equally well by millions of relays, and each of these must continue to perform reliably for millions and in some cases for more than a billion operations during its lifetime. Furthermore, in many cases the relay must function in a few thousandths of a second, or use little electrical power. The reliability required in telephone switching systems would be considered unreasonable in many other types of equipment and can be judged from the fact that a single failure in 5,000,000 operations is considered to be poor performance.

Stated another way, satisfactory operation for the average relay in modern telephone switching systems is less than one failure in forty years of operation. The measurement of such low trouble rates is, in itself, a difficult and challenging problem.

The need for such a high degree of reliability and for the associated requirement of high speed is evident from a few figures relating to

modern cross-bar switching systems. In these, establishing a single telephone connection between two subscribers makes use, for a very short time interval, of approximately 1,000 relay operations with a total of about 7,000 contacts.

As the telephone plant has become more mechanized with its local and toll dial systems, automatic message accounting, automatic trouble recorders, etc., the relays have been required to do more things with less trouble. Accordingly, the steady progress which has been made in the automatism of switching equipment has depended upon improving the performance of relays. To provide this improvement, relay design must be guided by a clear understanding of the physical relations among all aspects of performance and the construction specified by the design.

Some of the recent work in the area of relay design, production, service and measurements is covered by this issue which is devoted entirely to the analysis and measurement of relay performance, and to the economic considerations which govern optimum relay design. It is evident from the typical statistics given, that the successful and economical operation of switching systems and certain other existing automatic control systems demand the best in relay performance.

A. C. KELLER

Relay Measuring Equipment

By H. N. WAGAR

(Manuscript received September 25, 1953)

The wide variety of technical problems encountered in telephone relay design calls for quantitative measurements of many kinds, involving static performance and dynamic performance. The present article describes some of the more important measuring tools for this purpose, as used in Bell Telephone Laboratories. Instrumentation for evaluation of force, flux, displacement, time, or their combination, is described.

INTRODUCTION

Visitors to the switching development areas in the Bell Telephone Laboratories are often surprised at the extensive amount of measuring equipment used to study so simple a device as the telephone relay. This is because, though the relay itself may be simple, the problems requiring study are extremely complex. Because relays are used in such large quantities, their characteristics affect the economy of the telephone switching system. Not only is their manufacturing cost important; their quality also affects the central office cost. For example, an efficient design lowers the central office power plant cost, and faster-acting relays enable fewer common control units to handle more traffic, which further reduces the cost. These and related objectives pose many complex technical problems involving mechanics, magnetics, kinetics, heating, and the like.

The scope of relay analysis may be judged from the other articles in this issue. In every case there is need for measurement. Sometimes this quantitative work is needed to confirm an analytical relation, sometimes to learn more concerning the basic phenomena, and often to characterize the performance of a test model. This article will describe some of the most-used measuring tools for the study of relays.

Several kinds of measurements are required. In the first place, there are force measurements. Each relay must press the desired number of contacts together with a suitable force. To do this, a magnet must be provided which can move the springs, through whatever distance is required. Thus, measurement is needed of force-displacement characteris-

tics—both for contact arrangements, and for magnets. Another field of measurement involves magnetics. Both the mechanical output and the electrical characteristics of the electromagnet are determined by its magnetization relations, requiring the measurement of magnetic flux. Beside such static measurements of mechanical, electrical, and magnetic quantities, corresponding measurements must also be obtained as functions of time in studying relay behavior under the dynamic conditions of actual operation.

The many instruments which furnish such data may be grouped into those for static measurements, and those for dynamic measurements. Some of these tools are described in the following pages, particularly as they relate to force, current, flux, displacement, time, or their combination.

STATIC MEASUREMENTS

The Measurement of Force

For a complete understanding of the operation of relay designs, knowledge of how the forces vary as the magnet air-gap changes, or as the contact members are deflected through their stroke, is of course required. The measurements most needed concern the force versus distance of the contact loads which the magnet must operate; and the force versus air-gap characteristic of the associated magnet. For many years, such measurements were made by a process of hanging weights and setting a micrometer screw, point by point, until complete data were obtained. More recently special instrumentation has made available a pendulum-type tensile tester,[1] and a spring balance device,[2] which have greatly increased the convenience of making these measurements, point-by-point. Today, however, many such measurements may be made still more conveniently by a machine which automatically causes the gaps to vary and plots a curve of force versus distance. Machines of this type have been developed to a high degree of flexibility for tensile testing of materials such as metals, plastics, textiles, or paper.

One such machine, which is extremely convenient for relay measurements, is the Instron tensile testing machine shown in Fig. 1. This machine is manufactured by the Instron Engineering Corporation of Concord, Mass. In conjunction with suitable current supplies to control the behavior of the relay magnet, and with a few special circuits for automatically correcting for flexure of the relay parts, and making other similar adjustments, this machine has proven eminently suitable for observing all "static" force-deflection characteristics of relays. The manner

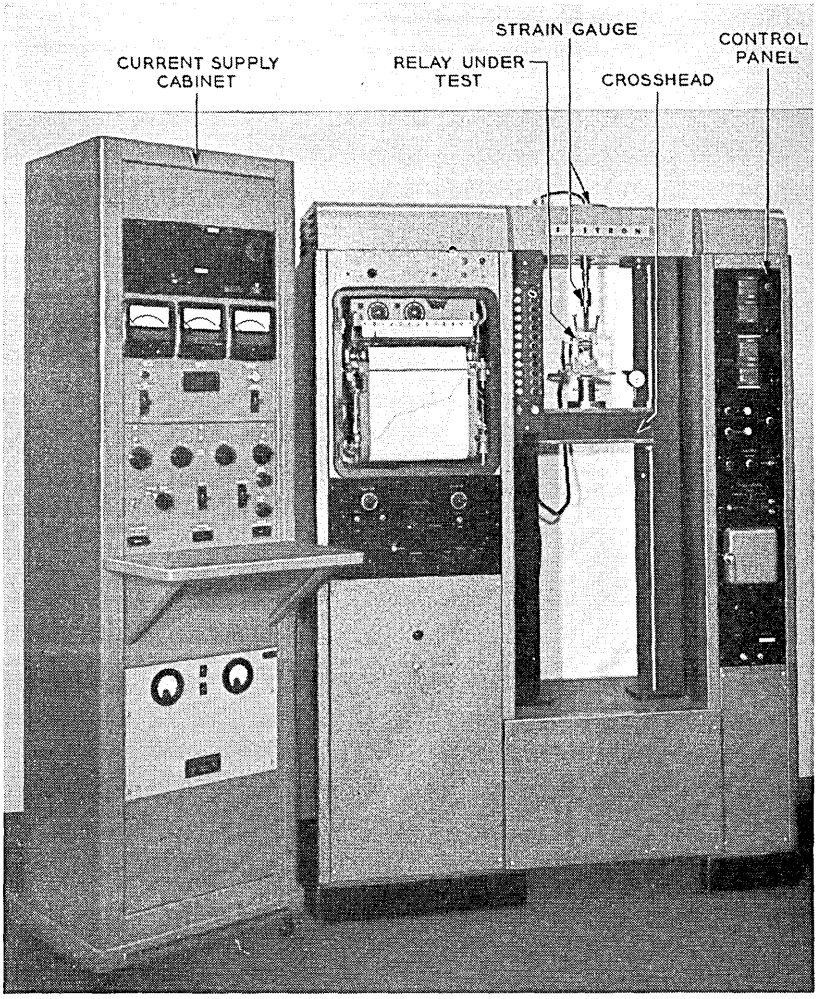


Fig. 1 — Tensile testing machine for force-displacement measurements.

of operation of this machine will be given only in outline as it is described elsewhere.[3] The force system to be measured, such as a relay magnet, is mounted on the crosshead of the machine, whose motion up or down may be controlled from the panel on the right. The other side of the force system being measured is connected to a strain gauge fastened to the top framework of the machine, and as the force on the strain gauge progressively changes, an electrical indication is given to the pen of the Leeds and Northrup X-Y recorder which then moves horizontally in proportion

to the force. The crosshead motion is transmitted electrically to a servo system which drives the paper up or down in proportion to the displacement of the mechanism involved. Thus, during the motion of the crosshead, the pen moves proportionately to force, the paper moves proportionately to distance, and a force-deflection curve is drawn. As a result, the tensile characteristics of mechanisms can be recorded very easily. For example, force variations of relay springs or of relay magnets may be measured, or a combination of the two, as desired.

Some typical measurements obtained on this machine are given in Fig. 2 showing on one chart the manner in which electrical contacts change their force against the magnet and the manner in which the magnet pull varies across the gap. Such charts as these, when completely analyzed, enable the designer to choose the proper magnet for a particular spring requirement. Among the useful features of this machine are the accuracy of recording which readily provides one or two per cent accuracy, and the ease of tracing and retracing a particular measurement. Also readily obtained are exact information on the mechanical hysteresis losses within the mechanism. For example, in curve 1 of Fig. 2, the load characteristic of the contact springs is seen to have two values. The upper one represents the force on the closure stroke while the lower one represents the force on separation. The area between these two curves is the hysteresis loss; friction at any point is readily estimated as one-half the difference between upper and lower force readings. Special information, such as the exact location of first closure or first separation of contacts can also be included, as the machine is provided with a circuit to insert a "pip" on the record when desired.

This instrument has marked advantages because of its accuracy, speed, convenience and wide range of uses. With almost equal ease, mechanisms whose full force variation is 2 grams, and those ranging up to half a ton can be measured. Once a suitable jig has been prepared to properly mount the parts, the Instron machine will furnish a complete set of data in less than an hour, whereas by previous methods it would have required a matter of days.

The time required for each individual curve in the figure is the time allowed for the crosshead to move through the relay stroke. This usually is about 30 to 60 seconds, which adequately simulates the static characteristics of the relay force system; i.e., those force properties that would be measured at a particular point if the armature were held there. These static characteristics are commonly used in relay design to establish performance criteria for ordinary relays. However, in detailed studies of relay dynamics, these static characteristics must be supplemented by

estimates or measurements of the inertia and other forces which also occur. The direct measurement of dynamic performance is described in a later section.

The Measurement of Flux

The mechanical output of a magnet depends on its magnetic quality. As is shown in companion articles in this issue, it is important for the designer to know the values of the closed gap reluctance $\mathcal{R}_0$, the leakage reluctance $\mathcal{R}_L$, and the effective pole face area A. Beside these important

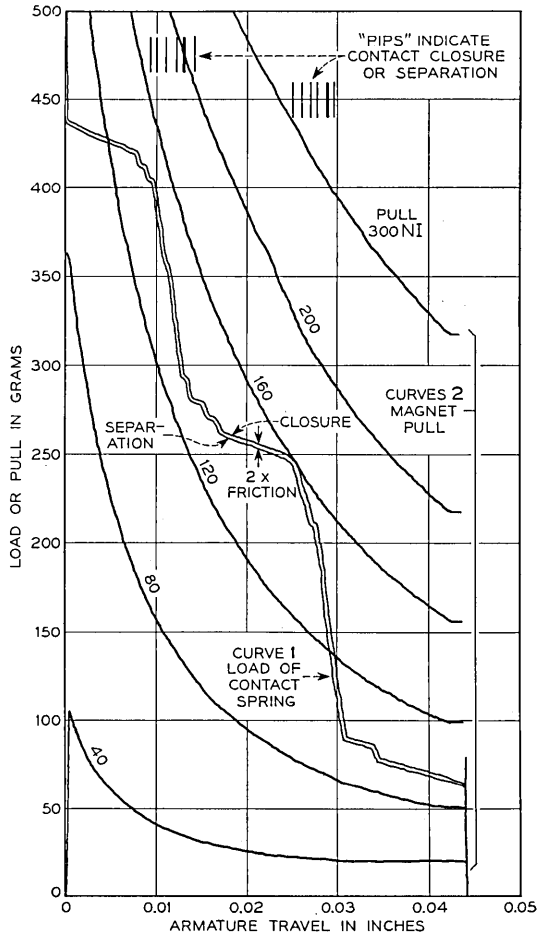


Fig. 2 — Typical recording of relay performance made on Instron machine.

figures of merit, values for saturation flux, φ'', and leakage flux ratios at various points in the magnetic circuit, are commonly needed. For such information, flux as a function of ampere turns must be measured. Perhaps the most effective method is to make determinations of the average flux in the magnet structure at a number of different air gaps; then by analysis determine values for the constants just enumerated. A series of magnetization curves for the relay magnet are required similar to those commonly taken on magnetic "ring samples," except that curves are obtained for several different air gaps. The ballistic galvanometer is the most familiar instrument for this purpose, still used for certain problems. In the Bell Laboratories, however, an extremely versatile recording fluxmeter was developed some ten years ago by P. P. Cioffi for use in research on magnetic materials.[4] One of these instruments is now in constant use for relay measurements because of its accuracy, versatility and the rapidity with which tests can be made. The complete unit is shown in Fig. 3, being made up of a power supply system, galvanometer-integrator unit, and X-Y recorder. Its operation will now be described very briefly.

The magnet to be tested is provided with a search coil. This search coil may be in the form of a winding physically in parallel with the main supply winding, a coil of few turns uniformly spread over the outside of the coil, or of a specially wound coil mounted at a particular point of interest. When current is progressively varied through the main winding, the voltage induced in the search coil by the magnet flux is transmitted to the galvanometer whose associated optical system divides a light beam between two photocells. The resulting photocell voltage unbalance is amplified and coupled back into the search coil circuit through a mutual inductance, so poled as to tend to restore the galvanometer deflection to zero. The feedback current in the mutual circuit is proportional to the flux, and is used to drive the pen on the X-Y recorder. The paper is driven proportionately to the current in the winding, with the result that flux versus current (or ampere turns) is plotted directly. The instrument gives accuracy of $\pm\frac{1}{2}$ per cent over a wide range of fluxes and currents, with provisions made for readily changing the scale to cover different windings, operating voltages or magnet shapes. Measurements are made in a few minutes compared to a considerably greater effort when using the ballistic galvanometer. Auxiliary features are also provided permitting convenient and complete demagnetization of the test magnets between readings. Readings are extremely stable and repeatable. Typical results are shown in Fig. 4 where a series of magnetization curves for a particular test magnet are shown. By

Fig. 3 — Recording fluxmeter.

methods of cross-analysis described in a separate article in this issue,[5] these data may be used to find the figures of merit for the magnet, and other information concerning its performance.

Often the measurement of magnetic potential at different locations in a magnetic circuit is required. Among the more versatile of such tools is the Ellwood magnetomotive force gauge.[6] A brief description of its operation appears in a companion article.[5]

The above measurements provide information on the static magnetic characteristics of a relay, but they do not provide all the information

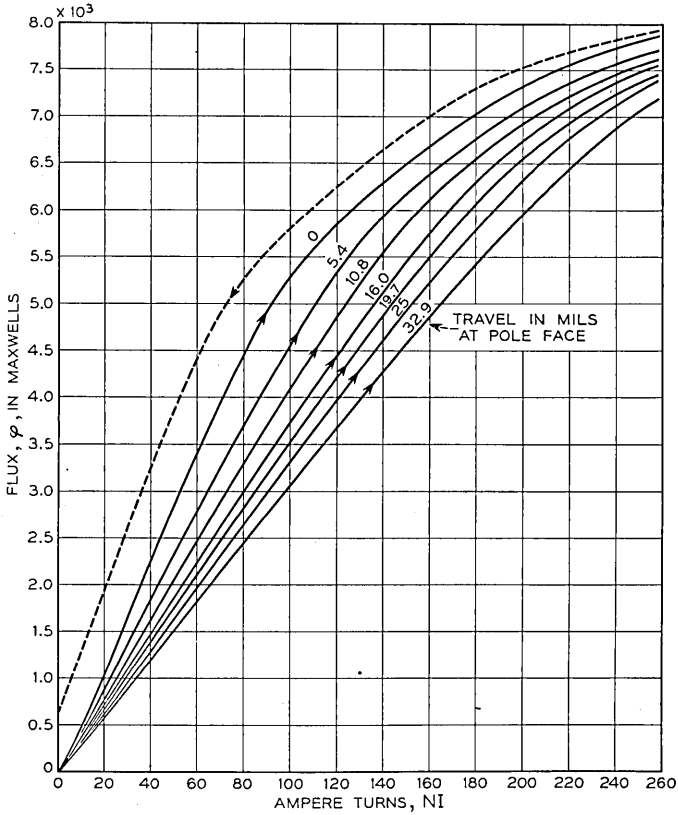


Fig. 4 — Typical recording of flux-ampere turn data on a relay magnet.

needed to describe the action while parts are in motion. For measurements involving dynamics, additional methods are described below.

DYNAMIC MEASUREMENTS

The measurements just described are indispensable in obtaining basic information on the relay's ability to perform its prescribed mechanical functions, which in turn depend upon its magnetic characteristics. They are needed in the every-day calculation of winding designs and the application of contact spring combinations which will be reliably operated by a given magnet. Cases arise in service, however, where very rigid requirements must be placed on closure and separation to insure the desired performance. Complete knowledge of the influence of voltage variations, mechanical adjustments, numbers of springs and their natural fre-

quencies is needed. Often impacts and slide of parts must be minimized in order to reduce mechanical wear. Furthermore, the design of high speed relays requires an understanding of how the flux rises and decays when current is connected or disconnected.

For the experimenter in these fields, many tools are available covering measurements of force, flux, current, and time, or combinations thereof. These measurements may involve shadowgraph, high speed motion picture, or various transducer techniques, and many special methods; some of these are described below.

The Measurement of Force

There is a definite need to know how forces in a relay structure vary with time. It would, for example, be very useful to measure the force experienced by the relay armature as it presses against its spring load during operation. While a satisfactory measurement technique for this problem has not yet been found, there are a number of similar type measurements which can be applied to individual mechanical problems in the structure. For example, with the higher-speed relay functions of today, it is found that relays made by older methods suffer excessive

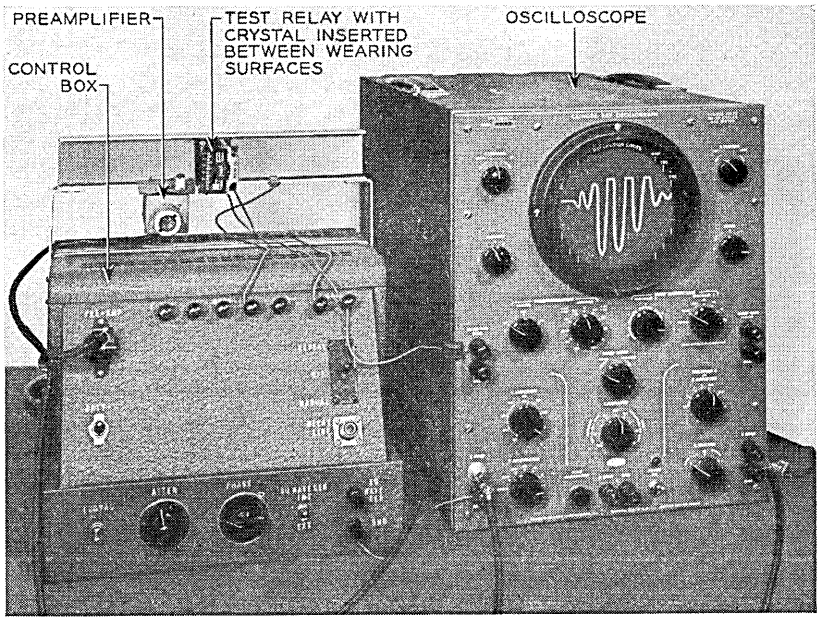


Fig. 5 — Barium titanate test set.

mechanical wear. In order to decide whether this is due to impacts, slide of parts, or vibration, methods have been developed to isolate some of the transient forces encountered in relay operation, and measure them individually.

Rapidly varying forces, such as impact forces, may be measured with the barium titanate crystal, which acts piezoelectrically to yield a voltage across its surfaces when subjected to compression or shear type forces. Since it can be obtained in very small sizes, it may be mounted within the relay as a substitute for the part to be studied. By observing the voltages that have been developed across it during relay operation, one may readily measure the forces involved.[7] Such an arrangement can be made to provide a faithful frequency response over a wide range, though it calls for careful amplifier circuit design. The unit shown in Fig. 5 provides for accurate frequency response for impact forces varying as high as 50 kilocycles. Fig. 6 shows a typical force-time relation obtained with this equipment. In some recent measurements on an experimental relay, the impact forces between the driving members were measured to be approximately five-fold the static force, but not sufficient to explain pulverizing and other damage to certain of the functioning parts.

Imposed Motion

Slide between parts has been found to be a very damaging source of wear in telephone relays. It has been studied in some detail by means of

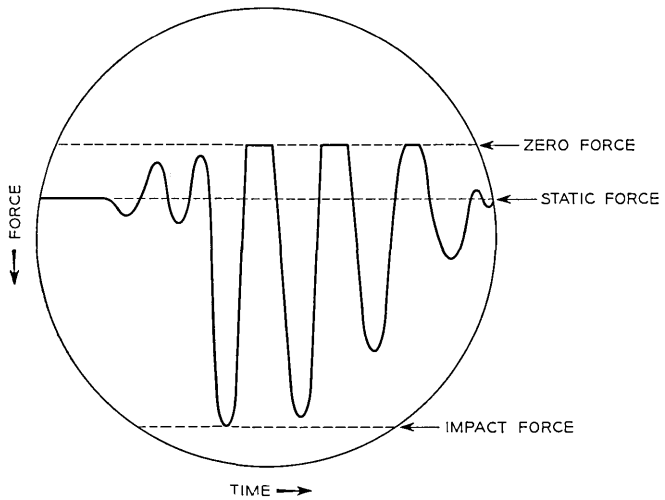


Fig. 6 — Oscillogram of impact in a relay.

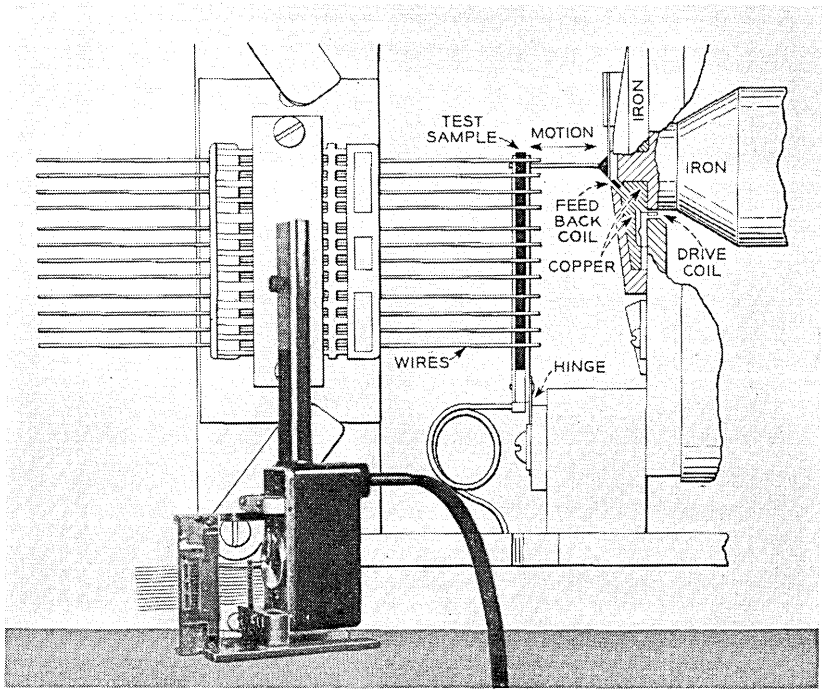


Fig. 7 — Photo of 1A recorder setup.

transducer elements whose motion is controllable to simulate a wide range of slide conditions typical of possible designs, from which wear measurements may be taken and analyzed. Such methods were also described in the previous reference.[7] The measurements may be made either by means of a barium titanate element or by means of a phonograph recording head such as the Western Electric 1A recorder. Because of the feedback coil and amplifier circuit, the latter transducer will generate the precise amplitude of motion desired for any particular test merely by properly setting its input current. It is normally driven at frequencies in the order of 1,000 cycles, permitting a large number of slides to be obtained in a short time and affording information to the designer on a very accelerated basis. Through the choice of different amplitudes, different materials to be tested, and different forces between them, a wide range of variables can be covered, to guide the designer in the proper choice of materials and conditions of use. A recorder setup for a typical test is shown in Fig. 7; data taken on a typical set of parts for an AF type relay are shown in Figure 8.

The Measurement of Flux

When attempts were first made to extend the speed performance of general purpose relays much below 8 or 10 milliseconds functioning time, it was found that the simplified eddy-current theory, which treated the core like a single short-circuited turn, was no longer sufficiently accurate. The subject has been clarified through the use of the dynamic fluxmeter which permits a direct determination of how the relay flux changes with time, under actual conditions. This fluxmeter, originally due to E. L. Norton and recently refined by M. A. Logan, was described in a previous issue,[8] and has proven most useful for determining those constants of the relay which the designer must understand in order to work much below 8 milliseconds functioning time. The instrument is shown in Fig. 9. It requires that the relay under test be pulsed, usually at a speed of about 10 to 20 cycles. When the main winding is alternately connected and disconnected, the search coil which surrounds the winding sends alternate positive and negative voltage impulses to a dc ammeter connected across its terminals. Now it can be shown that when the search coil and meter are disconnected during the interval from time zero to the moment of interest, then the resulting reading on the meter is exactly proportional to the flux at that moment. This switching function is provided by a timing circuit comprising a 40-kc frequency source driving a 3 decade counting ring, under control of switches permitting selection of the number of cycles of delay time. In this way the flux may be measured at intervals of 25 millionths of a second, with better than 1 per cent accuracy.

One interesting measurement has shown that in the short time intervals of present-day relays, eddy currents have less effect on the initial

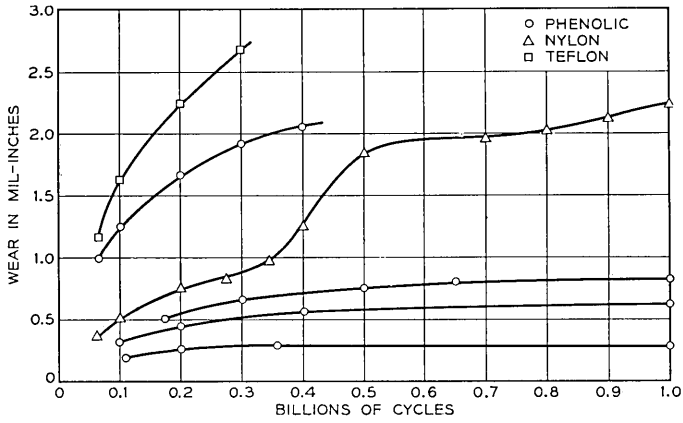


Fig. 8 — Slide data on AF relay.

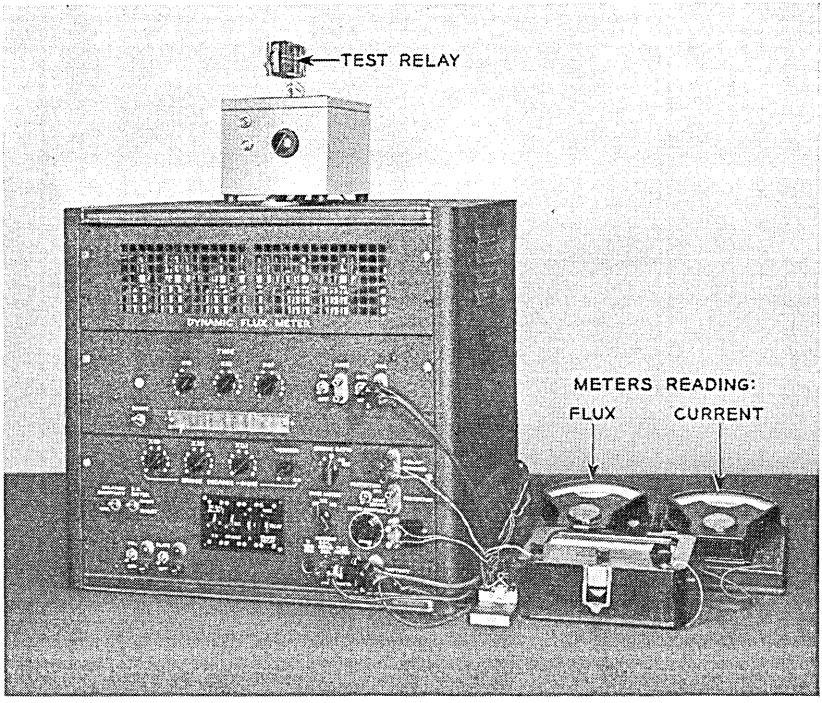


Fig. 9 — Photo of dynamic fluxmeter.

development and decay of the field than was previously thought to be the case.

The Measurement of Time — Electrical Effects

The actual functioning time of the relay is one of the more important behavior characteristics which must be tabulated for the particular design. Its measurement can be quite simple or quite complex depending on the problem involved.

One of the simpler time measurements is that permitted by means of the arrangement shown in Fig. 10. This particular equipment is composed of a control circuit for studying the various relay timing conditions, and a Berkeley Timer made by the Berkeley Instrument Company. The timer is provided with electronic controls which cause it to start counting cycles from a standard frequency when given an electrical impulse and to stop counting upon receiving a second impulse. The circuit may be arranged to give the first impulse when the relay winding is connected or disconnected and the second impulse when the contacts or other

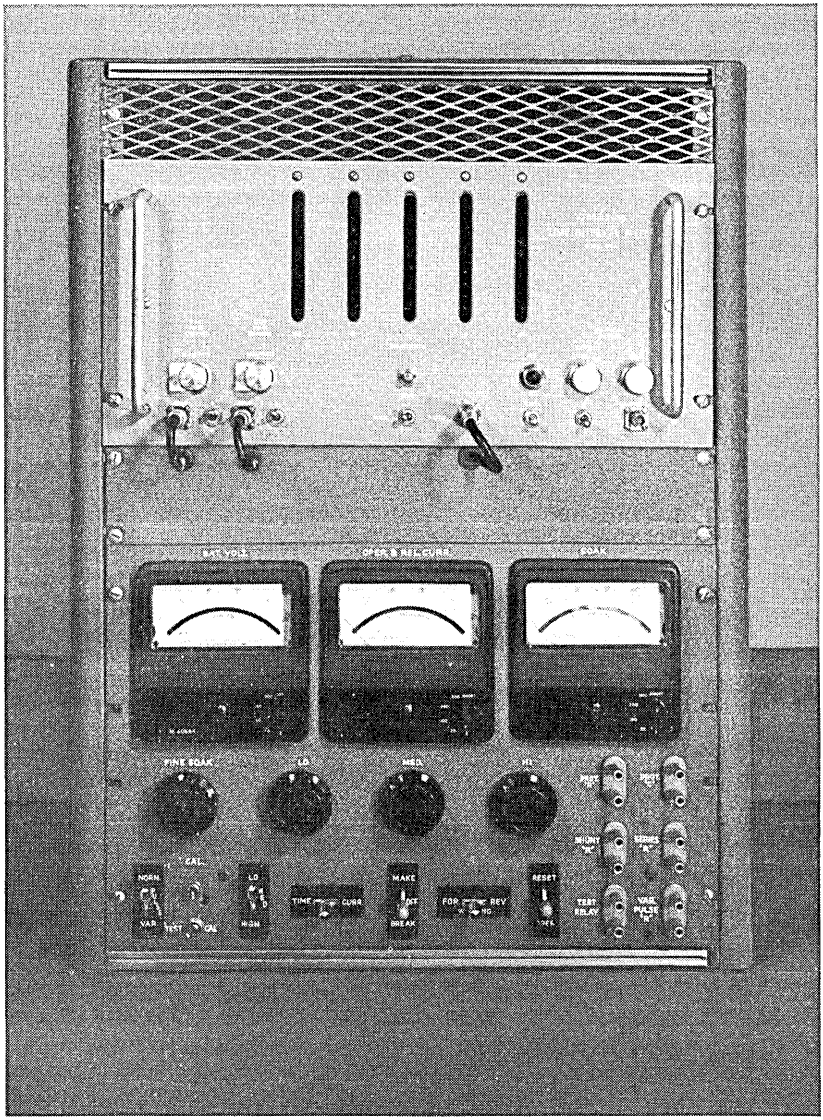


Fig. 10 — Photo of timing setup.

members either touch or separate. It is timed by a 100-kc oscillator allowing measurements to be made at intervals of 10 millionths of a second. It is in constant use for the every-day measurement of relay timing.

The simple timing measurements just described afford data on the

basic timing characteristics of the structure, but may give insufficient information on the timing performance of the relay contacts in a particular circuit. For example, contacts may close as desired but then reopen intermittently in a manner to cause either unwanted circuit behavior or undesirable contact sparking and consequent erosion. Sometimes the relay armature, in crossing the air-gap, encounters sudden loads as springs are successively picked up in the stroke, causing it to momentarily reverse its direction before pulling home. Also, when the relay armature releases, it may strike the backstop and rebound, recrossing a portion of its gap and causing undesirable reactuating of the contacts. Such momentary opening and closing of the contacts is classed as "contact chatter," and special means are needed to detect, measure, and understand its behavior.

The string oscillograph has been extremely useful for recording and studying such timing effects when the contact chatter was of comparatively long duration and low frequency. For relays whose functioning time exceeded 10 milliseconds and where chatter intervals were in the order of 2 or 3 milliseconds each for possibly 6 or 7 successive times, this instrument was most effective. For many of the faster relays, however, where chatter of this type has been completely eliminated, there remain problems of much higher frequency, shorter duration chatter, which are still of great importance to the circuit designer. In such cases the cathode ray oscilloscope is used.

A cathode ray oscilloscope arrangement for the study of chatter is shown in Fig. 11. The horizontal axis of the oscilloscope has a calibrated time base and the trace is displaced vertically to mark closing of the contacts or other events of interest. The horizontal sweep may be triggered at the initial closure of the contacts, but a variable time delay

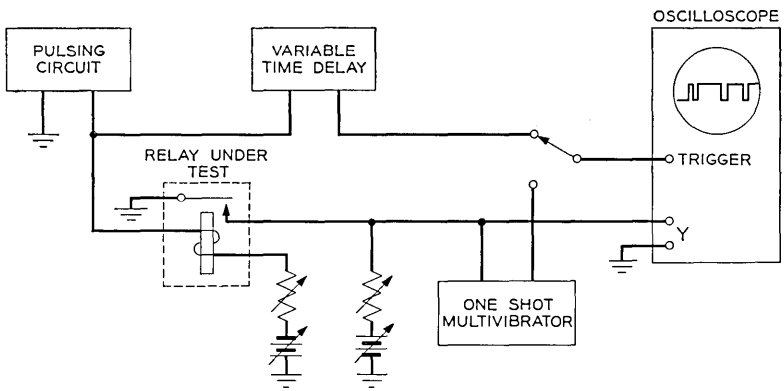


Fig. 11 — Cathode ray oscilloscope circuit for the study of contact chatter.

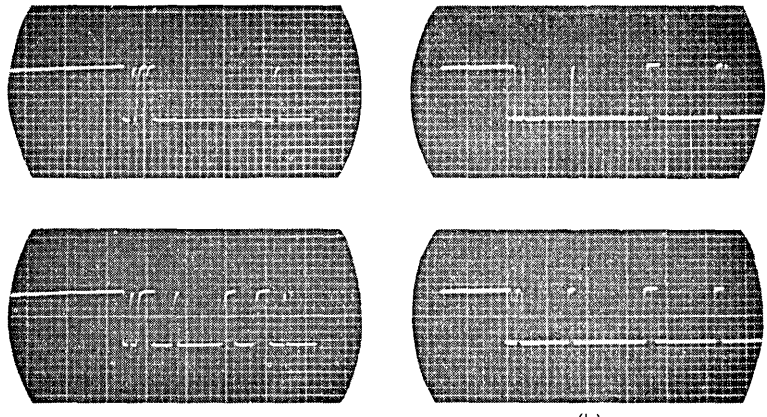


Fig. 12 — Views of chatter obtained by oscilloscope.

can be used to trigger the sweep at a later time so that any portion of the chatter can be observed in detail. Typical measurements are shown in Fig. 12. In each of the views, the horizontal lines starting at the upper left represent open contacts, the lower horizontal lines represent closed contacts, and any jumping between them indicates chatter. Fig. 12(a) shows the relatively high frequency chatter which occurs immediately after contacts close. This is called initial chatter and is caused by vibration of contact springs in their higher modes due to the impact of mating contacts. Initial chatter differs in character from the lower frequency shock chatter shown in Fig. 12(b). Shock chatter is caused by wire vibrations induced by the shock of the armature striking the core. The time of each reopen of the contacts and the time between opens is much longer than for initial chatter.

Many measurements of the type just described are made in the course of a relay development. They enable the relay designer to relate chatter performance to design characteristics of relays so that contact chatter can be reduced or eliminated. Such measurements also allow one to classify chatter performance according to the circuit application.

The Measurement of Time — Mechanical Effects

Electrical timing measurements previously described give data on over-all effective performance; to understand these results one often needs to measure displacement-time characteristics both alone, and in relation to current versus time and flux versus time. The string oscillo-

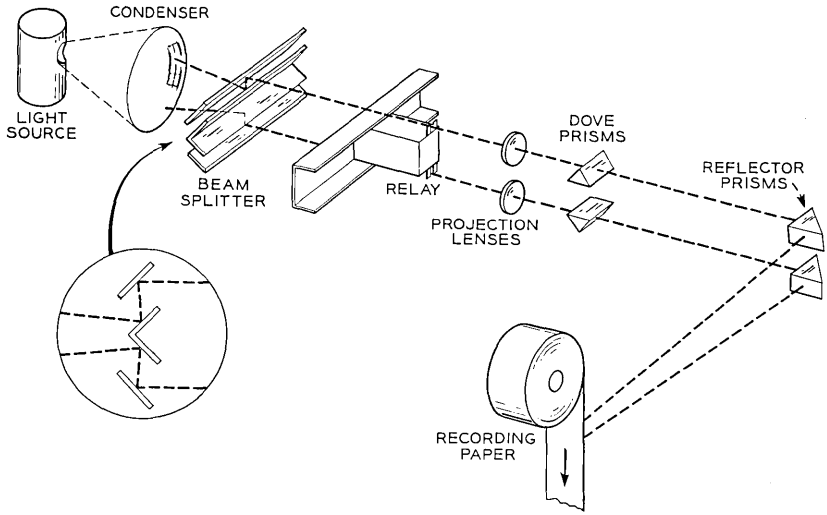


Fig. 13 — Schematic of shadowgraph action

graph mentioned above, has on many occasions been modified to also record a silhouette of the moving parts of the structure on a moving strip of photographic paper, and thus provide a simultaneous trace of displacement and current against time. One of the more advanced shadowgraph-oscillograph equipments is shown in Fig. 13, and a typical record made on a wire spring relay is given in Fig. 14. In this case the single record shows the current to the winding and the mechanical movement at the two opposite ends of the contact-operating card attached to the armature of this relay. From such measurements, velocities of impact, location of the parts at a given instant, stagger between parts, relative motion, unbalanced motion, and the like, may all be determined and correlated with electrical changes.

Motion of parts may also be studied optically to give displacement or velocity data, using the photocell. Properly placed flags on the moving

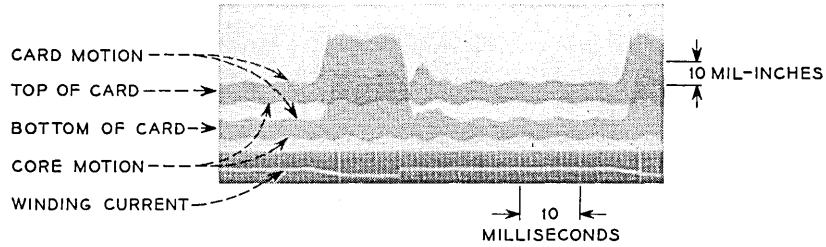


Fig. 14 — Typical shadowgram of 287 relay armature motion.

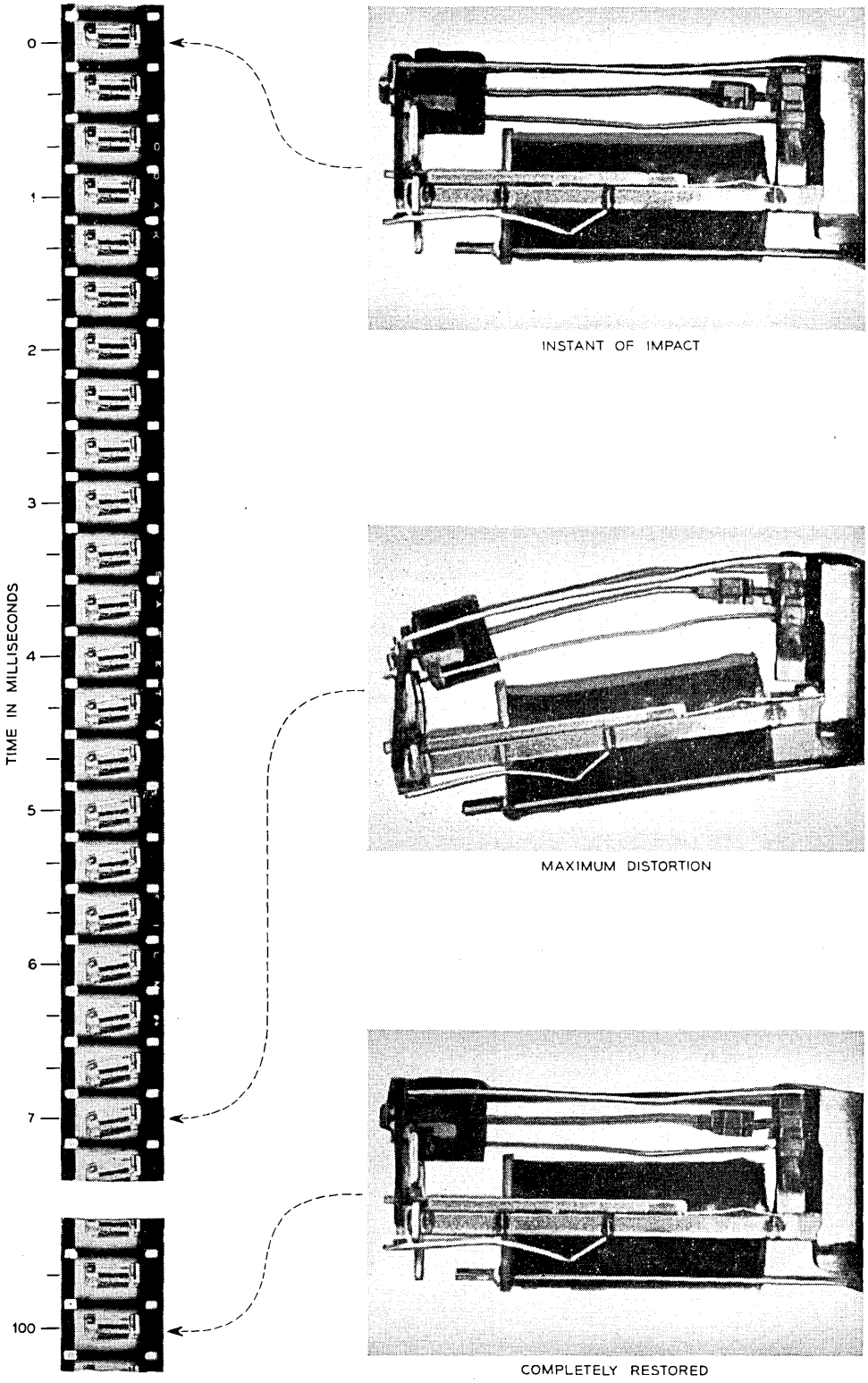


Fig. 15 — High-speed motion pictures of a falling relay at the moment of impact.

parts of the structure control the light falling on the cell, whose output can then be converted so as to read directly as displacement. The displacement data may then be differentiated as a function of time, in electrical differentiating circuits, to give very accurate measurements of velocity. A complete description of this method, originally due to E. L. Norton, has been published by M. A. Logan.[8] As a result of the accuracy and convenience of this method, it has recently been used extensively in relay studies correlating the amount of contact chatter with armature velocity.

The motion of parts may also be observed with various forms of transducers. One which is now finding application in telephone relay studies is a type of electrostatic gauge, measuring armature displacement by the changes in capacity between it and a fixed electrode. The entire scheme is to be described by T. E. Davis and A. L. Blaha in a forthcoming issue.

Some relay motions need to be studied in three dimensions. Then the high-speed motion picture gives best results. By photographing at up to 5,000 frames per second, and viewing at about 20 frames, a time magnification of 250 to 1 can be gained. If need be, such pictures may be scaled off to give displacement-time information, a somewhat tedious task. A recent study of the effect of dropping in shipment gives a striking picture of the utility of the method Fig. 15 gives views, photographed at 3,000 frames per second, of an AF relay as its mounting plate strikes a concrete block at the end of a six inch fall. Severe distortion, followed by recovery of the parts to normal, may be seen. With such information, relay designers can plan parts to stand the service stresses, and form a clear judgment of the margin of reliability built into the structure.

CONCLUSION

Although the relay is one of the oldest devices in the telephone business, many features of its operation are still imperfectly understood, even today. The number and complexity of the continuing technical problems may be judged from the other articles in this issue, on representative relay subjects. As improved relay operation becomes ever more important in the telephone system, the analytical and measuring technology for these devices must progress, in parallel. Some of the typical measuring equipment, as needed for modern relay design, has been described in this article.

REFERENCES

1. R. L. Peek, Jr., Measuring the Pull of Relays, Bell Lab. Record, June, 1953.
2. T. E. Davis, Measuring the Load-Displacement in Relays, Bell Lab. Record, June, 1953.

3. G. S. Burr, Servo-Controlled Tensile Strength Tester, Electronics, May, 1949, and E. G. Walsh, Continuously Recorded Relay Measurements, Bell Lab. Record, Jan., 1954.
4. P. P. Cioffi, Recording Fluxmeter of High Accuracy and Sensitivity, Rev. of Sci. Instr., July, 1950.
5. R. L. Peek, Jr., Analysis of Measured Magnetic and Pull Characteristics, page 79 of this issue.
6. W. B. Ellwood, A New Magnetomotive Force Gauge, Rev. Sci. Instr., **17**, p. 109, 1946.
7. W. P. Mason and S. D. White, New Techniques for Measuring Forces and Wear, B. S. T. J., May, 1952.
8. M. A. Logan, Dynamic Measurements on Electromagnetic Devices, B. S. T. J., Nov., 1953.

Magnetic Design of Relays

By R. L. PEEK, Jr., and H. N. WAGAR

(Manuscript received September 24, 1953)

The mechanical work and the speed of operation of telephone relays are determined in a large measure by the characteristics of the relay magnet. The underlying magnetic principles and the resulting design relationships for magnets are discussed in this article, which is in part a review of the background material and in part a description of its use in developing methods for magnet design and the analysis of magnet performance.

Basic energy considerations are shown to determine the relations between the work capacity and the magnetization characteristics, and analytical expressions for the latter are given in terms of the dimensions and materials of the magnet. These expressions are developed for the magnetic circuit approximation to static field theory, which is shown to provide an adequate representation of the field relations controlling performance. Methods are given for representing the magnetic circuit relations by means of a simple equivalent circuit. Expressions are derived for the mechanical output of the relay in parameters of this simple equivalent circuit, and these expressions used to determine optimum conditions for meeting specific design objectives.

INTRODUCTION

The complex switching equipment which handles the telephone traffic in automatic central offices is built up of simple component elements, of which the great majority are telephone relays. The large investment in these relays, of which tens of millions are made each year, has led to intensive effort to construct and use them as cheaply as possible, so that they will perform their function at a minimum over-all cost to the telephone system and thus to the subscriber. As the costs of use vary with the efficiency and speed, maximum economy requires the solution of technical problems in magnet design as well as the related problems of mechanical design for economy in manufacture. As a result, the technology of magnet design is under constant study at Bell Telephone Laboratories, directed to increased understanding of magnet performance

and to its improvement. This article gives the background of this technology as it applies to the dc telephone relay.

A telephone relay is an electromagnetic switch which actuates metallic contacts in response to signals applied to its coil. Its characteristics as a component of switching circuits are defined primarily by (1) the contact assembly, (2) the coil resistance, (3) the current flow requirements for operation, holding, and release, and (4) the operate and release times. The design of the contact assembly determines the number and nature of the contacts, the sequence of their operation, their current carrying capacity, and their reliability and life.

Structurally, the design of the relay, including both the electromagnet and the contact assembly, is governed not only by the performance requirements but also by the manufacturing considerations which determine the relay's initial cost, and by equipment considerations relating to the way it is mounted, wired, adjusted, and maintained. The ultimate objective is to perform a circuit function at a minimum over-all cost, including the costs of installation, power consumption, maintenance and replacement, as well as initial cost.

The design of the contact assembly determines a force-displacement characteristic representing the mechanical work which must be done by the electromagnet. The relations between this mechanical output and the electrical input to the coil are determined by the magnetic design of the relay, or specifically of its electromagnet. The electromagnetic design is therefore subject to the performance requirements, to the design of the contact assembly, and to the manufacturing and equipment considerations applying to the whole relay.

Magnetic design thus requires the ability to determine the effect on the performance of alternative choices of the configuration, dimensions, and materials of the electromagnet, of alternative load characteristics corresponding to variations in the design of the contact assembly, and of alternatives with respect to the coil dimensions and characteristics. The basic relations required for such prediction of performance are the magnetization relations, which express the field strength of the electromagnet as a function of the ampere turns, and the armature position.

The present article describes the evaluation of magnetization relations, and their use in relating the mechanical output with the electrical input to the coil. The time required for operation, which also depends upon the magnetization relations, is discussed in a separate article.[1] While much of the material given here has a broader application, the discussion of its use will be confined to the case of the simple neutral (non-polar) relay.

To illustrate the objectives of magnetic design, reference may be made to Fig. 1, which shows the load characteristic of the recently developed general purpose wire spring relay for a particular contact arrangement. The figure is taken from an article[2] describing this relay, which discusses the economic and other considerations which governed its design. Included in the figure are curves showing the pull exerted by this relay's magnet for various values of coil ampere turns. These pull curves determine the ampere turns required to operate a contact arrangement having a specific load characteristic, such as that shown. By means of the magnetization relations, these pull characteristics can be related to the design of the electromagnet.

The notation used in this article conforms to the list that is given on page 257.

1 THE COIL CONSTANT

The relay coil characteristics of interest in its use as a circuit component are its resistance and the current flow for operation. Subject to some qualifications as to available voltages and wire sizes, these may be summarized in a statement of the steady state power I^2R supplied in operation. As illustrated in Fig. 1, the coil quantity determined jointly by the load and the magnetic design is the ampere turn value NI. The power and the ampere turn value are related by the coil constant G_c or N^2/R (which equals $(NI)^2/(I^2R)$). This quantity is the equivalent single turn conductance of the coil, and is usually expressed in mhos. It is determined by the coil dimensions, as can be shown as follows:

Let A be the area of a cross-section of the coil as cut by a plane through the coil axis. Let m be the mean length of turn, the arithmetic mean of the lengths of the inner and outer turns. Then the coil volume S equals Am. If a is the cross-section of the wire, the number of turns N equals eA/a, where e is the copper efficiency, or fraction of the coil volume occupied by conductor. Substituting S/m for A, N equals $eS/(am)$. The wire length is Nm, and hence the resistance R equals $\rho Nm/a$, where ρ is the resistivity of the conductor. Hence N/R equals $a/(m\rho)$, and the coil constant is given by:

$$G_c = \frac{N^2}{R} = \frac{eS}{\rho m^2}. \tag{1}$$

The coil constant is thus independent of the wire size, except to the minor extent that the copper efficiency e decreases as the wire is made finer. With this qualification, and assuming copper wire to be used, the

coil constant is wholly determined by the coil dimensions and the type of insulation used. Thus the power required for relay operation depends upon (1) the load characteristic of the contact arrangement, (2) the magnetic characteristics which relate the pull to the ampere turns, and (3) the coil dimensions as defined by S/m^2, which determines the coil constant N^2/R, and thus the relation between ampere turns and power input.

The external dimensions of relays are largely determined by the coil

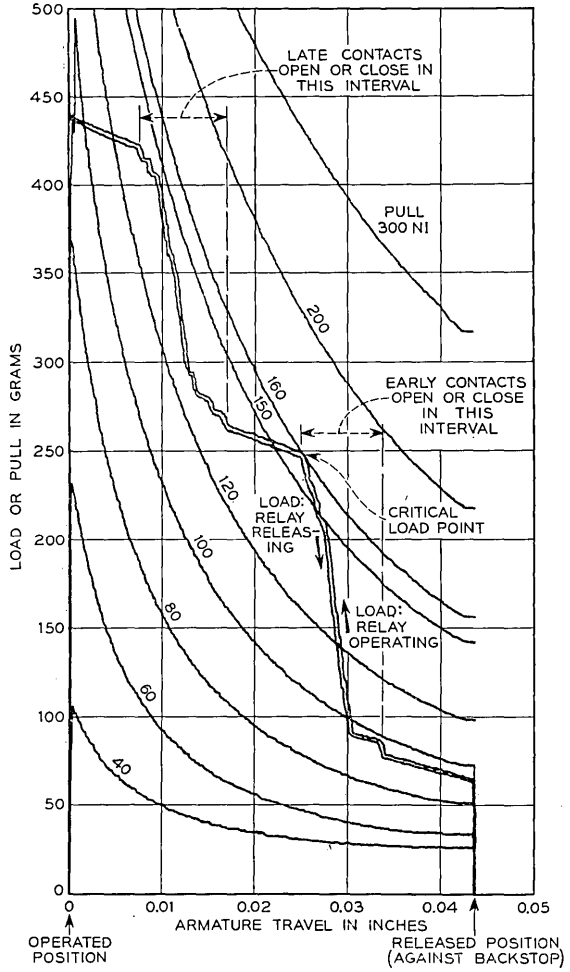


Fig. 1 — Typical load and pull characteristics of a wire spring relay with 12 transfer contacts.

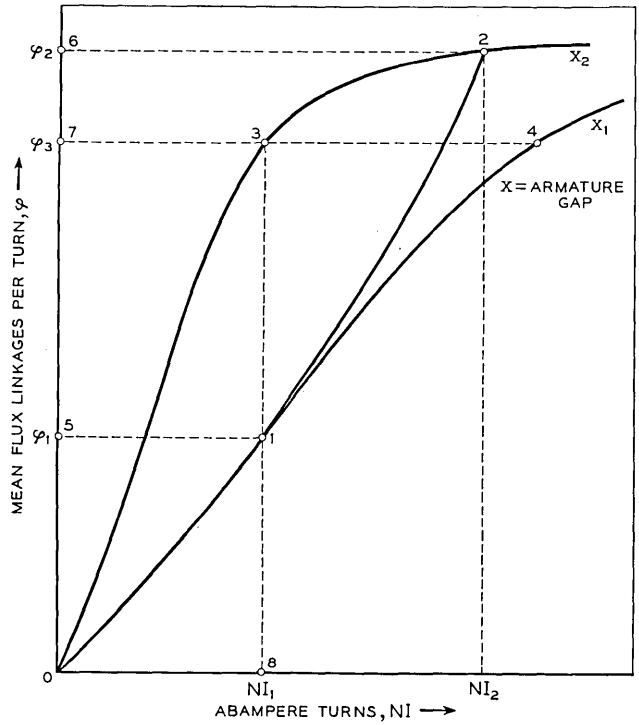


Fig. 2 — Energy relations in magnetization.

size, or winding space provided. The choice of these dimensions reflects an economic balance between the manufacturing costs of the magnet and its coil, which increase with these dimensions, and the savings in power cost resulting from increasing the coil constant.

2 MAGNETIZATION RELATIONS

The magnetization relations of an electromagnet define the steady state flux linkages of the coil as a function of the two variables: magnetomotive force $\mathfrak{F}$ (equal to $4\pi NI$), and the gap x, which specifies the armature position. They may be represented by a family of curves each giving the average flux linked per turn plotted against NI for a particular value of x. In Fig. 2, the curves marked x_1 and x_2 represent two such curves, where x_2 corresponds to a smaller gap than x_1.

To determine how the magnetization relations depend upon the dimensions and configuration of the electromagnet requires their interpretation in terms of static field theory. Such interpretation is needed in deter-

mining the design conditions for attaining a desired performance. For a specific structure, however, the observed magnetization relations, apart from any other interpretation, provide a record of the part of the electrical energy input to the coil which is stored in the electromagnet. The performance of the magnet with respect to the mechanical work which this stored energy can do may be determined directly from the way in which this energy varies with armature position. The experimental determination of the flux φ for particular values of NI and x involves a measurement of the electrical quantity $N\varphi$ defined by the equation:

$$N\varphi = \int_{i=0}^{i=I} (E - iR)\, dt. \tag{2}$$

The electrical energy U stored in making this measurement is the time integral of $i(E - iR)$, or:

$$U = \int_{i=0}^{i=I} (E - iR)i\, dt = \int_{0}^{N\varphi} i\, d(N\varphi).$$

Hence a plot of $N\varphi$ versus I gives a measure of the stored energy U represented by the area between the curve and the axis of $N\varphi$. This conclusion is quite independent of any physical meaning attached to $N\varphi$ other than the definition of equation (2). Plotting φ versus NI, rather than $N\varphi$ versus I, is merely a change of scale, which does not affect the value of the area measuring the stored energy U. It is convenient to make this change of scale because the relation between φ and NI is independent of the number of turns N, provided the location and dimensions of the coil are unchanged. The preceding expression for U may therefore be written as:

$$U = \int_{0}^{\varphi} Ni\, d\varphi. \tag{3}$$

Thus for $NI = NI_1$ in Fig. 2, U is measured by the area 0–1–5 for $x = x_1$, and by the area 0–3–7 for $x = x_2$.

Magnetization curves have the general character shown in Fig. 2. They are approximately linear for small values of φ, but at higher values bend over and approach a limiting value, φ''. This limiting value is the saturation flux, determined by the material and dimensions of the electromagnet, and designated throughout this paper by the double prime superscript.

The ratio of the magnetomotive force $\mathfrak{F}$ to the flux φ is the reluctance $\mathfrak{R}$, as defined by the equation:

$$\mathfrak{F} = 4\pi NI = \mathfrak{R}\varphi. \tag{4}$$

Over the linear portion of a magnetization curve, $\mathfrak{R}$ is a function of x only, or a constant for a particular curve. In this case, integration of equation (3) gives the following alternative expressions for the field energy U:

$$U = \frac{2\pi(NI)^2}{\mathfrak{R}} = \frac{\mathfrak{R}\varphi^2}{8\pi} = \frac{NI\varphi}{2}. \tag{5}$$

Over the upper curved portions of a magnetization curve, $\mathfrak{R}$ is a function of φ as well as of x, and equations (5) do not apply.

Decreasing Magnetization

In relay terminology, "operation," following closure of the coil circuit, is distinguished from "release," which follows opening of the circuit. The preceding discussion applies directly to the relations for increasing magnetization, as in operation. In release, the field energy and the current decrease together, giving a decrease in $N\varphi$ measured by a voltage time integral similar to the right-hand side of equation (2), but of opposite sign. The resulting decreasing magnetization curve is obtained by subtracting the decrease in $N\varphi$ from its initial value. The decreasing magnetization curve is higher than the magnetization curve, and the field energy recovered electrically is correspondingly less than that stored in magnetization; the difference corresponds to the loss of energy through hysteresis in the magnetic material.

Mechanical Output

Referring to Fig. 2, let the current have the steady value I_1, with the armature at rest at x_1. If the current increases to I_2 while the armature moves from x_1 to x_2, the flux φ varies with Ni along some curve such as 1–2 corresponding to the values of x and Ni concurrently attained. The electrical energy drawn by the coil in this process (aside from the heating loss) is given by the integral of $i\,d(N\varphi)$, or $Ni\,d\varphi$, taken along the curve 1–2. Part of this energy appears in the increase in the field energy from U_1 to U_2 as given by equation (3) for the points 1 and 2 respectively. The balance represents the mechanical work done, the integral of Fdx from x_1 to x_2 where F is the pull. Hence:

$$\int_{x_1}^{x_2} F\,dx = \int_{\varphi_1}^{\varphi_2} Ni\,d\varphi - (U_2 - U_1). \tag{6}$$

The first right-hand term is represented in Fig. 2 by the area 5–1–2–6 while U_1 and U_2 are represented respectively by the areas 0–1–5 and

0–3–2–6. Thus the mechanical work, the left-hand term of equation (6), is represented by 0–1–2–3–0, the area bounded by the two magnetization curves and the path followed by φ, x and Ni in the concurrent change from I_1 at x_1 to I_2 at x_2.

If armature motion occurs at constant flux, the first right-hand term in equation (6) is zero, and the mechanical work equals the change in the field energy U. If $\varphi = \varphi_3$ in Fig. 2 for example, the work done as the armature moves from x_1 to x_2 is $U_4 - U_3$, represented by the area 0–4–3–0. From equation (6), the pull F is then given by:

$$F = -\frac{\partial U}{\partial x} \quad (\varphi \text{ constant}). \tag{7}$$

If armature motion occurs at constant current, the first right-hand term in equation (6) becomes the change in $NI\varphi$. If $I = I_1$, for example, motion of the armature from x_1 to x_2 increases $NI\varphi$ from $NI_1\varphi_1$ to $NI_1\varphi_3$. Hence the mechanical work done is the difference between $NI_1\varphi_1 - U_3$, represented by the area 0–3–8 and $NI_1\varphi_1 - U_1$, represented by the area 0–1–8. The work done at constant current is therefore the change in the quantity W defined by the equation:

$$W = \int_0^{NI} \varphi \, d(NI),$$

which is represented by the area between the magnetization curve and the NI axis. From equation (6), the pull F is given by:

$$F = \frac{\partial W}{\partial x} \quad (I \text{ constant}). \tag{8}$$

The pull F can therefore be determined either from equation (7) or from equation (8), provided the magnetization curves are known. These equations may be applied graphically or numerically to compute the pull in specific cases. They may also be used, as shown in Section 7, to obtain expressions for the pull from expressions for the magnetization relations. In addition, they afford the following graphical interpretation of the dependence of the mechanical output upon the magnetization relations, the current, and the armature travel.

If x_1 in Fig. 2 represents the unoperated position of the armature, and x_2 its operated position, the work that can be done at constant current is $W_2 - W_1$, which varies with the value of I applying. For small values of I, for which both magnetization curves are linear, the work capacity $W_2 - W_1$ varies approximately as $(NI)^2$. At higher values of NI, the two magnetization curves approach each other as they approach the

limiting saturation flux, and the rate of increase of $W_2 - W_1$ becomes progressively smaller. The saturation flux therefore puts a ceiling on the mechanical work attainable.

If the armature travel is increased, increasing the unoperated gap x_1, the corresponding magnetization curve is lowered, reducing W_1 with a consequent increase in the work capacity. However large x_1 may be made, there remains a leakage field, to which corresponds a limiting magnetization curve. An extreme upper limit to the work capacity is represented by the area between this limiting magnetization curve and that for the operated position of the armature.

The magnetization curves thus suffice for the evaluation of both the field energy and the mechanical output associated with armature motion, and therefore completely define the static performance of the electromagnet. The problem of relating this performance to the design then reduces to the problem of relating the magnetization characteristics to the design.

3 THE MAGNETIC CIRCUIT CONCEPT

A rigorous determination of the magnetization relations from the dimensions and configuration of the electromagnet would require the solution of the static magnetic field equations. For the geometry obtaining in actual structures, such solutions can at best be obtained only for specific cases, and then only by tedious numerical or graphical methods. Approximate solutions, however, may be obtained by a procedure in which the actual distributed field is taken as confined to a limited number of paths, which together form a network analogous to an electrical circuit. The extent to which this procedure may provide a valid approximation is indicated in the following brief review of the basic postulates of static field theory. For the present purpose, these may be stated as follows:

The energy of a static magnetic field is the volume integral of the product of the magnitudes of the B and H vectors over the space containing the field. These vectors coincide in direction at all points, and are subject to the conditions: (1) that B can be represented by closed continuous lines, ("lines of induction"), whose density measures the magnitude of B, (2) that the ratio μ of B to H at any point is determined by the medium in which the point is located, (3) that the line integral of H around a closed path is equal to 4π times the total current in all circuits linking this path. This integral is the magnetomotive force $\mathfrak{F}$; it has the value $4\pi NI$ for a coil of N turns with current I flowing in them.

For a rigorous general treatment, these postulates must be expressed in differential form (the field equations), and solutions obtained in which they are satisfied at all points within the field. They may be applied directly, however, in certain cases of simple symmetry, and this same treatment, to some measure of approximation, may be used more generally. To do this, the lines of induction measuring B can be considered as grouped into tubes of induction. Each such tube is, from the properties of B, continuous, and the integral of B over a cross section of the tube is a constant quantity φ, characterizing this tube. Over a length of tube $\Delta\ell$ bounded at its ends by two surfaces over each of which H is constant exists a difference in magnetic potential, $\Delta\mathfrak{F}$, equal to the line integral $\int H d\ell$. Then $\Delta\mathfrak{F} = \varphi\Delta\mathfrak{R}$ where $\Delta\mathfrak{R}$ is the reluctance of this portion of the tube, determined by its dimensions and the permeability μ of the medium. In particular, for a tube of uniform cross-section a over length ℓ, $\Delta\mathfrak{R} = \ell/(\mu a)$. More generally, if two equipotential surfaces can be identified, bounding a region of constant permeability (such as an air gap), the solution of Laplace's equation for the region bounded by these surfaces permits the evaluation of the flux between them, and thus of the reluctance $\Delta\mathfrak{R}$.

It follows that if the pattern of the field can be recognized, so that the boundaries of some major tube of force can be determined, the reluctance of its several sections can be evaluated, and their sum $\Sigma\Delta\mathfrak{R}$ or $\mathfrak{R}$, is the ratio $\mathfrak{F}/\varphi$, where φ is the flux of the tube and $\mathfrak{F}$ the magnetomotive force of the coil linking it.

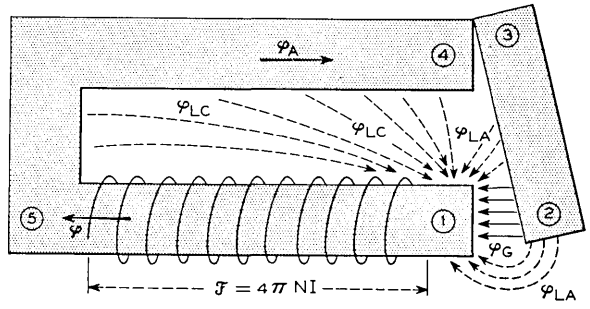
The possibility of recognizing the approximate pattern of the field results from the high values of the permeability of magnetic materials to that for air, μ_a. The ratio μ/μ_a is in excess of 1000 under normal conditions of operation. Hence the reluctances of air paths are large compared with those through magnetic material, the tubes of induction tend to follow a path through the iron, and the major changes in magnetic potential occur where they pass through air gaps. In an electromagnet such as that of Fig. 3(a), the major tube of induction follows a path through the core and armature, and the potential drops balancing the applied magnetomotive force $\mathfrak{F}$ appear principally at the air gap separating surface 1 from surface 2, and at the heel gap separating surfaces 3 and 4. There will be little potential difference between 4 and 5, but the large potential difference between this region and surface 1 will result in a leakage field between them.

Thus the total flux φ linking the coil can be considered as divided into tubes of induction φ_A and φ_{LC} following the paths indicated in Fig. 3(a).

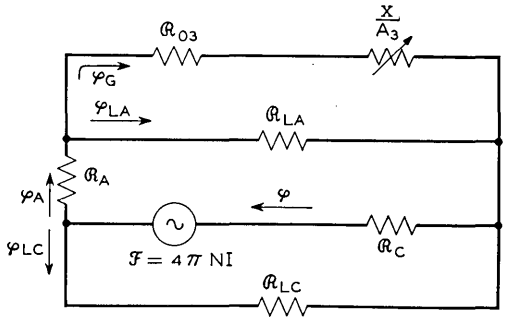
φ_A can be considered as subdivided in turn into φ_{LA} and φ_G representing respectively a leakage field from 2 to 1, and a field concentrated at the air gap between these surfaces.

These tubes of induction are analogous to the currents in an electrical network, and the applied magnetomotive force is analogous to the applied voltage producing these currents. The reluctances between equipotential surfaces are analogous to the resistances of the electrical network, and the potential differences between these surfaces are analogous to the voltage drops in the electrical resistances.

The field relationships involved may therefore be represented as in Fig. 3(b), in which the diagram of a resistance network indicates how the component reluctances determine the relation between the applied magnetomotive force $\mathfrak{F}$ and the resultant flux φ. The potential difference between the points 1 and 5 at the ends of the coil is $\mathfrak{F}$ less $\varphi\mathcal{R}_C$ where $\mathcal{R}_C$ is the reluctance of the core. This net potential balances the drop $\varphi_{LC}\,\mathcal{R}_{LC}$ across the leakage path and the drop through the armature path,



(a) MAGNETIC FIELD SCHEMATIC



(b) MAGNETIC CIRCUIT

Fig. 3 — Magnetic circuit representation of magnetic field relations.

comprising the drop $\varphi_A \mathfrak{R}_A$ through the iron part of this path in series with that across the two parallel paths followed by φ_{LA} and φ_G.

Obviously, this pattern can be elaborated and made more exact by more detailed consideration, particularly with respect to the leakage field across the coil, which is not wholly confined to that between its ends. In general, the field may be divided to any desired degree of refinement into tubes of induction, which may then be treated as a network. Thus, in principle, magnetic circuit representation may be used to evaluate the magnetization relations to any desired degree of accuracy.

Determination of the magnetization relations is thus reduced to the evaluation of the reluctances appearing in the magnetic circuit. These include both the reluctances of the iron parts for the major tubes of induction directed along the axes of these parts, and the reluctances of the air gaps and leakage paths. To evaluate these reluctances there are needed (1) values of the permeability of magnetic materials, as related to the flux density within them, and (2) expressions for the reluctances of gaps and leakage paths. These two topics are discussed in the sections following. While much of the material in these two sections is familiar, it is reviewed here with emphasis on the analytical formulation of the relations involved.

4 MAGNETIZATION CHARACTERISTICS OF MATERIALS

Magnetic properties are expressed either in terms of the relation between the permeability μ and the induction B, or of that between B and the potential gradient H, from which the $\mu - B$ relation is derived. The experimental determination of this relation is most commonly made with a uniformly wound ring sample, in which B and H are uniform throughout.

For increasing magnetization, as in relay operation, the pertinent $B-H$ relation is that obtained with an initially demagnetized specimen, shown as the solid line in the first quadrant in Fig. 4. This is the normal magnetization curve. It is nearly coincident with the locus of the terminal points (such as 1 and 2) of hysteresis loops obtained by cyclic magnetization and demagnetization.

The normal magnetization curve for magnetic iron is shown in Fig. 5 together with the corresponding $\mu - B$ curve. The permeability increases from its initial value μ_0 to a maximum value μ' at a density B' corresponding to the "knee" of the $B-H$ curve. (The single prime superscript is used throughout this paper to designate values of the various magnetic constants at maximum permeability.) Thereafter μ declines,

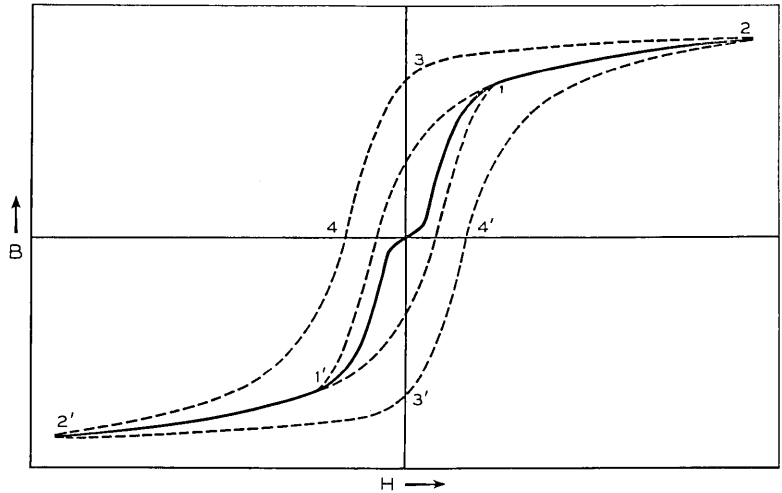


Fig. 4 — Cyclic magnetization relations.

and can be considered to approach zero as B approaches the saturation density B''.

In the design and operation of ordinary electromagnets, the aspects of the $\mu-B$ relation of practical importance are (1) the order of magnitude of μ through the central portion of the curve, which determines the approximate magnitude of the (minor) contribution of the iron path to the total reluctance, and (2) the value of B at which μ becomes relatively small (less, for example, than 1000). At or near this value of B, the magnetic path reluctance increases rapidly for small increments in B, limiting the field strength and pull attainable. In the core, this limiting value of B, together with the cross section a, limits the flux $\varphi(=aB)$ which the core can supply; conversely, this value of B establishes the core cross-section a required to attain a desired value of φ. The exact value of B which is effectively limiting varies with the design conditions, but is approximately measured by the density B_M at which μ equals 1000.

Demagnetization Relations

For decreasing magnetization, as in relay release, the pertinent $B-H$ relation is the return portion of the major hysteresis loop, the loop starting from a point on the normal curve near saturation, such as the loop 2–3–4 of Fig. 4. Nearly the same loop is obtained for any location of the point 2 well beyond the knee of the curve. The distinguishing feature of this relation is the existence of the remanence B_R at the point 3, where

$H = 0$. To account for this there must be added to the magnetic postulates cited above the assumption that there may be a movement of electric charge within the material which supplies an effective mmf per unit length measured by the coercive force H_C, represented in Fig. 4 by the distance from the origin to the point 4.

When the coil circuit is opened in relay release, the coercive force of the core material results in a residual flux passing through the air gaps in series with the iron path. A potential drop $\mathfrak{F}$ must exist across these gaps. In the absence of applied magnetomotive force, this must be balanced by an equal and opposite drop through the core. There is thus an imposed negative potential gradient $-H$ per unit length of the core. The remanence B_R is therefore governed by the $B-H$ relation in the second quadrant, the curve 3–4 of Fig. 4. This is the demagnetization relation, as illustrated separately in Fig. 6. The intercepts of this curve on the B and H axes are H_C and B_R.

If $\mathfrak{R}_E$ is the reluctance of the magnetic circuit external to the core, the external drop $\mathfrak{F}_E$ equals $\mathfrak{R}_E\varphi$, where φ is the flux. Taking B and H as uniform in the core, $\varphi = Ba$ and $\mathfrak{F}$ equals $H\ell$, where a and ℓ are the cross-sectional area and length of the core. The values of B and H in the core must therefore conform to the equation:

$$-\frac{B}{H} = \frac{\ell}{a\mathfrak{R}_E}. \tag{9}$$

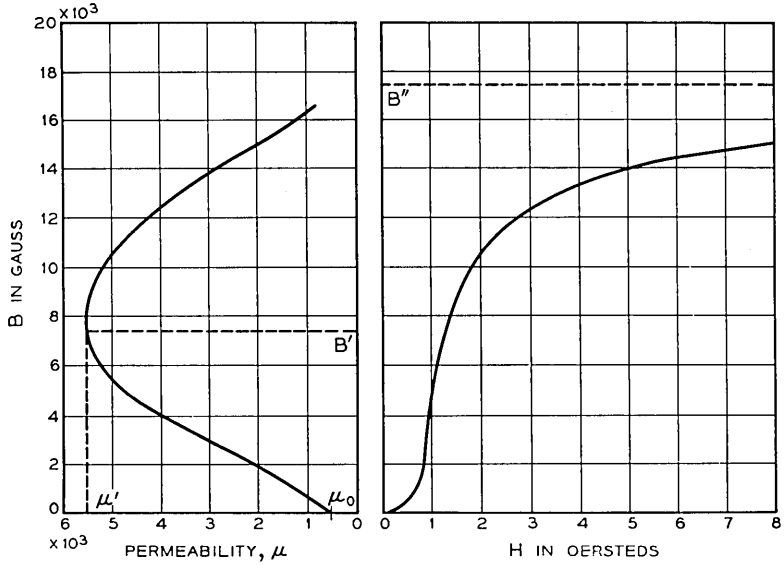


Fig. 5 — μ-B and B-H curves for magnetic iron.

As illustrated in Fig. 6, B and H are therefore determined by the intersection with the demagnetization curve of the line having a slope given by the right-hand side of (9). The complete magnetic circuit has a residual flux $\varphi = Ba$ resulting from the coercive mmf, $H_c\ell$, which equals the sum of the potential drop $H\ell(=\mathcal{R}\varphi)$ external to the core and the drop $(H_c - H)\ell$ in the core.

Properties of Magnetic Materials

The properties of the magnetic materials commonly used for electromagnets are shown in Table I. This includes, in addition to the magnetic properties, the working properties in manufacture, and the resistivity. The latter determines the eddy current delay in otherwise similar structures, and is a controlling factor where fast release is desired.

For design purposes, allowance must be made for variations in the material, which are relatively large in the case of magnetic properties. Minimum and maximum values are therefore given in Table I. As indicated by the footnote in this table, H_c is subject not only to initial variation, but to an increase with time: "aging." The relatively large magnitude of this change for magnetic iron is a serious disadvantage in

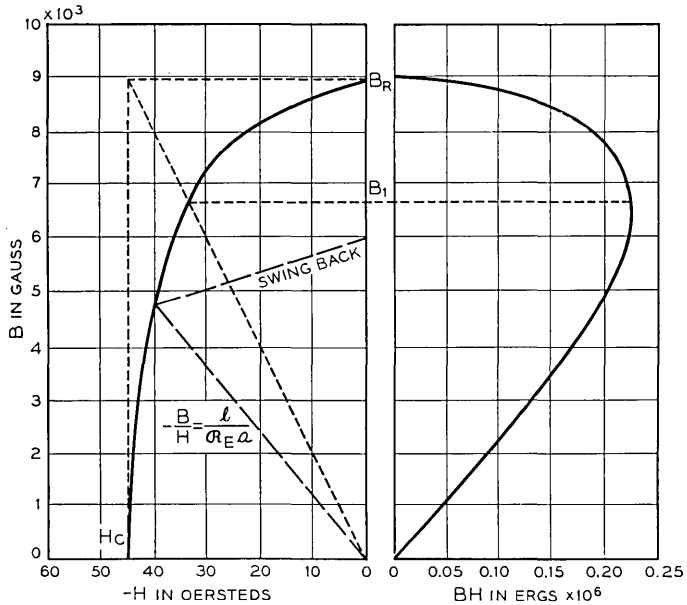


Fig. 6 — Demagnetization and energy product curves (1½ per cent chrome steel).

TABLE I — SOFT MAGNETIC MATERIALS

Material	Magnetization				Demagnetization				Resistivity Microhms cm^{-1}
	μ_{max}		B_M† (gauss†)		B_R (gauss)		H_C (oersteds)		
	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.	
Magnetic iron									
Sheet	4,300*	12,000	15,500	16,200	10,500	15,000	0.5	1.4*	11
Rod	4,300*	8,000	15,500	16,200	10,500	12,000	0.5	1.4*	11
Mild steel	2,200	7,500	14,000	16,000	7,800	15,000	0.8	2.5	12
Cast iron	(Nominal: 600)		(B'' = 10,000)		—	—	—	—	25
1% silicon iron	4,000	15,000	14,700	15,600	9,000	14,500	0.4	1.4	25
2½% silicon iron	4,000	12,000	14,000	15,000	8,000	12,000	0.4	1.4	40
4% silicon iron	5,000	12,000	13,500	15,000	8,000	12,000	0.3	1.1	60
H_2 anneal iron									
Sheet	7,000	15,000	15,500	16,500	14,000	15,000	0.5	0.9	11
Rod	4,500	8,000	15,500	16,500	12,000	14,000	0.7	1.0	11
45 permalloy	15,000	60,000	14,000	15,000	8,000	12,000	0.1	0.4	50
78 permalloy	50,000	250,000	10,000	10,400	5,000	7,500	0.02	0.1	16
Nickel	(Nominal: 600)		(B'' = 6,000)		—	—	—	—	8

* After aging, μ_{max} for magnetic iron may be as low as 3000, and H_C as high as 2.5.

† B_M: Density for μ = 1000. B'': Saturation density.

applications in which release performance is important. Aging also reduces the permeability. In materials other than magnetic iron, the aging effect is relatively minor.

Because it is cheap and easily fabricated, magnetic iron has been the most commonly used magnetic material in Bell System relays and switching electromagnets. Other materials, particularly 45 permalloy, have been used in special applications where the improvement in performance, as in higher sensitivity or faster release, warranted the increased cost. The superiority of the permalloys in these respects is offset for heavy duty applications by the lower level of flux attainable with a given core cross-section, as measured by B_M.

The silicon steels are comparable with magnetic iron in cost, and similar in magnetic properties, except for slightly smaller values of B_M. They are, however, superior in their relative freedom from aging, and in their higher resistivity. These advantages are offset, in the case of the 4 per cent material, by its hardness and brittleness. All the silicon steels are used in sheet form for the construction of transformers and electrical machines. The 1 per cent and 2½ per cent materials are available in rod and bar form, and have working properties intermediate between those of magnetic iron and 4 per cent silicon iron. Because of its advantages in aging and resistivity, 1 per cent silicon iron has been used in preference

to magnetic iron in the recently developed wire spring general purpose relay.

Magnetic iron is a very low carbon steel of high purity. Commercial mild steel, properly annealed, will serve as an adequate substitute where the spread in properties shown in Table I can be tolerated.

In magnet design, casting offers the possibility of providing a more complex one-piece configuration than can be obtained with punched and formed parts. Casting is rarely used in current magnet construction — an illustrative exception is the British Post Office stepping magnet for their step-by-step switch. Grey cast iron has been a preferred material for this type of construction.

The properties of nickel are included because of its use as a magnetic separator and hinge member.

Hyperbolic Approximation to Magnetization Curves

The variation of permeability with density makes it necessary to provide some formulation of the μ versus B relation in developing an analytical treatment of magnetization relations. The B versus H relation for decreasing magnetization, the loop 2–3–4 of Fig. 4, has a shape similar to that of a rectangular hyperbola, asymptotic to the line representing B''. This curve can therefore be represented approximately by the equation:

$$\frac{B}{B'' - B} = \frac{\mu''}{B''}(H_c - H), \tag{10}$$

in which μ'', B'', and H_c are constants.

This purely empirical relation is called the Froelich-Kennelly equation. In general, it does not give a satisfactory fit to the whole loop, but provides a satisfactory approximation for engineering use to the portions of the curve in either the first or second quadrants, using different values for the constants in the two cases. In addition, it may be employed to represent the upper portion of the normal, or increasing magnetization curve.

The expression for the permeability μ, or $B/(H_c + H)$ corresponding to (10) is:

$$\frac{1}{\mu} = \frac{1}{\mu''} + \frac{H_c + H}{B''}. \tag{11}$$

Hence, to the extent the B—H curve conforms to equation (10), the reciprocal of the permeability varies linearly with H. Alternately, the

permeability may be expressed in terms of B, giving the equation:

$$\mu = \frac{\mu''(B'' - B)}{B''} . \tag{12}$$

If this relation is applied to a part of a magnetic circuit, such as the core of a relay, the reluctance $\mathcal{R}_C$ of this part may be written as $\ell/(\mu a)$, where ℓ is the length and a is the cross-sectional area of the part. Then from (10), $\mathcal{R}_C$ is given by:

$$\mathcal{R}_C = \mathcal{R}_C'' \frac{\varphi''}{\varphi'' - \varphi} , \tag{13}$$

where $\mathcal{R}_C'' = \ell/(\mu'' a)$, $\varphi = Ba$, the flux through the part, and φ'' is the saturation value of φ.

For increasing magnetization, this expression is applicable only for values of B, or φ/a, beyond the knee of the magnetization curve, the point of maximum permeability. Thus (13) is applicable for values of B above B', the density at which μ has its maximum value μ'. Writing $\mathcal{R}_C'$ for $\ell/(\mu' a)$ and φ' for $B'a$, (13) may be written in the alternative form:

$$\mathcal{R}_C = \mathcal{R}_C' \frac{\varphi'' - \varphi'}{\varphi'' - \varphi} . \tag{13A}$$

For values of B below B' the permeability of initially demagnetized material varies greatly with B, as shown in Fig. 5. In normal use, however, an electromagnet is rarely operated from a fully demagnetized state. Fig. 7(a) shows the magnetization relations for an electromagnet in repeated operation, and Fig. 7(b) shows the corresponding relations for its core. The solid line corresponds to initial operation from a demagnetized condition, the dotted line to decreasing magnetization in release, and the dashed line to subsequent remagnetization. The latter corresponds to higher permeability and a lower core reluctance than those for initial magnetization. As a convenient approximation, linear magnetization may be taken as a representative condition for B less than B', in which case the core reluctance is constant for φ less than φ'. In this low density region, then, the magnetic circuit constants may be considered to be independent of the flux. This, of course, is never strictly true, but it is a satisfactory approximation for most ordinary electromagnets.

Hyperbolic Approximation to Decreasing Magnetization Curves

As the hyperbolic approximation is a purely empirical relation, it may be applied to the $\varphi - \mathfrak{F}$ relation for an electromagnet as well as to that

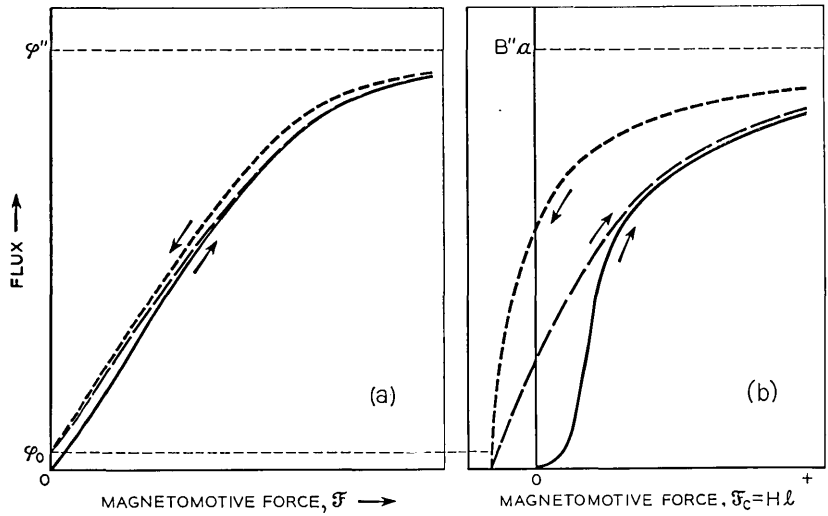


Fig. 7 — Repeated magnetization of an electromagnet and its core.

of a core or other part. It is convenient to use it in this way for the decreasing magnetization relation, as this is of interest only for a single value of the gap reluctance, that for the operated position.

The decreasing magnetization relations for an electromagnet and for its core are shown in Figs. 7(a) and 7(b) respectively. For decreasing magnetization, a residual flux φ_0 remains when $\mathcal{F}$ is reduced to zero, determined by the same equilibrium conditions as apply to permanent magnets. The curve for the magnet, Fig. 7(a), must pass through φ_0 and be asymptotic to the saturation flux φ''.

This relation between φ and $\mathcal{F}$ may be represented by an equation of the same form as equation (10), with B and B'' replaced by φ and φ'', with H and H_C replaced by $\mathcal{F}$ and $\mathcal{F}_C$, and the constant μ'' replaced by $1/\mathcal{R}''$. Using the condition that $\varphi = \varphi_0$ for $\mathcal{F} = 0$ to eliminate $\mathcal{F}_C$, the equation may be written in the form

$$\frac{\mathcal{F}}{\mathcal{R}''\varphi''} = \frac{\varphi}{\varphi'' - \varphi_0} - \frac{\varphi_0}{\varphi'' - \varphi_0}. \tag{14}$$

If the length ℓ and cross-section a of the core are known, together with the magnetic constants of the core material and the reluctance $\mathcal{R}$ of the return path (external to the core), the constant terms in (14) may be evaluted. φ'' is equal to aB'', where B'' is the saturation density of the core material. As previously noted, the values of B_M in Table I may be used as effective values of B''. φ_0 may be evaluated as the permanent

magnet flux supplied by the core to an external path of reluctance $\mathfrak{R}$, using equation (9) and the demagnetization curve for the core material. $\mathfrak{R}''$ may be determined from the initial slope $\mathfrak{R}_i = d\mathfrak{F}/d\varphi$ at $\mathfrak{F} = 0$. By differentiation of equation (14), this initial slope is given by the equation:

$$\mathfrak{R}_i = \left(\frac{\varphi''}{\varphi'' - \varphi_0}\right)^2 \mathfrak{R}''. \tag{15}$$

$\mathfrak{R}_i$ may be evaluated as the sum of the core reluctance and the external reluctance $\mathfrak{R}$. In determining the core reluctance, μ should be taken as the incremental permeability, or the slope of the demagnetization curve at $B = \varphi_0/a$. With $\mathfrak{R}_i$ thus evaluated, $\mathfrak{R}''$ is given by equation (15).

5 MAGNETIC RELUCTANCES AND CONDUCTANCES

As shown in Section 3, determination of the magnetization or $\varphi - \mathfrak{F}$ relations is equivalent to determining the reluctance $\mathfrak{R}$, or $\mathfrak{F}/\varphi$, and this in turn reduces to the determination of the component reluctances of the magnetic circuit. Some of the more useful expressions for the evaluation of these component reluctances are given in this section. Where parallel paths appear, the computations involve the reciprocals of reluctances. It is convenient to refer to the reciprocal of a reluctance as a magnetic conductance, or permeance.

Toroid

The magnetic field of a toroidal coil is shown in Fig. 8. Provided the medium within the coil is homogeneous, and the section diameter small compared with the toroid's diameter, symmetry requires the field to have the simple character shown. As $\mathfrak{F} = H\ell$ and $B = \mu H$,

$$\mathfrak{R} = \frac{\mathfrak{F}}{\varphi} = \frac{H\ell}{\mu H a} = \frac{\ell}{\mu a}, \tag{17}$$

which applies whether the path is in iron or air.

If the toroid is of iron, and the coil is concentrated over part of the length, some of the field appears outside the toroid. This effect is secondary, and (17) still applies to a close approximation. If a cut is made and an air gap introduced, the total reluctance is greatly increased, but the field within the toroid retains the same character except near the gap. The reluctance of the iron path, now in series with the air gap reluctance, is given by (17). This expression is therefore of general use in determining the reluctance of iron parts of length ℓ and uniform cross-section a.

Air Gap Between Parallel Planes

For magnetic paths in which the flux is uniform, it was shown above that the reluctance is given by (17). This applies to the case of an air gap between parallel planes of area A, as shown in Fig. 9, where the length of the path is the separation x. The reluctance of such a gap may therefore be written as:

$$\mathcal{R} = \frac{x}{A}.$$

In this, as in subsequent expressions for air gap reluctance, the multiplier $1/\mu_a$, which is the reciprocal of the permeability of air, is omitted for convenience. In CGS units, $\mu_a = 1$.

Special Gap Shapes

The simple relation just given provides a basis for the calculation of more complex shapes. For example, the reluctance of the wedge shaped gap of Fig. 10 may be found by assuming tubes of flux in parallel throughout the gap. Then the permeance of an elementary path is $\Delta\mathcal{P} = b\Delta r/r\theta$ and the total permeance of the gap is the sum of these elementary paths, or:

$$\mathcal{P} = \frac{b}{\theta}\int_{r_1}^{r_2}\frac{dr}{r} = \frac{b}{\theta}\ln\frac{r_2}{r_1}.$$

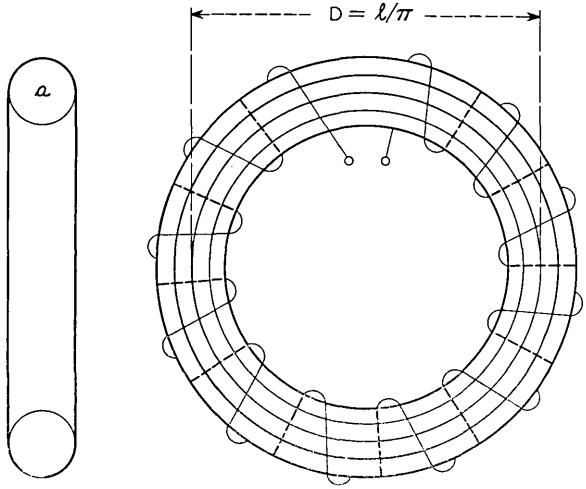


Fig. 8 — Field of a uniformly wound toroidal core.

The reluctance, as given in the figure, is the reciprocal of this.

For the gap of Fig. 11, it is more convenient to estimate reluctances, summing the elementary series contributions to the entire gap. For this case:

$$\Delta \mathcal{R} = \frac{\Delta r}{br\theta},$$

where

$$\mathcal{R} = \frac{1}{b\theta}\int_{r_1}^{r_2}\frac{dr}{r} = \frac{1}{b\theta}\ln\frac{r_2}{r_1},$$

as noted in the figure.

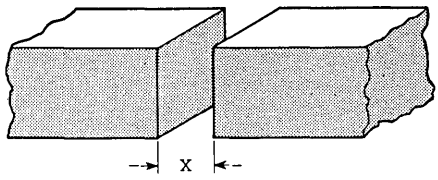


Fig. 9 — Reluctance between parallel plane surfaces.

Effective Pole Face Area of an Electromagnet

As illustrated in Figs. 9, 10, and 11, an individual air gap has a reluctance represented approximately by an expression of the form:

$$\mathcal{R} = \frac{x}{A},$$

where x is the separation measured at the centroid of the area A. Armature motion is usually rotary, and a convenient point for the measurement of armature position may be at some distance from the axis other than that of the centroid. If x is so measured, then the separation at the centroid is kx, where k is the ratio of the lever arms, and hence the reluctance $\mathcal{R}$ is kx/A, equivalent to that of a gap of separation x and pole face area A/k. This area A/k is called the effective pole face area referred to the point at which x is measured.

In general, an armature has at least two gaps, as illustrated in Fig. 12. The expression for the armature reluctance must include the terms for both gaps which can conveniently be combined as follows. Let x be the armature motion measured at some distance ℓ from the axis of rotation, as indicated in the figure. Let k_1x be the corresponding separation at the centroid of a_1, and k_2x the separation at the centroid of a_2. Then the

reluctance of the combination is:

$$\mathfrak{R} = x\left(\frac{k_1}{a_1} + \frac{k_2}{a_2}\right).$$

It follows that the reluctance of the combined gap is equivalent to that of a simple gap with an effective pole face area given by:

$$\frac{1}{A} = \frac{k_1}{a_1} + \frac{k_2}{a_2}. \tag{18}$$

Exact: $\mathfrak{R} = \dfrac{\theta}{bln\left(\dfrac{1 + \dfrac{a}{2r_0}}{1 - \dfrac{a}{2r_0}}\right)}$

Approximate: $\mathfrak{R} = \dfrac{x}{ab}$

Errors of approximation:

Less than 1 per cent if $a < \frac{1}{3}r_0$

and/or $x < \frac{1}{2}r_0$

Less than 10 per cent if $a < r_0$

and/or $x < \frac{4}{3}r_0$

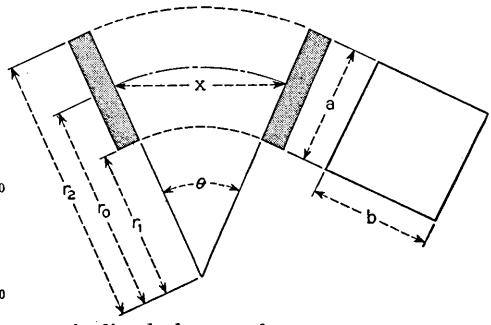


Fig. 10 — Reluctance between inclined plane surfaces.

Exact: $\mathfrak{R} = \dfrac{ln\,\frac{r_2}{r_1}}{b\theta}$

Where b = length parallel to axis

Approximate: $\mathfrak{R} = \dfrac{x}{br_0\theta}$

Errors of approximation:

Less than 1 per cent if $r_2 < 1.4r_1$,

Less than 10 per cent if $r_2 < 3r_1$,

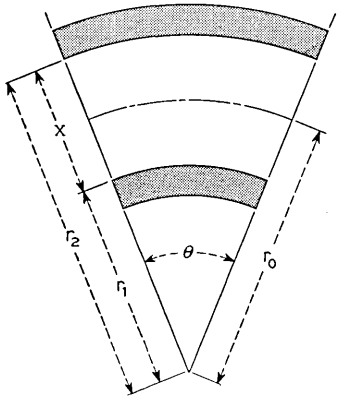


Fig. 11 — Reluctance of a cylindrical gap.

Leakage Reluctances

An accurate estimate of the reluctance between magnetic members requires a detailed knowledge of the flux paths. Since these are only known accurately in cases for which solutions of the field equations are available, approximations are obtained by assuming geometrical paths such as straight lines, arcs of circles, ellipses, and so forth. From these assumed paths the reluctance is calculated by means of the expression $\ell/\mu a$. The choice of suitable approximations depends largely upon a knowledge of the flux paths in certain simple cases which can be analyzed rigorously, and upon the experimental exploration of more complicated fields.

The method is satisfactory provided the separation between the magnetic members is small. It has been used to derive the relations given in Figs. 10 and 11. A further application of this method gives the reluctance between the side surfaces of coaxial cylinders, as shown in Fig. 13. This is useful in estimating the leakage reluctance shunting an air gap.

Where the separation between magnetic members is large, it is difficult to estimate the configuration of the flux paths. It is then necessary to employ the relations applying to the most nearly similar configuration for which a rigorous solution is known. Two such solutions are applicable to a number of problems. The first is the case of two infinitely long, parallel, equipotential, circular cylinders, shown in Fig. 14. The second is the case of two equipotential spheres of equal size, shown in Fig. 15.

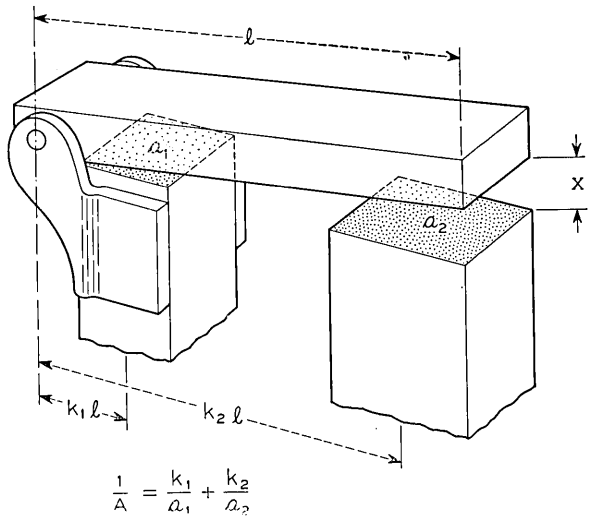
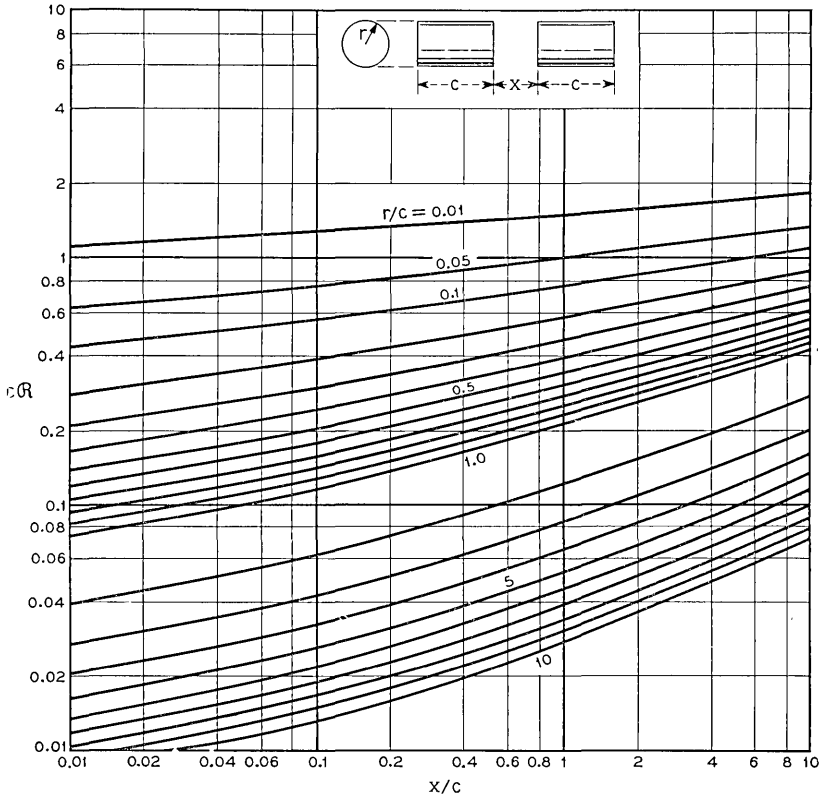


Fig. 12 — Effective pole face area referred to gap x at ℓ.



$$m = \ell n \left[1 + 2\frac{c}{x}\left(1 + \sqrt{1 + \frac{x}{c}}\right)\right],$$

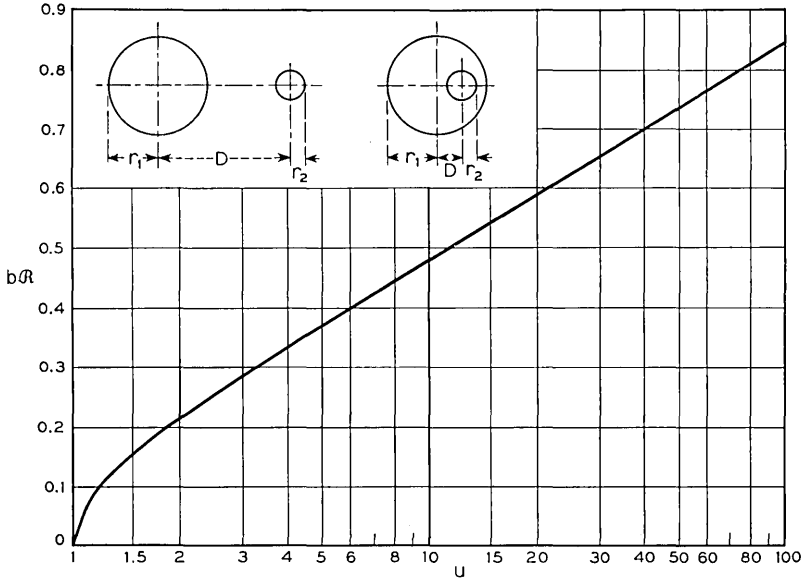
when $\frac{c}{rm} < 1$, $\qquad \Re = \dfrac{\cos^{-1}\left(\frac{c}{rm}\right)}{\pi rm \sqrt{1 - \left(\frac{c}{rm}\right)^2}}$,

when $\frac{c}{rm} = 1$, $\qquad \Re = \dfrac{1}{\pi c}$,

when $\frac{c}{rm} > 1$, $\qquad \Re = \dfrac{\ell n \left[\frac{c}{rm} + \sqrt{\left(\frac{c}{rm}\right)^2 - 1}\right]}{\pi rm \sqrt{\left(\frac{c}{rm}\right)^2 - 1}}$,

if cylinders are not circular, let $r = \dfrac{1}{2\pi} \times$ perimeter.

Fig. 13 — Reluctance between side surfaces of end-on coaxial cylinders.



$$\Re = \frac{\ell n\left(u + \sqrt{u^2 - 1}\right)}{2\pi b} = \frac{COSH^{-1}u}{2\pi b}$$

where $u = \left| \frac{D^2 - r_1^2 - r_2^2}{2r_1 r_2} \right|$

and b = length of each cylinder

Fig. 14 — Reluctance between parallel cylinders.

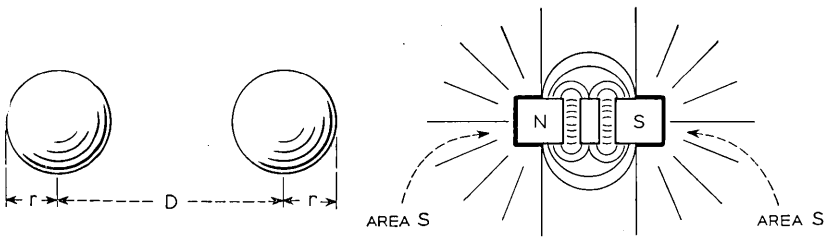
For other configurations which bear a reasonable resemblance to the above cases, the leakage reluctances can be estimated satisfactorily by judicious modifications of the expressions for the simple cases.

For example, consider the case of two parallel rectangular bars. They are roughly equivalent to two parallel circular cylinders provided the minimum separation is the same in both cases, and provided the perimeter of each cylinder is equal to the perimeter of the corresponding bar. Thus, to estimate the leakage reluctance between the rectangular bars, the radius of each equivalent cylinder is taken as $1/2\pi$ times the perimeter of the corresponding bar, and the center-to-center distance between the equivalent cylinders is taken as the minimum separation between the two bars plus the two equivalent radii.

Leakage between the legs of many magnet forms may be estimated by

the approximation just described. Consider the idealized form shown in Fig. 16, where the magnetic path consists mainly of two parallel cylinders connected at one end, the armature and working gap being at the opposite end. The leakage flux in parallel with the main gap flux is determined by the reluctance of the path between these cylinders. For that portion of the two cylinders appearing outside the coil and therefore at approximately constant potential, the reluctance is found from the relations given in Fig. 14. For the leakage reluctance over the length enclosed by the coil, the drop in potential along the core results in one-third the previous reluctance, as is shown in Section 6. The variation in permeance per unit length for such cases may be found in Fig. 17.

Leakage also occurs between the end sections of many magnet forms, and may be estimated by a procedure similar to those above, assuming the end surfaces to be equivalent to two hemispheres having the same diameter d as the cylinders. The quantity $C_2 d$ in Fig. 17 is one half the permeance between corresponding spheres, as determined from the relations of Fig. 15. From this the net leakage permeance of leg and end



Exact: $$\mathcal{R} = \frac{1}{2\pi r\left[1 + \left(\frac{r}{D}\right) + \left(\frac{r}{D}\right)^2 + \left(\frac{r}{D}\right)^3 + 2\left(\frac{r}{D}\right)^4 + 3\left(\frac{r}{D}\right)^5 + \ldots\right]}$$

Approximate: $$\mathcal{R} = \frac{1 - \frac{r}{D}}{2\pi r}, \qquad \text{when } D > 6r$$

$$\mathcal{R} = \frac{1}{2\pi r} = \frac{1}{\sqrt{\pi S}}, \qquad \text{when } D \gg r$$

where S = surface area of one sphere.

Latter approximation may be used to estimate leakage reluctance between back surfaces of pole pieces.

Fig. 15 — Reluctance between spherical surfaces.

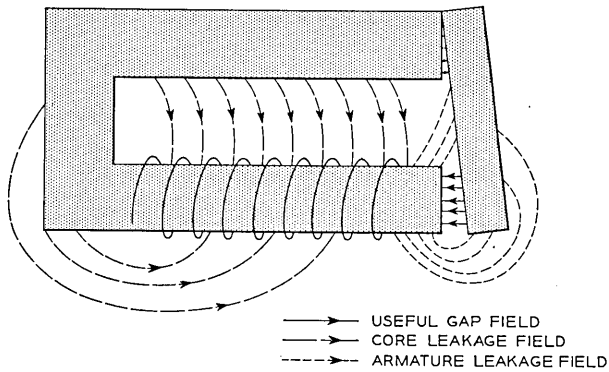


Fig. 16 — Distribution of the field of an electromagnet.

surfaces may be found as

$$\mathcal{P} = \frac{C_1 \ell_1}{3} + C_1 \ell_2 + C_2 d,$$

where values of C_1 and C_2 are given in curve form, and the equivalent value of d is as shown in Fig. 17. When the structure has two return paths, the reluctance is assumed to be one-half the value given by this last equation.

External Reluctance of a Bar Magnet

The relations of Figs. 13 through 17 suffice for the evaluation of leakage in most ordinary electromagnets. In special structures, particularly in polar relays, cases occur where the return path is predominantly an

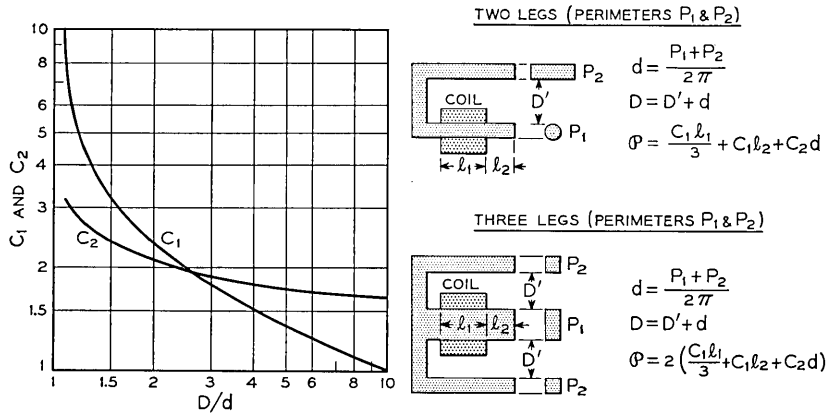


Fig. 17 — Effective reluctance between parallel bars joined at one end.

air path. A common case is that of a permanent magnet magnetized as a separate part prior to assembly in a polar structure. In such cases, the reluctance of the return path can be estimated as that of the external field of a bar magnet.

The reluctance of a bar magnet is closely represented by the reluctance of the same magnet in the form of a ring, in series with a reluctance representing all the flux return paths. Values for this reluctance may be assigned by measurements of magnetomotive force to produce a given flux in a ring sample and in a bar. The difference in magnetomotive force gives a measure of the reluctance of the air path. It has been found by

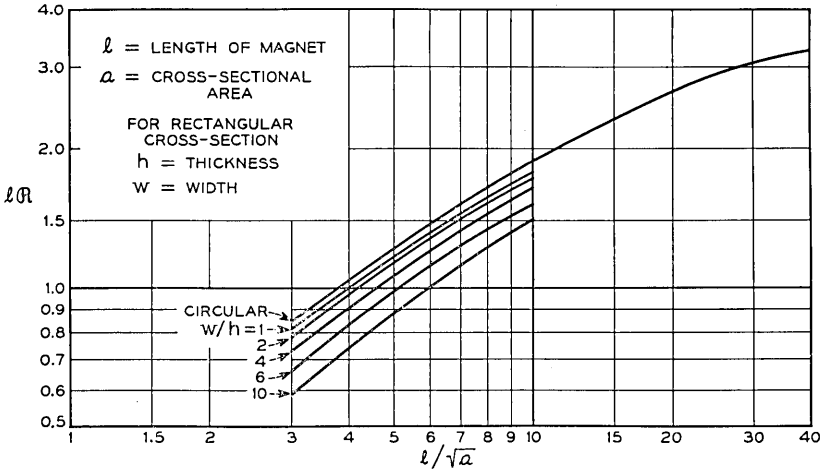


Fig. 18 — Effective leakage reluctance of a bar magnet.

Thompson and Moss[5] that the reluctance per unit length of bar depends on the shape of the bar cross-section, and is a function of the magnet dimensions $\ell/\sqrt{a}$, as indicated in Fig. 18. The reluctance so shown includes all flux lines emanating from the bar, and so may be thought of equally as leakage or effective reluctance.

Cases such as the bar magnet are difficult to estimate because of the variable flux density along the length of magnetic material. Calculations have been found possible only for the case of an ellipsoid, which is rarely met in practice.

Reluctance of a Solenoid

Another special case is that of the air field of a coil, which has the character shown in Fig. 19. This type of field obtains in the initial flux

decay of an electromagnet when the core is saturated. A similar condition applies to the leakage field of enclosed reed relays where the magnetic material constitutes only a small part of the inside cross-section of the coil. In such cases the reluctance of interest is that characterizing the complete magnetic circuit of an air core solenoid.

A number of simple relationships for various coil configurations follow from the identities between reluctance and inductance. Inductance is defined as the ratio of flux linkages to applied current:

$$L = \frac{N\varphi}{I} = \frac{4\pi N^2 \varphi}{4\pi N I},$$

from which:

$$\mathfrak{R} = \frac{\mathfrak{F}}{\varphi} = \frac{4\pi}{\left(\frac{L}{N^2}\right)}, \tag{19}$$

showing the inverse relationship between $\mathfrak{R}$ and the single-turn inductance L/N^2. Inductance for most of the usual coil configurations has been studied by Rosa and Grover[7] who give relations from which reluctance may be estimated. Most such expressions are quite involved; among the simpler is that for the case where b and c in Fig. 19 are small compared to a, given by:

$$\mathfrak{R} = \frac{0.2317a + 0.44b + 0.39c}{a^2},$$

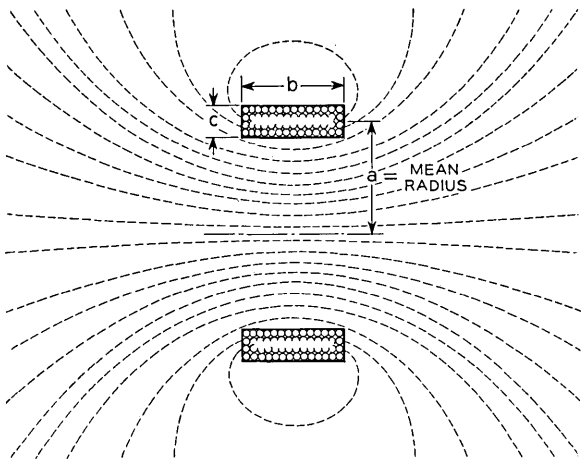


Fig. 19 — Magnetic field of a solenoid.

and Maxwell's approximation for coil of rectangular section:

$$\mathfrak{R} = \frac{1}{a\left(\log \frac{8a}{R} - 2\right)},$$

where R = geometric mean distance of coil cross-section.

The examples described in the present section cover those procedures which are used most frequently in estimating leakage reluctances. For a more extended treatment of this subject, reference may be made to S. Evershed[3] and H. C. Roters.[4]

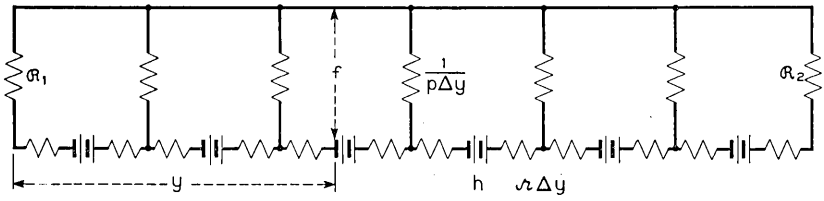
6 MAGNETIC CIRCUIT EVALUATION

In the discussion of the magnetic circuit concept in Section 3, it was noted that the accuracy of the representation varies with the extent to which the network of tubes of induction is sub-divided to correspond to the distributed nature of the actual field. The effect of sub-division upon the accuracy is greater in the high density region, where the reluctance of the iron parts is variable and increasing, than in the low density region, where the iron reluctance is small and approximately constant. Hence the choice of an adequate network is largely contingent upon the location of the iron parts having the highest flux density. Incipient saturation affects the reluctance of these parts, and thus the pattern of the field, while parts of lower density remain of low and approximately constant reluctance.

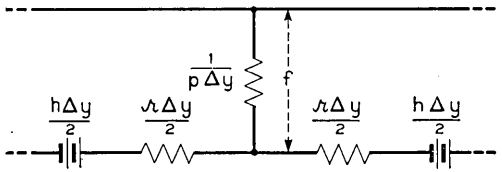
In ordinary electromagnets the core is the part of highest density, in which saturation limits the attainable field. It is good design practice to limit the core section to as small a value as is consistent with satisfactory performance, as this results in a minimum inside coil diameter. This is advantageous with respect to the coil constant, as shown by equation (1).

In the special case of high speed relays, it is advantageous to minimize the mass of the armature and hence its cross-section. In such relays armature saturation may control, or occur concurrently with core saturation. Saturation elsewhere than in the core or armature is of interest only in the diagnosis of faulty design, since the return members should have a section adequate to carry the maximum field at densities well below saturation.

Thus in most electromagnets, saturation occurs in the core, and incipient saturation affects its reluctance and the pattern of the associated field. The magnetomotive force varies along the length of the coil, and the core is therefore subject to variations along its length in magnetic



(a) TRANSMISSION LINE MAGNETIC CIRCUIT



(b) DIFFERENTIAL LINE ELEMENT

Fig. 20 — Transmission line analogy to the field of an electromagnet.

potential, in flux density, and (at high densities) in reluctance per unit length. To treat it as a single circuit element, as in the discussion of Fig. 3, is thus a highly simplified approximation. An understanding of the extent to which this approximation is valid may be gained by considering a more rigorous analysis, in which, as indicated in Fig. 20, the core and return path are treated as analogous to a transmission line with distributed constants.

Transmission Line Analogy

In Fig. 20(a), the core is shown as one side of the line, and the return path as the other side. The length of the line is taken as the length of the winding, which is assumed to be distributed uniformly. It is assumed that the line is terminated at its ends by lumped reluctances, $\mathcal{R}_1$ and $\mathcal{R}_2$. The characteristics of the line and the applied magnetomotive force are expressed in terms of the following quantities:

f = mmf difference between the two sides of the line,
$\mathcal{r}$ = reluctance per unit length of line,
p = leakage permeance per unit length of line,
h = impressed mmf per unit length of line.

Instead of trying to represent the entire magnetic circuit by a simple network of a few lumped reluctances, it is assumed only that an infinitesimal length (dy) of the magnetic line can be represented by the "T"

network shown in Fig. 20(b). The complete magnetic circuit, then, consists of an infinite number of these elementary networks connected end-to-end and terminated in lumped reluctances at the extreme ends. Note that this approximation to the magnetic circuit explicitly recognizes the distributed nature of the leakage flux.

Consider now the several quantities which enter into the approximation. Each of the terminating reluctances ($\mathcal{R}_1$ and $\mathcal{R}_2$) includes two component reluctances in parallel. The first component is simply the leakage reluctance across the end of the line. The second component is the sum of the reluctances of the magnetic members and series air gaps which complete the circuit from one side of the line to the other. For any single value of the working gap, $\mathcal{R}_1$ and $\mathcal{R}_2$ may be taken as constants.

The quantity p, which is the leakage permeance per unit length between the two sides of the line, depends upon the geometry of the structure, and may be calculated by the methods of Section 5. Provided the configuration of the magnetic line is uniform throughout its length, p is taken as a constant. The assumption that p is constant is equivalent to assuming that all leakage paths between the two sides of the transmission line lie in planes that are perpendicular to the core. This condition is not satisfied near the ends of the core, but correction for the end effects may be made in evaluating the terminal reluctances, $\mathcal{R}_1$ and $\mathcal{R}_2$.

The quantity $\mathcal{r}$, the series reluctance per unit length of line, involves magnetic material whose permeability varies with flux density. It is a variable whose magnitude depends upon the applied magnetomotive force, upon the terminating reluctances, and upon position along the line, since it is a function of φ. In the low density region, however, its value is substantially independent of φ. Application of the magnetic circuit relations results in the following equations:

$$\frac{df}{dy} = h - \mathcal{r}\varphi, \tag{20A}$$

$$\frac{d\varphi}{dy} = -pf, \tag{20B}$$

whence:

$$\frac{d^2\varphi}{dy^2} = -p(h - \mathcal{r}\varphi). \tag{21}$$

Subject to the validity of the original assumptions, the solution of equation (21) describes the way in which the flux φ varies with y, the distance along the line.

At low densities, where $\mathcal{R}$ is constant, (21) may be solved to give a picture of the flux distribution in the magnet. The solution is obtained in terms of hyperbolic functions. The resulting expression is somewhat cumbersome, but reduces to a much simpler form in the special case where $\mathcal{R}$ is small compared with h. Neglecting $\mathcal{R}\varphi$, equation (20A) reduces to:

$$\frac{df}{dy} = h \qquad \text{or} \qquad f = hy$$

In this case then, the magnetomotive force varies linearly along the core, as was assumed in the discussion of Fig. 17 in Section 5.

For $f = hy$, equation (20B) becomes:

$$\frac{d\varphi}{dy} = -phy,$$

whence:

$$\varphi = \varphi_0 - \frac{phy^2}{2}, \tag{22}$$

where φ_0 is flux at $y = 0$ (one end), and ph is a constant for a given magnet structure. Thus, for this approximate case, the core flux falls off along its length approximately as the square of the distance from one end. The second term in (22) represents the leakage flux. For the whole core, for which $y = \ell$, the leakage flux is given by $p\,h\ell^2/2$, or by $p\,\ell\mathfrak{F}/2$. The average flux linked per turn, however, is the integral

$$\frac{1}{\ell}\int_0^{\ell} \frac{phy^2}{2}\,dy,$$

corresponding to a leakage reluctance of $p\ell/3$. Hence the factor 3 is used to evaluate the average flux linked per turn in Fig. 17.

From this consideration of the more rigorous transmission line treatment, it is apparent that $\mathcal{R}\varphi$ must be small compared with h for the lumped core and leakage reluctance approximation to be valid. If $\mathcal{R}\varphi/h$ is small, the potential drop in the core, $\mathcal{R}_C\varphi$, is small compared to the applied magnetomotive force $\mathfrak{F}$. In most ordinary electromagnets, the ratio $\mathcal{R}_C/\mathfrak{F}$ is small in the low density region. For sensitive relays with long cores of small cross-section, this is not the case, and the lumped constant treatment is correspondingly limited in accuracy, even at low densities.

At high densities, however, the core reluctance increases even in ordinary electromagnets. To use the transmission line analogy at high densities, it is necessary to express $\mathcal{R}\varphi$ in (20A) in terms of the Froehlich-

Kennelly relation described in Section 4. A solution to the resulting equations can be obtained in series form, but this solution is too complex for convenient use in engineering estimates. From this formulation of the problem, however, it is apparent that the use of lumped values of core and leakage reluctances at high densities can only be a rough approximation, as the reluctance per unit length must vary along the core, and the pattern of the leakage field and hence the leakage reluctance are no longer constant. However, in representing the core reluctance as increasing from its low density value and approaching infinity as $\varphi \rightarrow \varphi''$, the lumped approximation correctly represents the limiting conditions. It therefore provides a rough approximation to the intermediate values.

Series-Parallel Magnetic Circuit

Subject to the limitations discussed above, the magnetic circuit of most ordinary electromagnets can be represented in the form shown in Fig. 21. The core reluctance $\mathcal{R}_C$ is in series with two parallel paths: a leakage path of reluctance $\mathcal{R}_{L2}$, and an armature path of reluctance $\mathcal{R}_{02} + x/A_2$. The subscript 2 is used with the constants of this particular circuit to distinguish them from those of the simpler approximation to be discussed below. The magnetic circuit of Fig. 3 reduces to that of Fig. 21 if the reluctance $\mathcal{R}_A$ of the former is ignored, or considered as part of the reluctance $\mathcal{R}_{02}$.

The circuit reluctance $\mathcal{R}$, or $\mathcal{F}/\varphi$, can be derived from the circuit by the procedure applying to resistances in an electrical circuit, and is given by:

$$\mathcal{R} = \mathcal{R}_C + \frac{\mathcal{R}_{L2}\left(\mathcal{R}_{02} + \dfrac{x}{A_2}\right)}{\mathcal{R}_{L2} + \mathcal{R}_{02} + \dfrac{x}{A_2}}. \tag{23}$$

The evaluation of the constants of this circuit may be described with

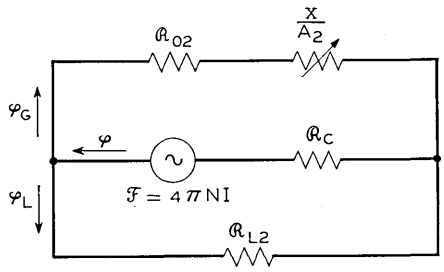


Fig. 21 — Series parallel magnetic circuit — the usual design analogy.

reference to the two structures shown schematically in Fig. 22, where Fig. 22(a) represents the familiar "end-on" armature type of construction, and Fig. 22(b) represents the "flat type" relay of Bell System use. The dashed lines indicate the path along which the length of the parts is measured, while the lined areas are those of the main, heel, and side gaps.

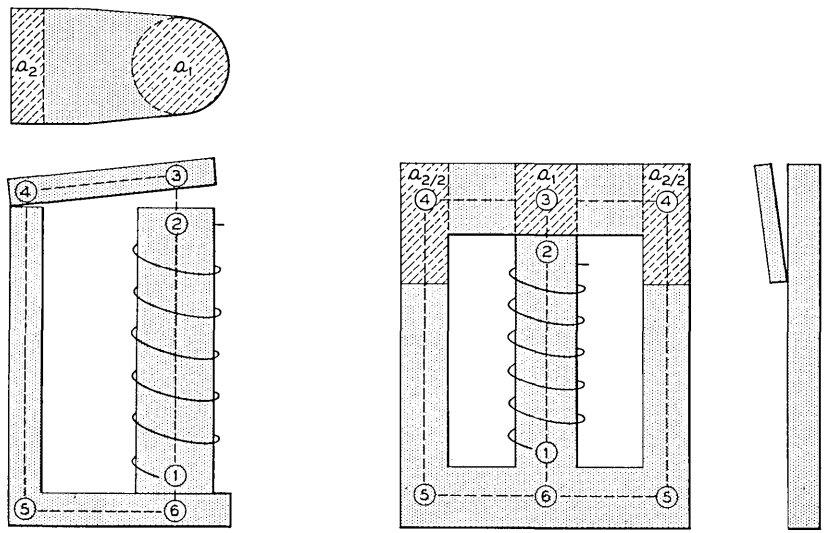
The core reluctance, $\mathcal{R}_C$, is determined from equation (17), taking the length as between the points 1 and 2 in both figures. The permeability is taken at the nominal maximum value throughout all iron parts.

The closed gap reluctance, $\mathcal{R}_{02}$, is the sum of the following:

A. The iron reluctance, determined from equation (17) taking the lengths of the several parts as measured around the path 2–3–4–5–6–1. In each term the cross-sectional area a is that of the part through which the flux passes. In Fig. 22(b) of course, the two side sections are added to give the total section.

B. The contact gap reluctances computed as x/a_1 and x/a_2, taking x as for an air gap of 0.005 cm. The reluctance of the joint at 1 in Fig. 22(a) is computed in the same way, and included in the sum.

C. The "stop pin" air gap reluctance, computed as x/A_2, where A_2 is the effective pole face area as determined below, and x is the separation at the measuring point for the stop pin opening.



(a) END-ON ARMATURE TYPE (b) FLAT TYPE

Fig. 22 — Magnetic circuit components of typical relay structures.

The effective pole face area, A_2, is determined by equation (18), following the procedure described in the discussion of Fig. 12.

The leakage reluctance, $\mathcal{R}_{L2}$, is determined in the case of Fig. 22(a) by the procedure discussed in Section 5 and indicated in Fig. 17. In using this, ℓ_1 is the length 1–2, while ℓ_2 is the length 2–3 in Fig. 22(b) and $\ell_2 = 0$ in Fig. 22(a). The reluctance terms $C_1\ell_2$ and C_2d correspond to the armature leakage reluctance $\mathcal{R}_{LA}$ of Fig. 3, here taken as in parallel with the core leakage reluctance in determining $\mathcal{R}_{L2}$.

It should be noted that this procedure provides for two flux paths across each gap: the flux through the reluctance x/A_2, varying linearly with gap, and the parallel leakage flux through a reluctance calculated as though the armature were absent. This representation allows for the effect of fringing, taking the total field across the gap as the sum of these two fields. It carries the implication that experimentally the two fields cannot be separated by search coil measurements.

For the magnetic circuit of Fig. 21, the low density reluctance is given by equation (23). The reluctance terms are calculated for maximum permeability μ', corresponding to density B', and hence for a total core flux $\varphi' = B'a$. As discussed in Section 4, these values are approximately applicable through the low density region, or for φ less than φ'. For φ greater than φ', $\mathcal{R}_C$ is taken as given by equation (13). Thus in the high density region, the total reluctance may be written as:

$$\mathcal{R} = \mathcal{R}_C + \mathcal{R}_E,$$

where,

$$\mathcal{R}_E = \frac{\mathcal{R}_L\left(\mathcal{R}_0 + \dfrac{x}{A}\right)}{\mathcal{R}_0 + \mathcal{R}_L + \dfrac{x}{A}},$$

and,

$$\mathcal{R}_C = \mathcal{R}_C'' \frac{\varphi''}{\varphi'' - \varphi},$$

in which,

$$\mathcal{R}_C'' = \frac{\ell}{\mu' a} \cdot \frac{\varphi'' - \varphi'}{\varphi''},$$

and $\varphi'' = B''a$, where ℓ and a are the length and cross-section of the core, respectively. The values of B_M in Table I may be taken as estimates of the effective value of the saturation density B''.

Equivalent Magnetic Circuit

Over the low density region, in which the magnetic circuit reluctances are substantially independent of the flux, the expressions for the reluctance can be simplified by a procedure analogous to that used in deriving electrical network equivalents. Provided the gap reluctance varies linearly with the gap, magnetic circuits such as those of Figs. 3 and 21 can be replaced by an equivalent circuit of the simple form shown in

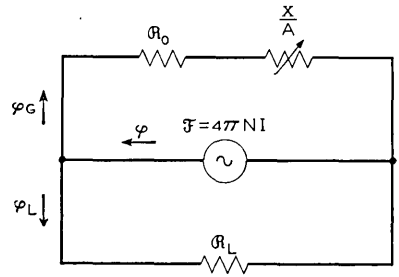


Fig. 23 — Equivalent magnetic circuit.

Fig. 23. For this simple parallel circuit, the total reluctance is given by:

$$\mathcal{R} = \frac{\mathcal{R}_L\left(\mathcal{R}_0 + \frac{x}{A}\right)}{\mathcal{R}_0 + \mathcal{R}_L + \frac{x}{A}}. \tag{24}$$

The simpler subscripts of these *equivalent* values of the magnetic circuit constants are used to distinguish them from the *design* values, applying to the magnetic circuit taken as representing the actual structure. In the usual case the design values apply to the two mesh circuit of Fig. 21, for which the additional subscript 2 is used. When a three mesh circuit is required to represent the structure, the design values are distinguished by the subscript 3, as in the constants of Fig. 3.

In the usual case in which the *design* values apply to the circuit of Fig. 21, expressions for the *equivalent* values may be obtained by comparison of the reluctance given by (23) with that given by (24). The former equation may be written in the form.

$$\mathcal{R} = \mathcal{R}_G + \mathcal{R}_{L2} - \frac{A_2\mathcal{R}_{L2}^2}{A_2\mathcal{R}_{02} + A_2\mathcal{R}_{L2} + x},$$

while (24) may be written.

$$\mathcal{R} = \mathcal{R}_L - \frac{A\mathcal{R}_L^2}{A\mathcal{R}_0 + A\mathcal{R}_L + x}.$$

By comparison, these two expressions are identical for all values of x, provided the following conditions are satisfied:

$$\begin{aligned} \mathcal{R}_L &= \mathcal{R}_C + \mathcal{R}_{L2}, \\ A\mathcal{R}_L^2 &= A_2\mathcal{R}_{L2}^2, \\ A(\mathcal{R}_0 + \mathcal{R}_L) &= A_2(\mathcal{R}_{02} + \mathcal{R}_{L2}). \end{aligned}$$

Upon collecting terms, the relations between design and equivalent circuits are given by:

$$\begin{aligned} \mathcal{R}_L &= \mathcal{R}_C + \mathcal{R}_{L2}, \\ A &= A_2/p^2, \\ \mathcal{R}_0 &= p^2\mathcal{R}_{02} + p\mathcal{R}_C, \end{aligned} \tag{25}$$

where:

$$p = 1 + \mathcal{R}_C/\mathcal{R}_{L2}.$$

Thus the magnetization relations in the low density region can be represented by the reluctance given by (24), corresponding to the simple parallel circuit of Fig. 23, provided the constants are evaluated from those of the design circuit by means of equations (25). The evaluation of the pull and of the field energy in the low density region can therefore be conducted in terms of the simpler relations applying to the equivalent circuit.

As these simpler relations suffice to define the performance, they may be readily evaluated experimentally, using the procedures described in a companion article.[8] While these equivalent constants provide a simple and convenient means for the description and analysis of performance, they are related to the dimensions of the structure only through the conditions of equivalence given by equations (25).

Special Magnetic Circuits

The reluctance of most ordinary electromagnets can be expressed in terms of the series parallel magnetic circuit of Fig. 21, as in the two cases of Fig. 22 discussed above. Differences in configuration may affect the detailed procedure for estimating the constants, but the same circuit schematic applies. There are, however, other magnetic structures requiring different magnetic circuits for their representation. For purposes of illustration, a summary discussion of three such cases is given here.

For *armature saturation*, or cases where the flux density in the armature exceeds that of the core, as in some high speed relays, the magnetic cir-

cuit must be taken as that of Fig. 3. In this case it is the armature flux φ_A, rather than the total flux φ, which approaches a saturation value limiting the mechanical output. In the low density region, where the reluctances are constant, the conditions of equivalence given by equations (25) may be used to evaluate a simple parallel equivalent to the armature path reluctance, reducing the circuit to the form of Fig. 21, and this may in turn be reduced to an equivalent circuit of the form of Fig. 23. In the high density region, however, it is the armature rather than the core reluctance which must be expressed in terms of the Froelich-Kennelly approximation.

For *reed relays*,[6] the structural schematic is as shown in Fig. 24. In some cases, the external shield may be omitted. The magnetic circuit involves an air path through the coil in parallel with a path through the reeds. With a new interpretation of the terms, the circuit of Fig. 3 may be taken as applying, with $\mathcal{R}_C$ taken as zero, and $\mathcal{R}_{LC}$ representing the air leakage field. With some correction for the effect of the shield, this may be estimated by means of the solenoid reluctance expressions of Section 5. The controlling path is that through the reeds, represented by $\mathcal{R}_A$ for the saturating section of maximum density in series with the parallel paths between the reeds: the useful path at the gap and the leakage field. The latter may be estimated from the relations of Fig. 13.

For *polar relays*, the network contains both the permanent magnet magnetomotive force $\mathcal{F}_M$ and the coil magnetomotive force $\mathcal{F}_C$. A typical case is shown in Fig. 25. The constants applying can be evaluated by the procedures described in Section 5, and the somewhat complex circuit equations resulting can be written by the application of Kirchoff's laws in a manner wholly analogous to that for the corresponding electrical circuit.

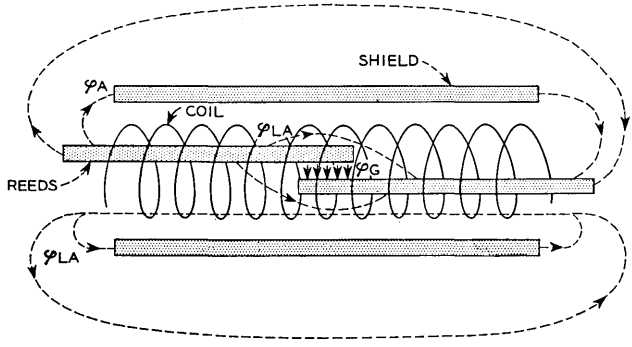


Fig. 24 — Magnetic field of a reed relay.

7 PULL EQUATIONS

Pull in Terms of Gap Flux

With the magnetization relations known, the pull can be determined from equations (7) or (8). Over the low density range, in which the magnetic circuit constants are independent of the flux, the field energy U is given by (5). Substitution of this in (7) gives the following expression for the pull F:

$$F = \frac{\varphi^2}{8\pi} \frac{d\mathfrak{R}}{dx}. \tag{26}$$

In terms of the equivalent circuit constants of Fig. 23, the reluctance $\mathfrak{R}$ is given by equation (24). Substituting this expression in (26) gives the equation:

$$F = \left(\frac{\mathfrak{R}_L}{\mathfrak{R}_0 + \mathfrak{R}_L + \dfrac{x}{A}} \right)^2 \frac{\varphi^2}{8\pi A}.$$

By comparison with (24), it can be seen that the bracketed term is the ratio $\mathfrak{R}/(\mathfrak{R}_0 + x/A)$, which equals the ratio of the gap flux to the total flux φ. Hence the preceding expression may be written in the form:

$$F = \frac{\varphi_G^2}{8\pi A} \tag{27}$$

By a parallel treatment it can be similarly shown that the application of equation (26) to the magnetic circuits of Fig. 3 and 21 gives expressions identical with (27) except that A is replaced by A_3 in the former case and by A_2 in the latter. Equation (27) is the familiar expression for

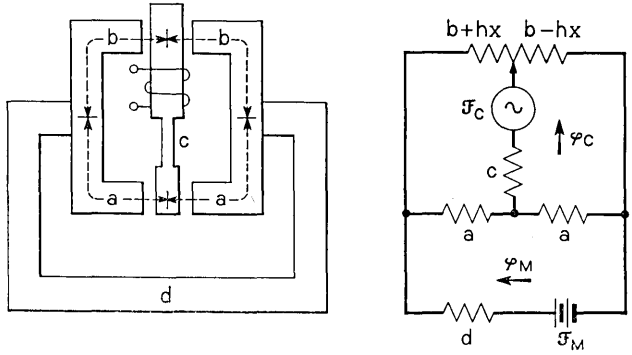


Fig. 25 — Magnetic circuit of a polar relay.

magnetic pull on two parallel planes with a field of uniform density φ_G between them. It can be directly derived from the fact that the gap energy is $\mathfrak{F}\varphi_G/(8\pi)$, with $\mathfrak{F} = x\varphi_G/A_2$. The change in this energy for a differential change dx of the gap equals $F\,dx$, so that F is given by equation (27). Derived in this way, it is apparent that (27) depends only on the field in the gap, and is quite independent of the density in the magnetic material.

Pull in Terms of Applied mmf

In the low density region, where $\mathfrak{R}$ is substantially a function of x only and the magnetization curves are linear, $W = U$, and hence the expression for F given by equation (8) becomes:

$$F = 2\pi(NI)^2 \frac{d}{dx}\left(\frac{1}{\mathfrak{R}}\right). \tag{28}$$

This expression for the pull in the linear region is known as the equation of Perrot and Picou.[9] As $NI = \mathfrak{R}\varphi/(4\pi)$, it is identical with (26).

For the magnetic circuit of Fig. 23, substitution in (28) of the expression for $\mathfrak{R}$ given by (24) gives the following expression for the pull:

$$F = \frac{2\pi(NI)^2}{A\left(\mathfrak{R}_0 + \frac{x}{A}\right)^2}. \tag{29}$$

In the low density region, the reluctance can always be expressed in terms of the equivalent values of Fig. 25. Using the equivalent values of closed gap reluctance $\mathfrak{R}_0$ and of pole face area A, the pull is given by equation (29). This expression is therefore of general application in the low density region, and is the most convenient equation to use for this purpose.

High Density Pull

It was shown above that equation (27) can be derived directly from the expression for the field energy associated with the gap. This expression is quite independent of the reluctance in the rest of the magnetic circuit, and is therefore equally applicable in the low and high density regions. This essentially physical argument shows that (26) and (27) are applicable through the full range of magnetization.

The same result can be obtained from equation (7) by substituting in it the expression for U given by (3), and substituting $\mathfrak{R}\varphi/(4\pi)$ for NI.

There is thus obtained:

$$F = \frac{\partial}{\partial x} \int_0^{\varphi} \frac{\mathfrak{R}\varphi}{4\pi} d\varphi.$$

For saturation confined to the core, as in the series-parallel circuit of Fig. 21, $\mathfrak{R} = \mathfrak{R}_C + \mathfrak{R}_E$. On substituting this expression for $\mathfrak{R}$ in the preceding equation, $\mathfrak{R}_C$ is independent of x and $\mathfrak{R}_E$ is independent of φ, so that the expression becomes:

$$F = \frac{\varphi^2}{8\pi} \frac{d\mathfrak{R}_E}{dx},$$

or, for the series parallel circuit of Fig. 21

$$F = \frac{(\mathfrak{R}_{L2}\varphi)^2}{8\pi A_2 \left(\mathfrak{R}_{02} + \mathfrak{R}_{L2} + \frac{x}{A_2} \right)^2}. \tag{30}$$

Evidently this same expression should apply equally to the equivalent circuit of Fig. 23. That this is the case can be shown by substitution of the equivalence conditions of equations (25), giving an expression for the pull identical with (30) except that the magnetic circuit constants $\mathfrak{R}_{02}$, $\mathfrak{R}_{L2}$, and A_2 are replaced by their equivalent values $\mathfrak{R}_0$, $\mathfrak{R}_L$, and A.

Thus when saturation occurs in the core, as in most ordinary electromagnets, (30) gives the pull through the full range of magnetization, and the expression is invariant to a change from the design constants of Fig. 21 to the equivalent constants of Fig. 23. If φ is taken as φ'', the saturation flux, (30) gives the upper limit to the attainable pull. It is also useful, as shown in the companion article[1] on relay speed, in estimating the pull effective in rapid operation, when φ may be nearly constant during the latter stage of armature motion.

In estimating the steady state pull characteristics, equation (29) is applicable throughout the low density range. In the high density range, the pull may be determined from (27), or, provided saturation occurs in the core, from (30). In cases of armature saturation, the pull must be determined from (27), the expression of most general application.

8 SENSITIVITY AND WORK CAPACITY

Ampere Turn Sensitivity

Aside from any margin for rapid operation, the pull of the electromagnet must exceed the static mechanical load at all values of gap x. As illustrated in Fig. 1, this may be studied graphically by comparing

the load curve with the family of pull curves for various values of NI. The just operate ampere turn value is that for the pull curve which just exceeds the load curve at all points. For design purposes, estimated pull and load characteristics must be used in this comparison, and conditions determined for minimizing the ampere turns required to operate a specified load.

The work V done against the load is represented by the area under the load curve. Its maximum possible value is equal to the work W done by the magnet, represented by the area under the pull curve. Within the region of substantially linear magnetization, the pull is given by (29), which may be written in the form:

$$F = \frac{F_o}{(1 + u)^2}, \tag{31}$$

where:

$$F_0 = \frac{2\pi(NI)^2}{A\mathfrak{R}_0^2}, \tag{32}$$

and $u = x/(A\mathfrak{R}_0)$, or the ratio of the gap reluctance x/A to the closed gap reluctance $\mathfrak{R}_0$.

The work W done by the magnet in the travel from an initial gap x_1 to $x = 0$ is the integral of $F\,dx$, or of $A\mathfrak{R}_0\,F\,du$. From (31), this integral is given by:

$$W = \frac{u_1}{1 + u_1} W_{\max}, \tag{33}$$

where $u_1 = x_1/(A\mathfrak{R}_0)$ and $W_{\max} = A\mathfrak{R}_0F_0$. For a large initial gap (u_1 large), W approaches $W_{\max}$, which therefore measures the upper limit to the output for this ampere turn value; the total area under the pull curve. In general, the difference between the force displacement characteristics of the load and pull curves permits only a fraction of $W_{\max}$ to be used to operate the load. The potential output of the magnet, however, depends solely upon $W_{\max}$. From (32), in which $A\mathfrak{R}_0F_0$ equals $W_{\max}$, the ampere turn sensitivity, or potential output for a given ampere turn value, is given by:

$$\frac{W_{\max}}{(NI)^2} = \frac{2\pi}{\mathfrak{R}_0}. \tag{34}$$

Thus the ampere turn sensitivity depends solely upon the closed gap reluctance $\mathfrak{R}_0$ throughout the region of linear magnetization.

Sensitivity for Specific Loads

The part of W_{max} that can be realized depends upon the shape of the load curve. It also depends upon the travel through which the load must be moved. In principle, the latter may be adjusted to any desired value by the choice of a lever arm determining the ratio of the travel at the point of load actuation to the travel at the point where x has been measured in determining the magnetic circuit constants. The use of a lever arm linkage, of course, changes the force required, but the work V and the shape of the load characteristic remain unchanged.

In Fig. 26 is shown the pull curve relation of equation (31), together with two simple load curves: a constant load and one varying linearly with travel. In the first case, let F_1 be the constant load required, and x_1 the required travel. The pull curve to operate this load must have $F = F_1$ at $x = x_1$. Hence, from (31), $F_1/F_0 = 1/(1 + u_1)^2$. The work V_c done against the load is $F_1 x_1$, and the ratio of V_c to W_{max} is therefore $F_1 x_1/(A\mathcal{R}_0 F_0)$, or $u_1 F_1/F_0$. Hence the part of W_{max} realized against a constant load is given by:

$$\frac{V_c}{W_{max}} = \frac{u_1}{(1 + u_1)^2}. \tag{35}$$

The ratio V_c/W_{max} is shown plotted against u_1 in Fig. 27. Its maximum value is at $u_1 = 1$, where the travel $x_1 = A\mathcal{R}_0$. Maximum sensi-

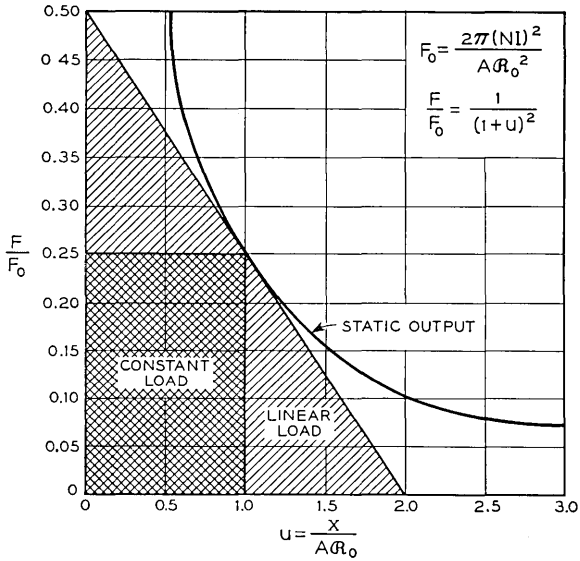


Fig. 26 — Relations between load and pull curves.

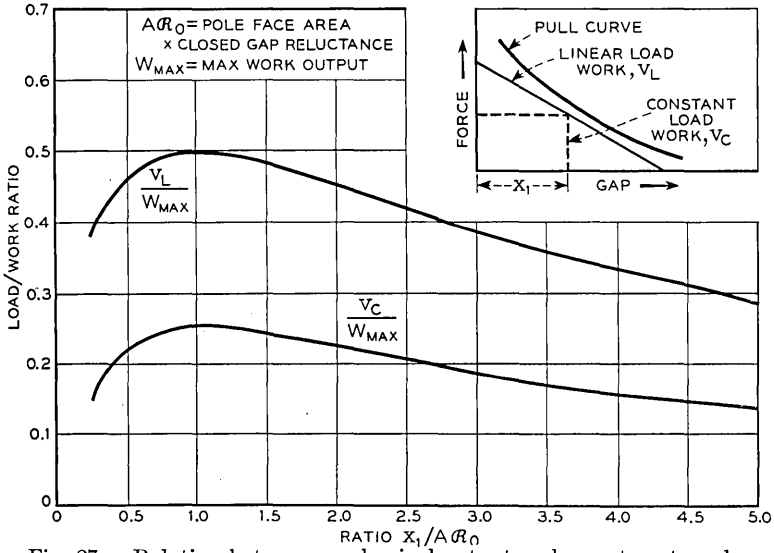


Fig. 27 — Relation between mechanical output and armature travel.

tivity is attained if the lever arm is chosen to satisfy this condition, and in this case $V_C = W_{\max}/4$.

For the case of a linear load, varying from F_2 at $x = 0$ to zero at $x = x_2$, the load is given by:

$$F = F_2\left(1 - \frac{u}{u_2}\right),$$

where $u_2 = x_2/(A\mathcal{R}_0)$. For the pull curve to be tangent to this load curve, the values of both F and dF/dx given by this last equation and by (31) must be equal at the point of tangency, x_1. These two conditions give two expressions for the ratio F_2/F_0. Equating these, there is obtained the following equation for the point of tangency:

$$u_1 = \frac{2u_2 - 1}{3} \tag{36}$$

The work done against the load is $F_2x_2/2$, and the ratio of this to $W_{\max}$ is therefore $F_2x_2/(2A\mathcal{R}_0F_0)$. Substituting the expression for F_2/F_0 obtained as described, and the expression for u_1 given by (36), there is obtained the following expression for the fraction of $W_{\max}$ realized against a linear load:

$$\frac{V_L}{W_{\max}} = \frac{(3u_1 + 1)^2}{4(u_1 + 1)^3}. \tag{37}$$

The ratio V_L/W_{max} is shown plotted against the point of tangency u_1 in Fig. 27. Its maximum value is at $u_1 = 1$, when $V_L = W_{max}/2$. For this optimum condition, $u_2 = 2$, so that the total travel, x_2, equals $2A\mathcal{R}_0$, while the point of tangency, x_1, is at half that travel. For this linear load case, under the optimum condition, the maximum work that can be usefully applied is, then, $V_L = W_{max}/2$.

Actual load curves seldom conform to either of the simple cases of Fig. 26, and more commonly have the irregular character illustrated in Fig. 1. Most of them, however, show a point of closest approach to the pull curve either at the junction of two segments, as in Fig. 28(a), or at a point of tangency to a linear segment, as in Fig. 28(b). In the former case, the coordinates of the junction point may be taken as F_1 and x_1, as indicated, and the relations for a constant load for which $V_C = F_1x_1$ applied. In the other case, the tangent segment may be extended, as indicated, to intersect the axes at F_2 and x_2, and the relations for a linear load for which $V_L = F_2x_2/2$ applied.

Quite generally, therefore, the work V done against the load curve is proportional to W_{max}, and the proportionality constant is some function of $u_1 = x_1/(A\mathcal{R}_0)$, where x_1 is the point of closest approach of the load and pull curves. Writing $f(u_1)/(2\pi)$ for the ratio V/W_{max}, it follows from equation (34) that the ampere turn sensitivity, $V/(NI)^2$ is given by:

$$\frac{V}{(NI)^2} = \frac{f(u_1)}{\mathcal{R}_0}, \tag{38}$$

where $f(u_1)$ is a maximum for $u_1 = 1$, and is similar to the curves of Fig. 27. For the particular load conditions represented by constant load

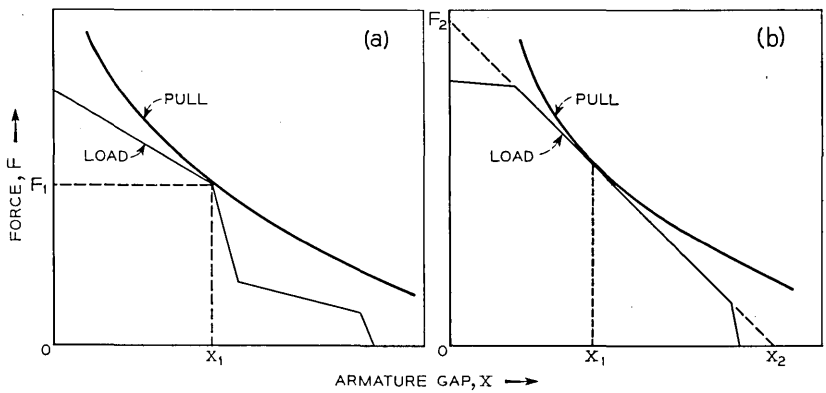


Fig. 28 — Determination of pull curve to match load.

(V_C) and linear load (V_L) characteristics, the highest values of useful work were shown to be $\pi(NI)^2/(2\mathfrak{R}_0)$, and $\pi(NI)^2/\mathfrak{R}_0$, respectively.

These relations establish the important design requirement that the point of closest approach of the load and pull curves should be at $u_1 = 1$, or where the gap reluctance x_1/A equals the closed gap reluctance $\mathfrak{R}_0$. To the extent permitted by space requirements and manufacturing considerations, this may be met by a proper choice of lever arm ratio or of pole face area. In the latter case $\mathfrak{R}_0$ is not a wholly independent parameter but includes a term varying with A, so that allowance must be made for the change in $\mathfrak{R}_0$ in changing A. The curves of Fig. 27 show that $V_C/W_{\max}$ and $V_L/W_{\max}$ decrease slowly for values of u_1 between 1 and 2. The ampere turn sensitivity is therefore close to its maximum value if u_1 lies in this range.

These relations are particularly useful in preliminary design estimates, as they permit the ampere turn requirement to be estimated for a known load merely from estimates of $\mathfrak{R}_0$ and A. For this purpose it is only necessary to determine $V/W_{\max}$ by the procedure outlined above. With V known, this determines $W_{\max}$, which equals $A\mathfrak{R}_0F_0$. Thus F_0 can be evaluated, and NI determined from equation (32).

Core Cross-Section

These relations are formally applicable only in the range of linear magnetization. For the estimates to be valid, however, it is only necessary that this condition be satisfied at the point of closest approach of load and pull curves, as an approach to saturation at smaller gaps will reduce the pull in a region where, in most cases, it is well in excess of the load curve.

For linear magnetization to obtain at the point of closest approach ($u = u_1$), the core flux should not materially exceed φ', or aB', where B' is the density for maximum permeability and a is the core cross-section. The flux is equal to $4\pi NI/\mathfrak{R}(u_1)$, and the required value of a is therefore given by:

$$a = \frac{4\pi NI}{B'\mathfrak{R}(u_1)}. \tag{39}$$

Here the values of NI and u_1 applying are those determined as described above, and $\mathfrak{R}(u_1)$ is given by equation (24) for $x = u_1A\mathfrak{R}_0$. To determine the core section a needed therefore requires an estimate of $\mathfrak{R}_L$, as well as of $\mathfrak{R}_0$ and A.

If the expression for NI given by (38) is substituted in (39), it is apparent that a varies as the square root of $V\mathfrak{R}_0/f(u_1)$ and inversely as

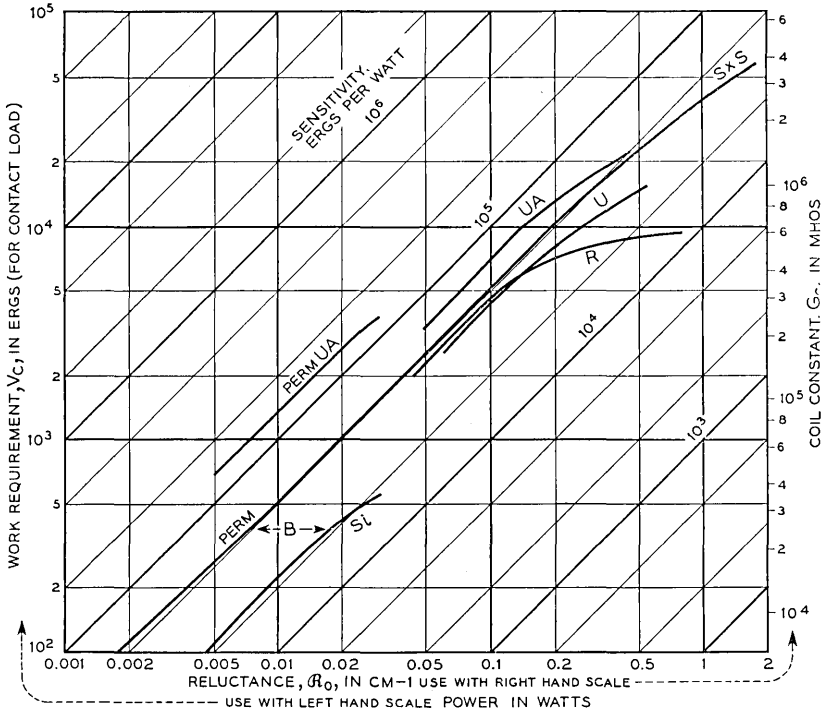


Fig. 29 — Work-power relations in terms of magnetic and coil constants.

$\mathcal{R}(u_1)$. The latter is approximately proportional to $\mathcal{R}_0$, so the core section a varies approximately as $1/\sqrt{\mathcal{R}_0}$, and hence increases with increasing ampere turn sensitivity as well as with increasing load. For values of u_1 near unity, where $f(u_1)$ is nearly constant, a decreases as u_1 increases.

Power Sensitivity

The power sensitivity is the ratio of the work V done against the load to the steady state power I^2R. From equations (1) and (38), this is given by:

$$\begin{aligned} \frac{V}{I^2R} &= G_c f(u_1)/\mathcal{R}_0 , \\ &= \frac{eS}{\rho m^2} \frac{f(u_1)}{\mathcal{R}_0} , \end{aligned} \tag{40}$$

where e/ρ is a constant determined by the coil construction, S is the coil volume, and m the mean length of turn. Depending on the coil depth,

the latter reflects the inside coil dimensions, which vary with the required core section a. The variation of the latter with u_1 discussed above has a minor effect on the power sensitivity, partially offsetting the effect of the term $f(u_1)$, so that $V/(NI)^2$ is nearly independent of u_1 in the range $1 \leqq u_1 \leqq 2$.

Ignoring these secondary variations, equation (40) shows that the steady state power required to operate a given load varies directly as the closed gap reluctance $\mathcal{R}_0$, and inversely as S/m^2, where S is the coil volume and m the mean length of turn. Fig. 29 shows the relationship between the work that can be done, and the steady state power that must be supplied to the coil at this time. Since, in ordinary practice, use cannot be made of the entire area under the force-deflection curve, the "constant-load work," V_c is often used in comparing magnet designs. The maximum value for this term, as determined for a number of Western Electric relays, is shown in the figure. According to equation (40) the ratio of V_c to I^2R should here have the value

$$\frac{V_c}{I^2R} = \frac{\pi G_c}{2\mathcal{R}_0},$$

since for constant-load output $f(u_1) = \pi/2$, when the optimum work condition has been chosen.

Work Capacity

In applications where maximum sensitivity is not required, higher output may be obtained by operating the electromagnet through much of its travel in the non-linear region. The controlling factor is the flux developed at the point of closest approach of the load and pull curves, so the attainable output is controlled by the pull curve for constant flux, as given by equation (30). In what follows the flux will be taken as the saturation flux φ'', corresponding to the upper limit of attainable output. The same relations may be used, however, to estimate the output attainable with any value of flux approaching φ'' in magnitude. For $\varphi = \varphi''$, (30) may be written in the form:

$$\frac{F}{F_0''} = \frac{C_L^2}{(C_L + u)^2}, \tag{41}$$

where $u = x/(A\mathcal{R}_0)$, $C_L = (\mathcal{R}_0 + \mathcal{R}_L)/\mathcal{R}_0$, and F_0'' is given by:

$$F_0'' = \frac{(C_L - 1)^2\varphi''^2}{8\pi AC_L^2}. \tag{42}$$

Equation (41) is of the same form as equation (31), with u replaced

by u/C_L, and F_0 by F_0''. Integration therefore gives an expression for the work W similar to equation (33), with u_1 replaced by u_1/C_L and W_{max}, or $A\mathcal{R}_0F_0$, replaced by $W_{sat.}$, or $C_LA\mathcal{R}_0F_0''$. There is thus obtained:

$$W = \frac{u_1}{C_L + u_1} W_{sat.} \tag{43}$$

As (43) is of the same form as (33), with u replaced by u_1/C_L, the relations between the load and pull curves are the same in the two cases. Thus the ratios $V_C/W_{sat.}$ and $V_L/W_{sat.}$ may be obtained from (35) and (37), or from Fig. 27, by using u_1/C_L to replace u. Maximum output is thus obtained when $u_1/C_L = 1$, and the gap reluctance x_1/A at the point of closest approach equals C_L. Thus the optimum leverage for maximum output differs from that for maximum sensitivity by the factor C_L.

9 DISCUSSION

Applications

The applications of the magnetization relations considered in this article are confined to the static characteristics: the sensitivity and work capacity attainable when no specific timing requirements are imposed. As the same relations control both the electrical and mechanical response of an electromagnet under dynamic as well as under static conditions, the material outlined here has further applications in other aspects of magnet performance, as illustrated in the companion articles appearing in this issue of the JOURNAL.[1,10]

The sensitivity and work capacity discussed above relate essentially to the operate characteristics. The other static characteristics of relay performance are those for release, and for marginal performance. The release characteristics are primarily dependent upon the operated load in relation to the pull at the closed gap. The rising pull characteristic as the gap is closed tends to give an operated pull well in excess of the operated load, giving a release ampere turn value small compared with the operate value. The coercive force tends to further decrease the release value, and imposes the need for stop pins to assure release. A high ratio of release to operate can be attained by providing a rising load characteristic to match the pull, by using high stop pins (thus increasing $\mathcal{R}_0$), or by using a core or armature section that saturates in the latter part of the travel. A detailed discussion of the provisions for marginal operation is outside the scope of the present article, but obviously in-

volves application of the same relations that apply to the operate performance.

The procedures outlined for the estimation of sensitivity and work capacity, or of the magnetic parameters required to give a desired performance in these respects, have their primary use in preliminary design. They make it possible to determine from the performance requirements definite dimensional criteria which must be satisfied by the design. After models have been made and measured, the load characteristics may be compared directly with the observed pull. Study of the magnetization relations, and comparison of observed and estimated magnetic circuit constants, however, is of value through the whole course of development, and even in the engineering use of an established design. Such study relates the performance directly to the dimensions and materials used. In particular it permits extrapolation of the performance observed on particular models, providing estimates of the range of performance variation corresponding to the range of tolerances in dimensions and material properties. An essential phase of such studies is the experimental determination of the magnetic circuit constants by the methods discussed in a companion article.[8]

Validity

The solutions to the static field equations based on the magnetic circuit treatment are inherently approximations. The discussion of this subject in many engineering texts implies that the approximation is rather crude. In principle, as shown above, the magnetic circuit method may be applied to as close an approximation as desired, subject to the accuracy with which the pattern of the field can be recognized.

The criterion of satisfactory estimation of the magnetic circuit parameters is the ability to predict the magnetization relations actually observed. Analysis of observed magnetization relations by the method described in the companion article[8] has shown that these relations can be satisfactorily represented by the magnetic circuit relations given here. Satisfactory agreement has been found between observed and estimated values of the magnetic circuit parameters when the latter are determined by the procedures described above. These conform in large measure to those initiated by Evershed,[3] and are distinguished principally by explicitly recognizing the existence of leakage paths shunting any air gaps, and by recognizing that these cannot be identified by search coil measurements, so that the field between two members bounding a gap must be regarded as the vector sum of a constant and a variable field.

Design Considerations

The relations controlling the sensitivity and work capacity provide a direct indication of the effect of the configuration and dimensions of an electromagnet upon its performance. Here these will be considered only in general terms. Other things being equal, the sensitivity requirement determines the coil size, and hence the over-all dimensions, while the work capacity determines the cross-sections of the magnetic path. Within the framework of these generalizations, the dependence of performance on size and shape can be most readily considered in terms of the individual magnetic parameters.

The core reluctance $\mathcal{R}_C$ is characterized primarily by its minimum value $\mathcal{R}'_C$, and by the saturation flux φ'' which sets the upper limit to the work capacity. φ'' varies with the core cross-section a, as does φ', the flux level for minimum reluctance. $\mathcal{R}'_C$ is given approximately by $\ell/(\mu' a)$, where ℓ is the core length, which varies with the coil size. For heavy duty relays with a large cross-section a, $\mathcal{R}_C$ is a minor component of the total reluctance except for φ very close to saturation. For highly sensitive light duty relays, ℓ is large and a small, making $\mathcal{R}_C$ large, and a must be made large enough for φ to be close to φ', so that $\mathcal{R}_C$ may be close to its minimum value $\mathcal{R}'_C$.

The leakage reluctance $\mathcal{R}_L$ depends upon both the dimensions and the shape of the magnetic path. The leakage per unit length depends upon the ratio of the perimeter of the magnetic members to their separation, and is a minimum for a square magnet outline as in the two coil telegraph type relay. Space and accessibility considerations cause telephone relays to have a long narrow single coil configuration, increasing the leakage per unit length by a factor of from two to four over that of the square outline. Aside from this shape factor, the leakage reluctance varies inversely as the length. With respect to the static characteristics the leakage flux (1) increases the core cross-section required to attain a given work capacity, (2) decreases the equivalent pole face area, as shown by equations (25), (3) increases the equivalent closed gap reluctance $\mathcal{R}_0$, as shown by the same equations, thus decreasing the sensitivity. In addition, as discussed in the companion articles[1,10] it delays operation and release by increasing the total field linked by the coil. In all these effects, the controlling factor is the ratio $\mathcal{R}_L/\mathcal{R}_0$, rather than the absolute value of $\mathcal{R}_L$, so these effects are most readily minimized by making $\mathcal{R}_0$ small.

The closed gap reluctance $\mathcal{R}_0$ is the sum of the reluctances of the return members and of the closed gaps and joints. The reluctance of the return members is comparable with that of the core but smaller, and in most

cases minor compared with that of the gaps and joints. The latter are equivalent to an air gap of 0.005 cm (2 mil-in) over the area of the joint, giving a marked advantage to one-piece construction of core and return members. The heel and main gaps necessarily introduce reluctances of similar magnitudes, further increased at the main gap by the height of the stop pins required to reduce the residual flux to the release level. It is therefore advantageous to use large areas for both heel and main gaps. When a small value of $\mathfrak{R}_0$ is thus obtained, providing high sensitivity, its value is the more sensitive to variations in fit and alignment at the heel and main gaps, and high sensitivity therefore requires close tolerances on the dimensions controlling the fit of the armature to the pole pieces. Heavy section magnets tend to high sensitivity, partly because of the reduced reluctance of the magnetic members, but principally because heavy sections facilitate the provision of large areas at the gaps and joints.

As shown above, the effective pole face area is the reciprocal of the coefficient of x in the expression for the variable reluctance term. It therefore depends upon both the heel and main gaps, though the contribution from the heel gap is small when the armature is hinged there. For optimum sensitivity, or optimum work capacity, there are optimum values of the gap reluctance x_1/A as shown above, which can be obtained either by a choice of leverage to the load or by a choice of pole face area. It is preferable to attain these optima by varying the leverage, using as large a pole face area as possible, in order to make $\mathfrak{R}_0$ small.

10 SUMMARY

The material given in this article provides a basis for relay design through the relationships between mechanical output and electrical input. From the indicated relations, one may find the mechanical work from a given magnet design, or find the magnetic design needed to provide a required mechanical output. Relationships for gaining optimum performance are also given.

The first step was to show that the mechanical work depends upon the field energy of the magnet, which is a consequence of the magnetomotive force provided. The magnetomotive force in turn is furnished by the electrical circuit, being determined jointly by the number of turns in the coil and the circuit resistance. The electrical output and the magnetic input were thus equated as

$$I^2R = \frac{\mathfrak{F}^2}{16\pi^2 G_c},$$

where $G_c = N^2/R$, and is shown to be dependent on the coil dimensions and the conductivity of the wire in the winding.

The magnetic field energy is described by "magnetic flux" which, averaged for the entire magnet, is related to magnetomotive force through

$$\mathfrak{F} = \varphi\mathfrak{R},$$

where $\mathfrak{R}$ is an expression accounting for the dimensions and materials of the magnet and its air gaps, called "magnetic reluctance." The validity of this magnetic circuit concept is discussed at some length, leading to the proof that for many cases a very simplified "equivalent two-mesh circuit" may be used to represent the more complicated actual cases. As a result, performance may be expressed in terms of the equivalent values: $\mathfrak{R}_0$, "closed gap reluctance"; $\mathfrak{R}_L$, "leakage reluctance"; and A, "pole face area." Values for these terms may be estimated for cases of initial design (magnet synthesis), or be measured so as to characterize completed models (magnet analysis). Evaluation of reluctances of various magnetic circuit components is described, with numerous relations given for gaps and for magnetic materials, including an approximate relation covering the non-linear behavior of ferrous materials.

In terms of the magnetic circuit variables, force and work may then be expressed as

$$F = \frac{\mathfrak{F}^2}{8\pi A\mathfrak{R}_0^2}\,\frac{1}{(1+u)^2},$$

$$W = \frac{\mathfrak{F}^2}{8\pi\mathfrak{R}_0}\,\frac{u}{1+u},$$

where $u = x/A\mathfrak{R}_0$, the ratio of air gap to the product $A\mathfrak{R}_0$. It is thus found that greatest magnet output for force systems common in relays is obtained when the critical load is picked up at a gap $x_0 = A\mathfrak{R}_0$, which may be accomplished by choice of lever arm. The maximum work that may thus be considered useful, $V_{\max}$, depends somewhat on the particular load characteristic, and for the typical constant-load case, V_c, is related to ampere turns, and to power input as follows:

$$V_c/(NI)^2 = \pi/2\mathfrak{R}_0,$$

$$V_c/I^2R = \pi G_c/2\mathfrak{R}_0.$$

Greatest useful magnet output is thus seen to depend directly on G_c, the "coil constant," and inversely on $\mathfrak{R}_0$, the equivalent closed gap reluctance.

Additional cases are given permitting design relations to be extended into the saturated range for the iron.

REFERENCES

1. Estimation and Control of Operate Time of Relays, Part I — Theory, R. L. Peek, Jr., page 109 of this issue. Part II — Applications, M. A. Logan, page 144 of this issue.
2. A. C. Keller, A New General Purpose Relay for Telephone Switching Systems, B.S.T.J., **31,** p. 1023, Nov. 1952.
3. S. Evershed, Permanent Magnets in Theory and Practice, Institute of Electrical Engineers Journal (England) **58,** 1919–1920.
4. H. C. Roters, *Electromagnetic Devices*, John Wiley and Sons, New York, 1941.
5. S. P. Thompson and K. W. Moss, Proc. Physical Society of London, **21,** p. 622, 1909.
6. W. B. Ellwood, Glass Enclosed Reed Relay, Elec. Eng. **66,** p. 1104, Nov., 1947.
7. E. B. Rosa and F. W. Grover, Bulletin Bureau of Standards, No. 169, 1912.
8. R. L. Peek, Jr., Analysis of Measured Magnetization and Pull Characteristics, page 79 of this issue.
9. Perrot and Picou, Comptes Rendus, **175,** Dec. 9, 1922.
10. R. L. Peek, Jr., Principles of Slow Release Relay Design, page 187 of this issue.

Analysis of Measured Magnetization and Pull Characteristics

By R. L. PEEK, Jr.

(Manuscript received September 21, 1953)

It is shown in this article that the observed magnetization relations of most ordinary electromagnets conform to simple expressions which can be interpreted as the flux vs. magnetomotive force equations of a reluctance network, analogous to the current-voltage equations of a resistance network. To the extent of such conformity the magnetic circuit constants characterizing the network suffice for the evaluation of the field energy and pull characteristics of the electromagnet. The agreement between the observed magnetization and these simple relations is close in the region of linear magnetization, and is adequate for engineering purposes at higher flux densities, but the extent of agreement in the latter range varies with the type of structure and the location of the magnetic parts which first approach saturation.

Specific analytical and graphical procedures are given for the evaluation of the magnetic circuit constants from both pull and magnetization measurements. These procedures employ relations which give linear plots indicating the degree of conformity of the observed relations to the expressions used to fit them. The relation of the measured constants to those which can be estimated in design[1] is discussed, as is the use and application of the measured constants in development and engineering studies.

1 INTRODUCTION

In the design of telephone relays and similar switching apparatus, the characteristics of the electromagnet which serves as motor element may be distinguished from those of the mechanical system of contact springs and actuating members which it operates. The performance of the electromagnet is characterized statically by the mechanical work done for a given coil energization, and dynamically by the time required to actuate the mechanical load, including in this the inertia of the moving parts.

Both the potential work output of the electromagnet and the energy stored in developing its field can be evaluated from its magnetization

relations. The consequent dependence of the pull and timing characteristics upon the magnetization relations makes the measurement and analysis of the latter fundamental to the understanding of performance and its relation to design.

Procedures are described in this article for the evaluation from measured magnetization relations of a few parameters which suffice for the determination of the pull and of the field energy under given conditions. These parameters, the magnetic circuit constants, characterize the electromagnet to which they apply. They are used as measures of performance in comparing different structures, or in studying the effect of dimensional or material variations in a particular design.

In addition to their significance as parameters summarizing measured magnetization relations, the magnetic circuit constants may be interpreted as the observed values of quantities postulated in the magnetic circuit approximation to static field theory. This approximation is discussed in a companion article,[1] which describes methods of estimating the values of the magnetic circuit constants from the configuration, dimensions, and materials constant of the electromagnet. Comparison of observed and estimated values of these constants has served as a guide in developing these methods of estimation. A complete design methodology is provided by the ability to both estimate and measure the magnetic circuit constants characterizing the performance of electromagnets.

In the experimental evaluation of the magnetic circuit constants described in this article, the basic method is that employing measured magnetization relations. Because of the dependency of the pull upon the magnetization relations, pull measurements may also be used, subject to certain limitations, for the evaluation of the magnetic circuit constants. The article includes description of procedures for doing this.

The notation used in this article conforms to the list that is given on page 257.

2 MAGNETIZATION RELATIONS

The magnetization relations of an electromagnet give the average flux φ linked per turn of the winding as a function of the two determining variables: applied ampere turns NI and armature position x. The relations are usually shown, as in Fig. 1, as a family of curves giving φ versus NI for various values of x. Armature motion is usually rotary, and the choice of the location at which x is measured is a matter of convenience. The curves of Fig. 1 apply to the electromagnet shown in Fig. 2, the AJ (heavy duty) type of wire spring relay described in an article

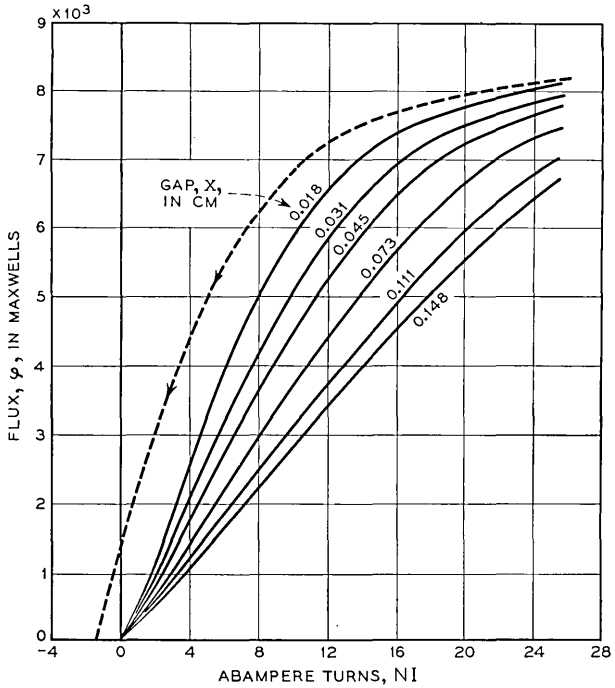


Fig. 1 — Magnetization curves.

by A. C. Keller.[2] In this case it is convenient to measure x at the center line of the actuating card, whose location with respect to the axis of rotation is indicated in Fig. 2.

Apart from interpretation in terms of magnetic theory, any observed point on a magnetization curve represents a measurement of the electrical energy U stored in the coil, as given by the integral

$$\int_0^I ei\, dt,$$

where $e = d(N\varphi)/dt$. It is represented by the area between the magnetization curve and the axis of φ, and is given by:

$$U = \int_0^{\varphi} NI\, d\varphi. \tag{1}$$

This stored field energy may, in principle, be recovered in the decay of the field, except for the portion lost by hysteresis in the material, represented by the area between the increasing and decreasing mag-

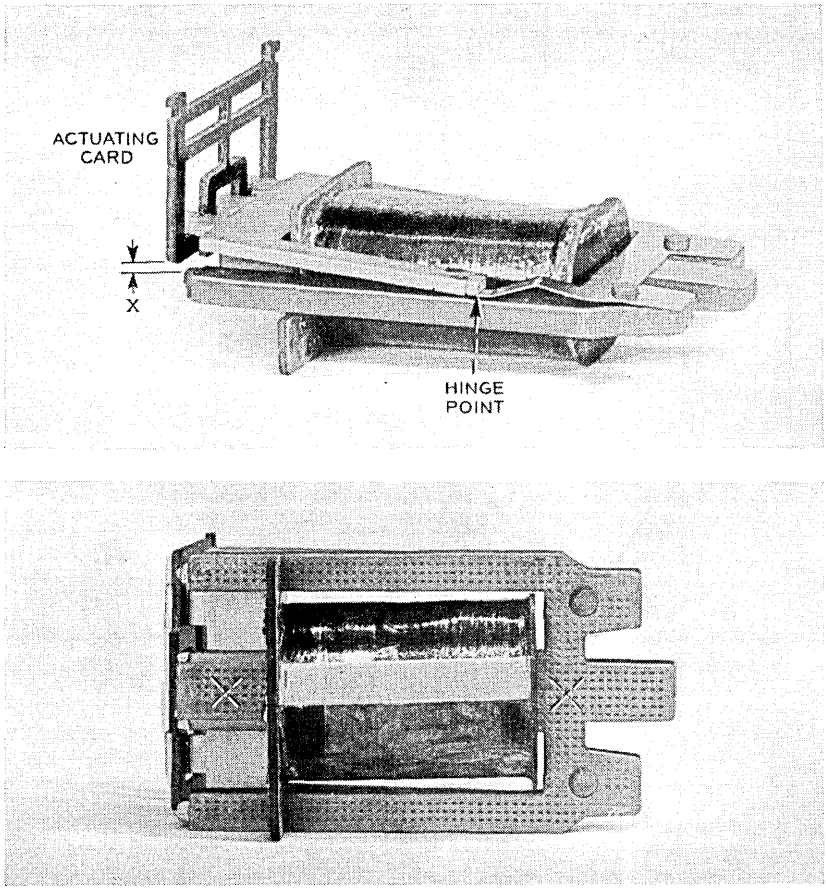


Fig. 2 — Magnetic parts, coil, and actuating card of AJ relay.

netization curves. The dashed curve of Fig. 1 is the decreasing magnetization curve for the same gap as the adjacent magnetization curve.

U is a function of x, and changes as the armature moves. If this motion occurs at a constant value of φ, there is no induced voltage and hence no electrical transfer of energy. A decrease in U must therefore represent mechanical work done on the armature. For a differential change dx in armature position, the mechanical work is Fdx, where F is the pull exerted on the armature. A general expression for F is therefore given by:

$$F = \frac{\partial U}{\partial x}_{(\varphi \text{ constant})} . \tag{2}$$

It follows from these considerations of energy balance that the magnetization relations determine the relations among the electrical input to the coil, the stored energy, and the mechanical output attainable. Hence, any generally applicable and simple expression for the magnetization relations, even if purely empirical, would provide a convenient means for evaluating and describing the performance characteristics.

The expressions given below for the magnetization relations are not purely empirical, but are those obtained as approximate solutions to the magnetic field equations by the magnetic circuit method. When experimentally evaluated, however, their utility in defining the characteristics of the electromagnet to which they apply is independent of this interpretation, and depends only on their conformity to the observed magnetization relations.

It is convenient to formulate the expressions for the magnetization curves in terms of the reluctance $\mathfrak{R}$, the ratio $\mathfrak{F}/\varphi$, where $\mathfrak{F}$ is the magnetomotive force $4\pi NI$. The observations of Fig. 1 are plotted in Fig. 3 in the form of curves giving $\mathfrak{R}$ vs. NI for various values of x. A constant value of φ is represented in such a plot by a straight line through the origin. The radial lines used as supplementary co-ordinates are spaced to give a convenient scale for φ.

The reluctance curves of Fig. 3 are similar in character to those applying to most ordinary electromagnets. Each curve has a relatively flat characteristic in the vicinity of a minimum located on a common flux

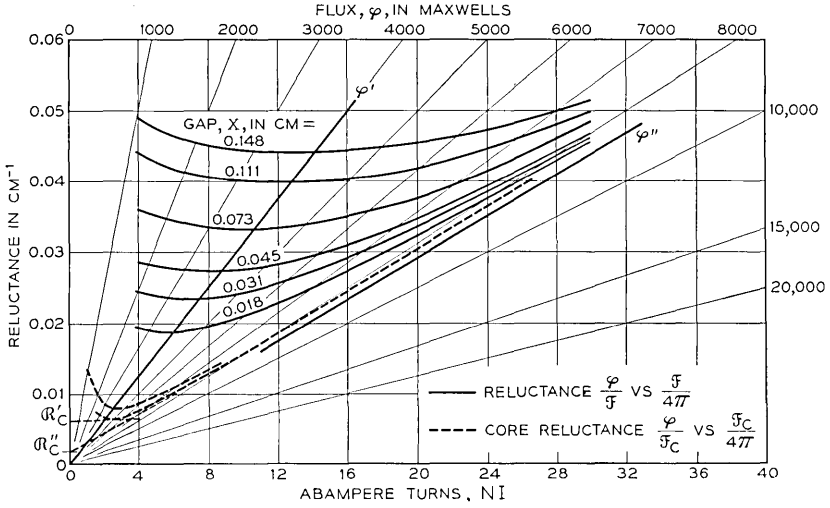


Fig. 3 — Reluctance curves.

line, that marked φ' in the figure. At values of φ above this minimum, the reluctance increases at an increasing rate, indicating an upper limit to the value of φ approached as NI becomes very large. The observed value of φ' may be interpreted as that for which the core density corresponds to maximum permeability, while the indicated upper limit may be interpreted as the saturation flux φ''.

The observations plotted in Figs. 1 and 3 were obtained with the sample initially demagnetized, a convenient reference condition for measurement. In actual use, relays and other electromagnets have been previously operated, and the applicable φ versus $\mathfrak{F}$ relation is that for repeated magnetization. In this case there is little or no increase in reluctance below φ'. For engineering purposes, the reluctance below φ' may be taken as constant at the value observed at φ' in measurements made from an initially demagnetized condition. This assumption is equivalent to taking the φ versus $\mathfrak{F}$ relation as linear up to the "knee" of the curve, the point of tangency with a line through the origin.

It follows that expressions for the magnetization relations may take $\mathfrak{R}$ as a function of x only, independent of φ, in the low density region, $\varphi < \varphi'$, while for $\varphi > \varphi'$, modified expressions must be employed which show $\mathfrak{R}$ increasing with φ, and approaching $\mathfrak{F}/\varphi''$.

Magnetic Circuit Schematics

The reluctance expressions for the magnetization relations can be conveniently defined by the magnetic circuit schematics shown in Figs. 4, 5, and 6. In these the reluctance $\mathfrak{R}$, or $\mathfrak{F}/\varphi$, is represented by a network of component reluctances, analogous to a network of electrical resistances, with flux analogous to current, and magnetic potential analogous to voltage. Thus the reluctance corresponding to Fig. 4 is given by:

$$\mathfrak{R} = \frac{\mathfrak{R}_L\left(\mathfrak{R}_0 + \frac{x}{A}\right)}{\mathfrak{R}_L + \mathfrak{R}_0 + \frac{x}{A}}. \tag{3}$$

The interpretation of the parameters $\mathfrak{R}_0$, $\mathfrak{R}_L$, and A in terms of the magnetic circuit concept is discussed in the companion article cited above.[1] In the analysis of experimental results they are parameters which summarize measurements in which the observed values of $\mathfrak{R}$ conform to (3). From this viewpoint Fig. 4 indicates only that the total flux φ is the sum of a flux φ_L for which the reluctance is constant, and a flux φ_G for which the reluctance varies directly with x.

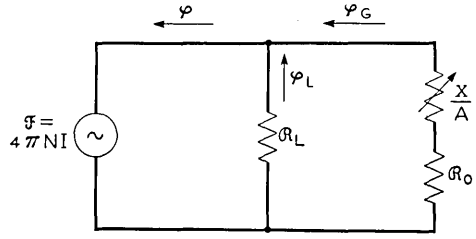


Fig. 4 — Equivalent magnetic circuit.

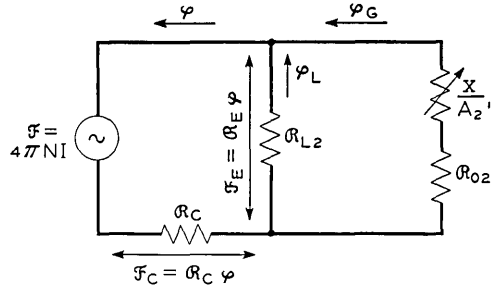


Fig. 5 — Magnetic circuit for core saturation.

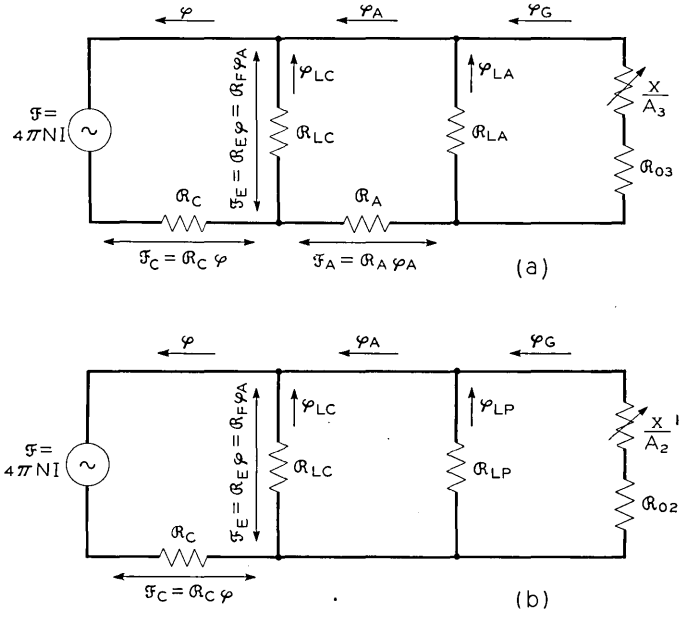


Fig. 6 — Magnetic circuit for armature saturation.

In Fig. 4, $\mathcal{R}_0$, $\mathcal{R}_L$, and A are constants, so that $\mathcal{R}$ is a function of x only, independent of φ. Equation (3) can then apply only to the low density region, $\varphi < \varphi'$. For most ordinary electromagnets, the magnetization relations for $\varphi < \varphi'$ conform to (3), and hence to the schematic of Fig. 4.

For $\varphi > \varphi'$, the expression for $\mathcal{R}$ must provide for the variation with φ. It usually suffices to use an expression in which the only term varying with φ is that corresponding to the path in which saturation first occurs. The most common case is that in which saturation first occurs in the core. The reluctance can then be taken as conforming to the schematic of Fig. 5, in which $\mathcal{R}_{02}$, $\mathcal{R}_{L2}$ and A_2 are constants, and $\mathcal{R}_C$ is a function of φ only. The total reluctance is given by:

$$\mathcal{R} = \mathcal{R}_C + \mathcal{R}_E, \tag{4}$$

where:

$$\mathcal{R}_E = \frac{\mathcal{R}_{L2}\left(\mathcal{R}_{02} + \dfrac{x}{A_2}\right)}{\mathcal{R}_{L2} + \mathcal{R}_{02} + \dfrac{x}{A_2}}. \tag{5}$$

If the variation with φ is taken as conforming to the empirical Frohlich-Kennelly equation,[6] $\mathcal{R}_C$ is given by:

$$\mathcal{R}_C = \mathcal{R}_C'' \frac{\varphi''}{\varphi'' - \varphi} = \mathcal{R}_C' \frac{\varphi'' - \varphi'}{\varphi'' - \varphi}, \tag{6}$$

where φ'' is the saturation flux and φ' is the flux for maximum permeability or minimum reluctance $\mathcal{R}_C'$. If these assumptions apply, the minimum values of $\mathcal{R}$ for all values of x must lie on φ', as for the results of Fig. 3. This common minimum is thus evidence that the core is controlling with respect to the variation with φ, and that (4) and (6) are applicable. As (6) applies only for $\varphi > \varphi'$, $\mathcal{R}_C'$ is the value of $\mathcal{R}_C$ not only at $\varphi = \varphi'$, but throughout the low density region $\varphi < \varphi'$. In the alternative form given by (6), $\mathcal{R}_C''$ is defined by this equation, and represents merely the intercept of a plot of $\mathcal{R}_C$ extended below the region to which (6) applies.

In some electromagnets saturation occurs in the armature rather than in the core. This is the case, for example, in high speed relays in which the armature section is minimized to reduce its mass. In such cases, the reluctance conforms to the schematic of Fig. 6(a), in which $\mathcal{R}_A$ represents the reluctance of the armature, which varies with φ_A. As the ratio φ_A/φ

then varies with x, the minimum values of $\mathcal{R}$ do not lie on a common value of φ in this case.

In Fig. 6(a), $\mathcal{R}_A$ may be taken as given by an expression of the same form as (6), with φ' and φ'' replaced by the minimum and saturation values of φ_A, and $\mathcal{R}_C'$ replaced by the minimum value of $\mathcal{R}_A$. $\mathcal{R}_C$ may be taken either as given by (6), or as a constant if the variation in $\mathcal{R}_A$ is dominant. The other parameters of Fig. 6(a) are constants. The expression for $\mathcal{R}$ applying to Fig. 6(a) may be obtained in the same manner as those given for the circuits of Figs. 4 and 5.

The magnetic circuits of Figs. 5 and 6 are called "design" circuits, because their constants may be estimated from the dimensions and material constants of the design, as discussed in the companion article.[1] Estimates of the corresponding values of the equivalent magnetic circuit constants of Fig. 4 are obtained from these design constant estimates by the equivalence equations given below.

Conditions of Equivalence

Whatever magnetic circuit is taken as applying, the component reluctances are independent of φ in the low density region, $\varphi < \varphi'$, and are therefore constants except for the gap reluctance. If this varies directly as x, as assumed in the schematics shown, the reluctance for any circuit reduces to an expression of the form of (3), applying to the circuit of Fig. 4. As shown in the companion article,[1] the reluctance of the circuit of Fig. 5 when $\mathcal{R}_C$ is constant is given by (3) when the parameters of this equation are given by:

$$\begin{aligned} \mathcal{R}_L &= \mathcal{R}_{L2} + \mathcal{R}_C , \\ A &= \frac{A_2}{p^2} , \\ \mathcal{R}_0 &= p^2\mathcal{R}_{02} + p\mathcal{R}_C , \end{aligned} \tag{7}$$

where:

$$p = 1 + \frac{\mathcal{R}_C}{\mathcal{R}_{L2}} .$$

These relations are the conditions of equivalence, for which the reluctances of Figs. 4 and 5 are identical for all values of x. With them, the reluctance given by (4) can be reduced to the simpler form of (3) when $\mathcal{R}_C$ is a constant, as throughout the low density region.

Similar relations apply to the magnetic circuit of Fig. 6(a). In particu-

lar, when $\mathcal{R}_A$ is a constant, the reluctance $\mathcal{R}_F$, or $\mathcal{F}_E/\varphi_A$, represented by the series parallel path of φ_A in Fig. 6(a), may be represented by the simple parallel path of φ_A in Fig. 6(b). By analogy with equations (7) the constants applying are related by:

$$\begin{aligned} \mathcal{R}_{LP} &= \mathcal{R}_{LA} + \mathcal{R}_A , \\ A_2 &= \frac{A_3}{p^2} , \\ \mathcal{R}_{02} &= p^2\mathcal{R}_{03} + p\mathcal{R}_A , \end{aligned} \tag{7A}$$

where

$$p = 1 + \frac{\mathcal{R}_A}{\mathcal{R}_{LA}} .$$

The circuit of Fig. 6(b) is identical with that of Fig. 5 with $1/\mathcal{R}_{L2}$ equal to the sum of $1/\mathcal{R}_{LC}$ and $1/\mathcal{R}_{LP}$. Thus the circuit of Fig. 6(a) may be reduced to that of Fig. 5 when $\mathcal{R}_A$ is constant. The resulting circuit may be reduced in turn to the equivalent circuit of Fig. 4 when $\mathcal{R}_C$ is constant. Thus the equivalent circuit may be used to represent all cases in the low density region, where the component reluctances are independent of φ, provided the gap reluctance varies linearly with x.

3 ANALYSIS OF MAGNETIZATION MEASUREMENTS

In analysing the observed magnetization relations, it is convenient to treat the low density and high density relations in separate and successive steps. For the former, the analysis is based on the equivalent magnetic circuit of Fig. 4.

EVALUATION OF EQUIVALENT MAGNETIC CIRCUIT CONSTANTS

For a given value of x, the total reluctance in the low density region is taken as constant at the minimum value $\mathcal{R}'(x)$ observed in measurements made from the demagnetized condition. If the observations are plotted as reluctance curves, as in Fig. 3, these values of $\mathcal{R}'(x)$ may be read directly. When the recording fluxmeter[3] is used, $\mathcal{R}'(x)$ can be obtained from the φ versus $\mathcal{F}$ curve as the slope $\mathcal{F}/\varphi$ of the line through the origin tangent to the curve. The reciprocals $P(x)$ of the values of $\mathcal{R}'(x)$ thus evaluated may be plotted against x. Values of $P(x)$ corresponding to the minimum reluctance values of Fig. 3 are plotted in this way in Fig. 7. The origin of x is taken as for a gap of 0.025 cm., corresponding to the operated position of the actuating card for maximum stop pin height.

Values of x measured from this origin are termed travel, as distinguished from gap values, which are measured from the position of iron to iron contact.

If the relation between $\mathfrak{R}'(x)$ and x conforms to (3), the values of $P(x)$ must conform to:

$$P(x) = \frac{1}{\mathfrak{R}_L} + \frac{A}{A\mathfrak{R}_0 + x}. \tag{8}$$

Then if x_c is a central reference value of x, and $P(x_c)$ the corresponding value of $P(x)$, the expression for $P(x_c) - P(x)$ given by (8) may be written in the form:

$$\frac{x - x_c}{P(x_c) - P(x)} = A\left(\mathfrak{R}_0 + \frac{x_c}{A}\right)^2 + \left(\mathfrak{R}_0 + \frac{x_c}{A}\right)(x - x_c). \tag{9}$$

Thus conformity with (3) requires the ratio given by (9) to vary linearly with x. This ratio is plotted against $x - x_c$ in Fig. 7, referred to the upper and right hand scales. The values of this ratio are sensitive to deviations in the values of x and $P(x)$ near x_c, and minor variations from linearity are to be expected here. Allowing for this, the results agree with (9), and the relation between $\mathfrak{R}'(x)$ and x therefore conforms to (3).

From (9), the slope and intercept of the linear plot may be interpreted as indicated in Fig. 7. Then A may be evaluated by dividing the intercept

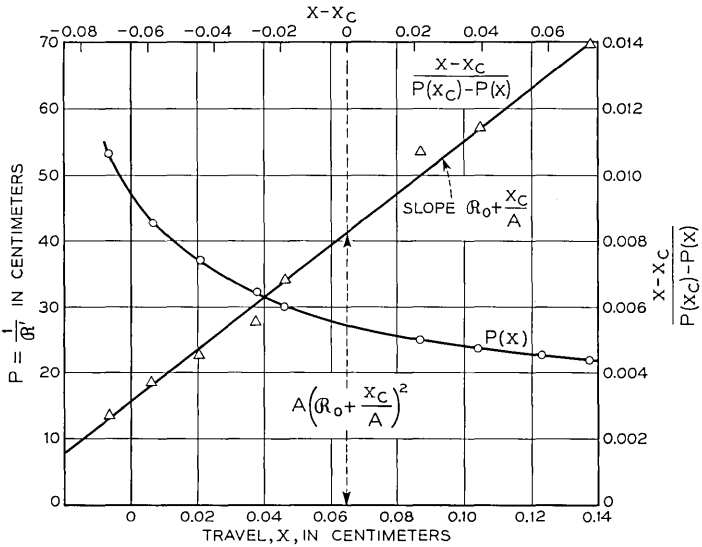


Fig. 7 — Evaluation of equivalent circuit constants.

value by the square of the slope value, and the latter then serves to evaluate $\mathcal{R}_0$ by subtracting x_C/A. On substituting these values of $\mathcal{R}_0$ and A in (8) at $x = x_c$, $\mathcal{R}_L$ may be evaluated. For the case plotted in Fig. 7, the values thus obtained are:

$$\mathcal{R}_0 : 0.0300 \text{ cm}^{-1}$$

$$A: 1.34 \text{ cm}^2$$

$$\mathcal{R}_L : 0.0690 \text{ cm}^{-1}$$

These values of the equivalent circuit constants, substituted in (3), suffice to accurately characterize the magnetization relations and thus the electromagnet's performance through the low density region.

Evaluation of High Density Relations

For cases of core saturation, the high density relations may be analysed in terms of the magnetic circuit of Fig. 5. The corresponding reluctance is given by (4), in which $\mathcal{R}_C = \mathcal{R}'_C$ for $\varphi = \varphi'$. If $\mathcal{R}'_C$ can be determined, the three parameters determining $\mathcal{R}_E$ in (5) may be evaluated by the equivalence equations (7). To determine $\mathcal{R}_C$ through the high density region from (6) requires evaluation of $\mathcal{R}'_C$, φ', and φ''. Thus the analysis of the high density relations requires evaluation of these three quantities in addition to the three parameters of the equivalent circuit evaluated above.

A method for evaluating $\mathcal{R}'_C$, φ', and φ'' directly from the magnetization curves is described subsequently, following a description of a preferred method which requires additional measurements. These measurements are determinations of the drop $\mathcal{R}_C\varphi$ in magnetic potential through the core, and are conveniently made with the magnetomotive force gauge developed by W. B. Ellwood.[4]

Magnetomotive Force Measurements

The Ellwood mmf gauge measures the difference in magnetic potential between the outer ends of the two reeds of a dry reed switch, as determined by the coil current required to open the switch. The end of one reed is used as a probe, which is brought into contact with the magnetic structure being measured. The end of the other reed, at a distance of about 5 cm from the probe, then lies on a relatively distant surface surrounding the structure: this surface is substantially at the same potential for any position of the gauge. Then the difference between two such measurements made at different points on the magnetic structure measures the difference in magnetic potential between them.

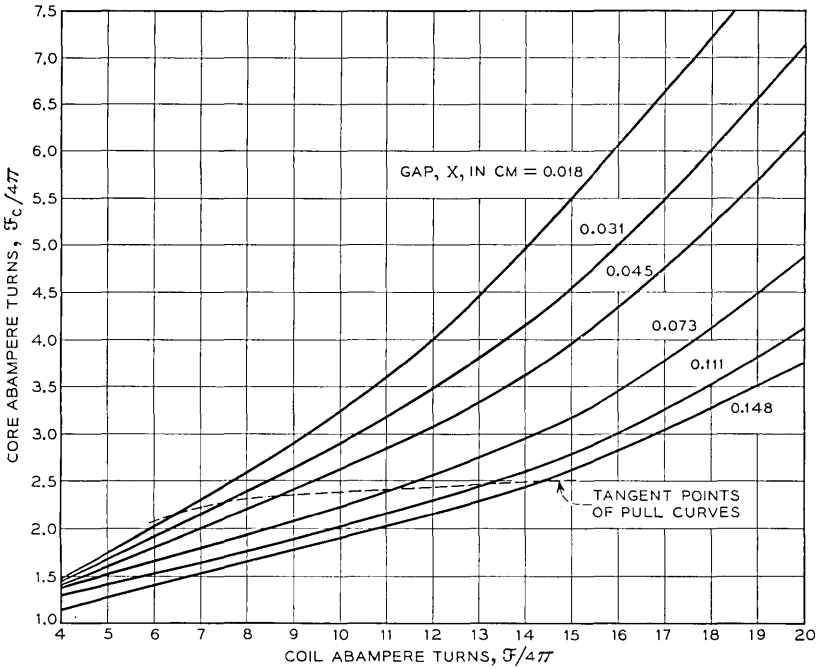


Fig. 8 — Relation of core potential drop to applied magnetomotive force.

To the extent that the physical structure may be identified with the corresponding magnetic circuit element, the drop in magnetic potential in the core may be identified with $\mathcal{R}_C\varphi$ in Fig. 5. Assuming this identification to apply, magnetic probe measurements may be made at the two ends of the core, as at the points marked X in Fig. 2. The potential drop observed in these measurements is the magnetic potential $\mathfrak{F}_E$ applied to the external magnetic circuit. $\mathfrak{F}_E$ is equal to the applied magnetomotive force $\mathfrak{F}$(or $4\pi NI$) less the potential drop $\mathfrak{F}_C = \mathcal{R}_C\varphi$ in the core. Thus $\mathfrak{F}_C = \mathfrak{F} - \mathfrak{F}_E$. Values of $\mathfrak{F}_C/(4\pi)$ thus determined for the relay of Fig. 2 are shown plotted against $\mathfrak{F}/(4\pi)$ for various values of x in Fig. 8. These curves are substantially linear in the low density region: their upward concavity at higher values of $\mathfrak{F}$ is evidence that saturation first occurs in the core.

Evaluation of Core Relucatance Constants

The magnetization curves of Fig. 1 and the core magnetomotive force curves of Fig. 8 were obtained with the same model. For given values of $\mathfrak{F}$ and x there can be read from these two figures corresponding values of

φ and $\mathfrak{F}_C$, giving the corresponding values of core reluctance $\mathfrak{R}_C = \mathfrak{F}_C/\varphi$. Such values of $\mathfrak{F}_C$ have been determined for two values of x (the smallest and largest included in these figures), and are shown plotted against $\mathfrak{F}_C/(4\pi)$ in the dashed curves included in Fig. 3. Values of $\mathfrak{R}_C$ for intermediate values of x are intermediate between these two curves, whose similarity supports the assumption of Fig. 5 and equation (4) that $\mathfrak{R}_C$ is substantially independent of x.

For $\varphi > \varphi'$ the plot of $\mathfrak{R}_C$ versus $\mathfrak{F}_C$ is substantially linear. In this it conforms to (6) which, by substitution of $\mathfrak{F}_C/\mathfrak{R}_C$ for φ, may be written in the form:

$$\mathfrak{R}_C = \mathfrak{R}_C'' + \frac{\mathfrak{F}_C}{\varphi''} = \left(1 - \frac{\varphi'}{\varphi''}\right)\mathfrak{R}_C' + \frac{\mathfrak{F}_C}{\varphi''}. \tag{10}$$

Thus φ'' may be evaluated from the slope of the plot of $\mathfrak{R}_C$ versus $\mathfrak{F}_C$, and $\mathfrak{R}_C''$ from the intercept of the dotted line extension, as indicated in Fig. 3. The observed relation deviates from (10) in the vicinity of φ', and the linear relation conforming to (10) does not intersect $\mathfrak{R}_C'$ at φ', but at some higher value of φ. Thus (6) more nearly represents the observed relation if φ' is taken as this higher value, rather than at the flux for minimum $\mathfrak{R}$ previously taken as φ' and marked at such in Fig. 3. The error in using this latter value of φ' in (6), however, is minor, and of little significance in engineering estimates.

Alternative Method of Determining Core Reluctance Constants

When magnetomotive force measurements are not made, the core reluctance constants can be determined directly from the magnetization curves as follows:

As can be seen in Fig. 3, the line representing φ'' can be directly identified as the apparent asymptote of the reluctance curves: specifically by determining the flux line parallel to the tangent to the upper portion of the lowest reluctance curve. As φ' is the value of φ at which the reluctance curves have their common minima, φ' and φ'' can be readily evaluated, and only $\mathfrak{R}_C'$ remains to be determined.

On substituting $\mathfrak{F}/\mathfrak{R}$ for φ in (6), and substituting this expression for $\mathfrak{R}_C$ in (4) the resulting equation can be solved to give:

$$\mathfrak{R} = \frac{1}{2}\left(\mathfrak{R}_E + \mathfrak{R}_C'' + \frac{\mathfrak{F}}{\varphi''} + \sqrt{\left(\mathfrak{R}_E + \mathfrak{R}_C'' - \frac{\mathfrak{F}}{\varphi''}\right)^2 + 4\mathfrak{R}_C''\frac{\mathfrak{F}}{\varphi''}}\right).$$

Let $\mathfrak{R}''$ be the particular value of $\mathfrak{R}$ for which

$$\mathfrak{F} = \mathfrak{R}'\varphi'' = (\mathfrak{R}_C' + \mathfrak{R}_E)\varphi''.$$

As both $\mathfrak{R}_c'/\mathfrak{R}'$ and $\mathfrak{R}_c''/\mathfrak{R}'$ are small, substitution of $\mathfrak{F} = \mathfrak{R}'\varphi''$ in the preceding expression gives as an approximation:

$$\mathfrak{R}'' = \mathfrak{R}' + \sqrt{\mathfrak{R}_c''\mathfrak{R}'},$$

from which:

$$\mathfrak{R}_c'' = \frac{(\mathfrak{R}'' - \mathfrak{R}')^2}{\mathfrak{R}'}. \tag{11}$$

Thus, as indicated in Fig. 9, the value of $\mathfrak{R}'$ may be read from any one of the reluctance curves, and the corresponding value of $\varphi''\mathfrak{R}'$ computed. The value of $\mathfrak{R}$ corresponding to $\mathfrak{F} = \mathfrak{R}'\varphi''$ is read from the curve, and taken as $\mathfrak{R}''$. The values of $\mathfrak{R}''$ and $\mathfrak{R}'$ are then substituted in (11) to determine $\mathfrak{R}_c''$. This procedure may be followed for all the reluctance curves, and a mean value of $\mathfrak{R}_c''$ computed. $\mathfrak{R}_c'$ is then taken as

$$\varphi''\mathfrak{R}_c''/(\varphi'' - \varphi').$$

Evaluation of External Reluctance Constants

The procedures described above provide for the determination (a) of the equivalent constants of Fig. 4, (b) of the minimum value $\mathfrak{R}_c'$ of $\mathfrak{R}_c$ in Fig. 5. The remaining constants of Fig. 5 can then be computed by means of the equivalence equations (7). Table I lists, for the measurements of Figs. 1 and 8, the values of the equivalent constants deter-

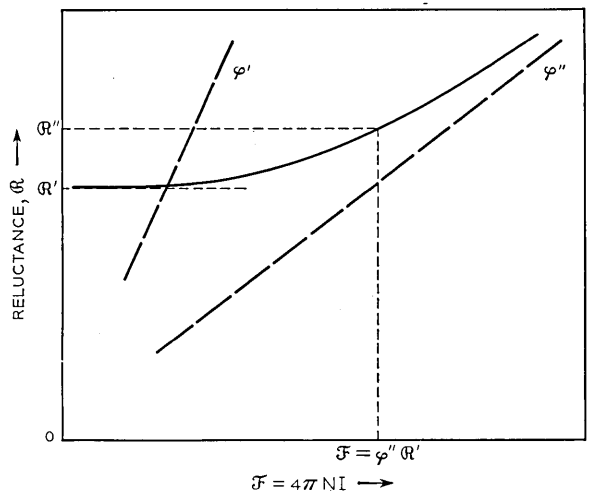


Fig. 9 — Alternative method of evaluating core reluctance.

TABLE I — EVALUATION OF MAGNETIC CIRCUIT CONSTANTS

Equivalent Values (Fig.4)	Design Values (Fig. 5)
From Fig. 7 $\mathcal{R}_0$: 0.0300 cm^{-1} A: 1.34 cm^2 $\mathcal{R}_L$: 0.0690 cm^{-1}	From Fig. 3 $\mathcal{R}_C'$: 0.0070 cm^{-1} From equations (7) $\mathcal{R}_{02}$: 0.0181 cm^{-1} A_2 : 1.66 cm^2 $\mathcal{R}_{L2}$: 0.0620 cm^{-1}

mined in Fig. 7, and the value of $\mathcal{R}_C'$ determined in Fig. 3. These values have been substituted in equations (7) to determine the component terms of $\mathcal{R}_E$.

The constants applying to Fig. 5 are of interest primarily for comparison with the values computed from the design, as discussed in the companion article.[1] While formally required to evaluate the high density magnetization relations, they are not explicitly required in estimating the high density pull, as is shown in a later section.

The reluctance $\mathcal{R}_E$ of equations (4) and (5) can be evaluated experimentally through the full range of the magnetization measurements when the magnetomotive force measurements are available. Thus Figs. 1 and 3 can be used to determine values of φ and $\mathcal{F}_E$ (or $\mathcal{F} - \mathcal{F}_C$) for corresponding values of $\mathcal{F}$ and x. In this way there have been obtained the curves of $\mathcal{R}_E$ plotted against $\mathcal{F}_E$ for various values of x in Fig. 10. These curves are included here merely to illustrate the approximate validity of (4), in which $\mathcal{R}_E$ is taken as independent of φ. The departures from this condition are minor except for the smaller gaps at high values

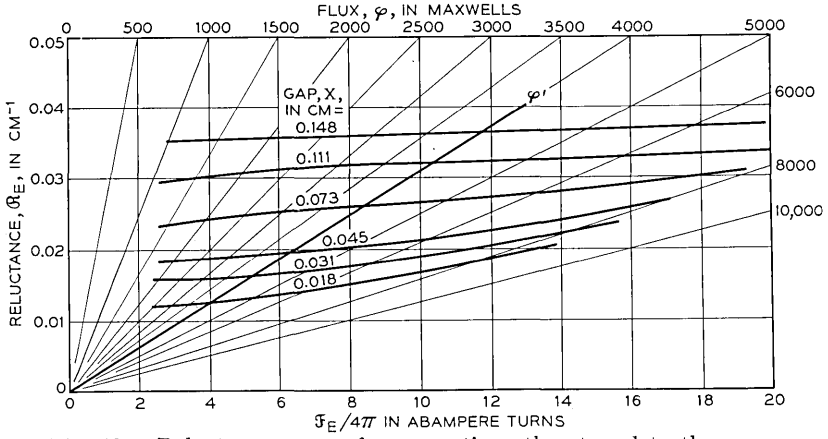


Fig. 10 — Reluctance curves for magnetic path external to the core.

of φ. Here the increase in flux density in the iron parts external to the core results in a significant increase in $\mathcal{R}_E$ as φ increases. By comparison with the plot of $\mathcal{R}_C$ included in Fig. 1, however, this increase in $\mathcal{R}_E$ is minor, and does not affect the fact that the limiting reluctance is $\mathcal{F}/\varphi''$, where φ'' is the saturation flux of the core.

It may be noted in passing that the values of $\mathcal{R}_E$ lying on φ' must conform to (5). As this is of the same form as (3), these values of $\mathcal{R}_E$ may be used to evaluate the component terms of (5) by the procedure applied to the values of $\mathcal{R}'$ in Fig. 7. As the same data are employed, the values of $\mathcal{R}_{02}$, $\mathcal{R}_{L2}$, and A_2 thus determined agree with those computed by means of equations (7) within the accuracy of the computations.

Case of Armature Saturation

The analysis of the low density relations is, of course, independent of where saturation occurs, and involves merely the determination of the equivalent magnetic circuit constants by the procedure described above. In the high density region, however, the quantities appearing in Fig. 6(a) must be evaluated when armature saturation occurs. This requires measurements not only of $\mathcal{F}$ and φ, but also of $\mathcal{F}_E$ and φ_A.

$\mathcal{F}_E$ is given by the Ellwood mmf gauge measurements previously described. To measure φ_A requires the use of a search coil wound on the armature, located over the region of maximum density. For twin return path structures, such as the relay of Fig. 2, twin search coils must be used, connected to measure the total armature flux. If the variation of $\mathcal{R}_A$ with φ_A is to be determined directly, additional mmf gauge measurements must be made of the potential drop $\mathcal{R}_A\varphi_A$ through the armature.

With φ_A determined as a function of $\mathcal{F}_E$ for various values of x, the reluctance $\mathcal{R}_F$, or $\mathcal{F}_E/\varphi_A$ may be analysed by the procedure described above for the analysis of the reluctance $\mathcal{R}$ or $\mathcal{F}/\varphi$. The curves of $\mathcal{R}_F$ versus $\mathcal{F}_E$ have minima $\mathcal{R}_F'$ at φ_A'. The reciprocals of these values of $\mathcal{R}_F'$ are plotted and analysed as in Fig. 7 to evaluate the equivalent constants $\mathcal{R}_{LP}$, $\mathcal{R}_{02}$, and A_2 of Fig. 6(b). The constants $\mathcal{R}_A'$, φ_A', and φ_A'' characterizing the relation between $\mathcal{R}_A$ and φ_A are determined either from mmf gauge measurements or by the method of Fig. 9. Then the constants $\mathcal{R}_{LA}$, $\mathcal{R}_{03}$, and A_3 of Fig. 6(a) can be evaluated by means of the equivalence equations (7A).

To complete the determination of the quantities appearing in Fig. 6(a), $\mathcal{R}_{LC}$ is evaluated as the ratio $\mathcal{F}_E/(\varphi - \varphi_A)$, while $\mathcal{R}_C$, which may either be constant or conform to (6), is evaluated from the relation between φ and $\mathcal{F}_C$.

4 PULL RELATIONS

As shown in Section 2, the pull F can be determined from the field energy U by means of equation (2). If the gap reluctance varies linearly with x, this equation reduces to:

$$F = \frac{\varphi_G^2}{8\pi A}, \tag{12}$$

as shown in the companion article.[1] Equation (12) is Maxwell's law for the pull between parallel plane surfaces of area A. As used here, $1/A$ has the more general sense of the coefficient of x in a linear expression for the gap reluctance. This equation is therefore applicable to all the cases previously discussed, with A, of course, taken as A_2 or A_3 for the relations applying to Figs. 5 and 6. As with the magnetization relations, it is convenient to give separate consideration to the pull at low densities and that at high densities.

Pull for Linear Magnetization

At low densities, where the magnetization relations are approximately linear, the reluctance conforms to the equivalent magnetic circuit of Fig. 4. In this case, φ_G is given by $\mathfrak{F}/(\mathfrak{R}_0 + x/A)$, and equation (12) becomes:

$$F = \frac{\mathfrak{F}^2}{8\pi A\left(\mathfrak{R}_0 + \frac{x}{A}\right)^2},$$

which may be written in the more familiar form:

$$F = \frac{2\pi(NI)^2}{A\left(\mathfrak{R}_0 + \frac{x}{A}\right)^2}. \tag{13}$$

For conformity with (13) the pull for a given value of x should vary as $(NI)^2$, giving a linear plot with a slope of two when plotted against NI on logarithmic paper. Fig. 11 shows pull measurements of the relay of Fig. 2 plotted in this way. The dashed lines tangent to the curves have a slope of two, and conform in this respect to (13). It is convenient to denote the pull indicated by these dashed lines as F'. Near and below the point of tengency the actual pull F differs little from F'. The pull, like the magnetization, is measured under the condition of initial demagnetization, which gives lower magnetization, and hence less pull, than normally applies in actual use when the magnet has been previously

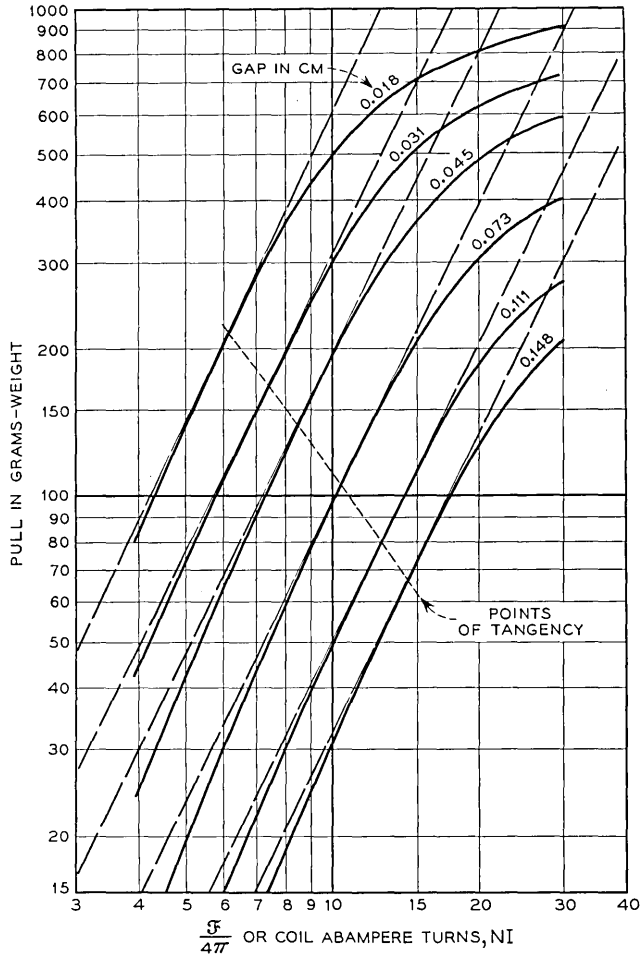


Fig. 11 — Pull curves.

operated. Hence F' satisfactorily represents the pull applying in practice through the low density region. In the high density region, the pull falls off as the result of incipient saturation, and the ratio F'/F increases as $\mathfrak{F}$ (or $4\pi NI$) is increased.

High Density Pull — Core Saturation

It is convenient to express the high density pull in terms of the ratio F'/F, where F is the actual pull, and F' the indicated pull conforming to (13) for the same value of $\mathfrak{F}$ (or $4\pi NI$). As F'/F is unity at low densi-

ties, it follows from (12) that F'/F is equal to $\bar{\varphi}_G^2/\varphi_G^2$, where φ_G is the actual value, and $\bar{\varphi}_G$ the fictitious value that would obtain if the magnetic circuit reluctances remained constant at all values of $\mathfrak{F}$. For the case of core saturation, to which the magnetic circuit of Fig. 5 applies, the ratio φ_G/φ is constant for a given value of x, and hence in this case F'/F equals $\bar{\varphi}^2/\varphi^2$ where φ is the actual value and $\bar{\varphi}$ the fictitious value that would obtain if the reluctance were $\mathfrak{R}'$. Thus for core saturation,

$$\frac{F'}{F} = \left(\frac{1}{\varphi}\frac{\mathfrak{F}}{\mathfrak{R}'}\right)^2,$$

and hence:

$$\frac{F'}{F} = \left(\frac{\mathfrak{R}}{\mathfrak{R}'}\right)^2. \tag{14}$$

As values of $\mathfrak{R}'$ and $\mathfrak{R}$ can be read directly from magnetization curves, this relation provides a simple means of computing the high density pull for specific values of x and $\mathfrak{F}$, provided the low density pull has been measured or estimated from (13), using the values of $\mathfrak{R}_0$ and A determined from the low density magnetization relations.

An approximate expression for the right-hand side of (14) can be obtained in a convenient form from the two limiting conditions: $\mathfrak{R} = \mathfrak{R}'$ for $\mathfrak{F} = \mathfrak{R}'\varphi'$, and $\mathfrak{R}$ approaches $\mathfrak{F}/\varphi''$ as $\mathfrak{F}$ becomes very large. A simple expression satisfying these conditions is:

$$\mathfrak{R}^2 = \mathfrak{R}'^2\left(1 - \left(\frac{\varphi'}{\varphi''}\right)^2\right) + \frac{\mathfrak{F}^2}{\varphi''^2}.$$

Aside from meeting the limiting conditions, this, like the Frohlich-Kennelly relation used previously, is a purely empirical expression. Assuming it to apply, the high density pull in the case of core saturation is, from (14), given by:

$$\frac{F'}{F} = 1 - \left(\frac{\varphi'}{\varphi''}\right)^2 + \left(\frac{\mathfrak{F}}{\mathfrak{R}'\varphi''}\right)^2. \tag{15}$$

As F' is given by (13), this expression gives the high density pull in terms of the magnetomotive force $\mathfrak{F}$, or explicitly:

$$F = \frac{\mathfrak{F}^2}{8\pi A\left(\mathfrak{R}_0 + \frac{x}{A}\right)^2\left(1 - \left(\frac{\varphi'}{\varphi''}\right)^2 + \left(\frac{\mathfrak{F}}{\mathfrak{R}'\varphi''}\right)^2\right)}, \tag{15A}$$

where $\mathfrak{R}'$ is given by (3), so that the constant terms in this expression are confined to φ', φ'', and the equivalent magnetic circuit constants A, $\mathfrak{R}_0$, and $\mathfrak{R}_L$.

High Density Pull — Armature Saturation

A similar treatment is applicable in the case of armature saturation, where F'/F is equal to φ_G^2/φ_G^2 as before. In this case, however, it is φ_G/φ_A rather than φ_G/φ that is constant, so that the expression corresponding to (14) is:

$$\frac{F'}{F} = \left(\frac{\mathcal{R}_F}{\mathcal{R}'_F}\right)^2, \tag{16}$$

where $\mathcal{R}_F$ is the reluctance $\mathcal{F}_E/\varphi_A$ of Fig. 6(a), and $\mathcal{R}'_F$ is its minimum value. The limiting expression for $\mathcal{R}_F$ is $\mathcal{F}_E/\varphi''_A$, and the approximation corresponding to (15) is therefore:

$$\frac{F'}{F} = 1 - \left(\frac{\varphi'_A}{\varphi''_A}\right)^2 + \left(\frac{\mathcal{F}_E}{\mathcal{R}'_F\varphi''_A}\right)^2. \tag{17}$$

In general, there is no simple expression for $\mathcal{F}_E/\mathcal{F}$, and (17) is of little interest as a general expression. When saturation is wholly confined to the armature, however, $\mathcal{R}_C$ is a constant and usually a minor term, and $\mathcal{F}_E$ is then nearly equal to $\mathcal{F}$. In this case, the high density pull is given by the approximation:

$$F = \frac{\mathcal{F}^2}{8\pi A\left(\mathcal{R}_0 + \frac{x}{A}\right)^2\left[1 - \left(\frac{\varphi'_A}{\varphi''_A}\right)^2 + \left(\frac{\mathcal{F}^2}{\mathcal{R}'_F\varphi''_A}\right)^2\right]}. \tag{17A}$$

The application of this approximation may be simplified by developing an approximate expression for $\mathcal{R}'_F$ in terms of the equivalent magnetic circuit constants. The expression for $\mathcal{R}'_F$ can be read from Fig. 6(b), where $\mathcal{R}_{LP}$ is usually large compared with $\mathcal{R}_{02} + x/A_2$, so that the latter term is an approximate expression for $\mathcal{R}'_F$. As (17A) applies only for $\mathcal{R}_C$ small, and as Fig. 6(b) is then of nearly the same form as the equivalent circuit of Fig. 4, the equivalent constants $\mathcal{R}_0$ and A are in this case approximately equal to $\mathcal{R}_{02}$ and A_2, respectively. Hence in the case of saturation confined wholly to the armature, to which (17A) applies, R'_F can be taken as given by $\mathcal{R}_0 + x/A$, and (17A) then involves only φ'_A, φ''_A, and the equivalent magnetic circuit constants $\mathcal{R}_0$ and A.

5 ANALYSIS OF PULL MEASUREMENTS

The following discussion relates primarily to the use of pull measurements in analytical studies as a means of evaluating the magnetic circuit constants.

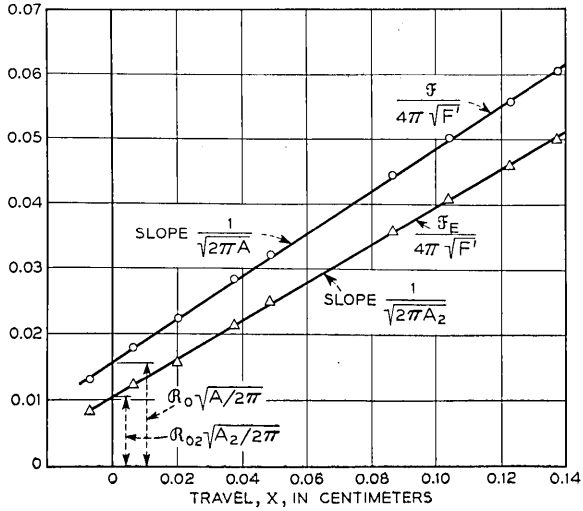


Fig. 12 — Evaluation of magnetic circuit constants from pull data.

Evaluation of Equivalent Magnetic Circuit Constants

The low density pull is given approximately by equation (13), to which the dashed F' lines of Fig. 11 conform. These lines coincide with the actual pull curves at the points of tengency, which must correspond to the minimum reluctance values. Taking F as F', (13) may be written in the form:

$$\frac{\mathfrak{F}}{4\pi\sqrt{F'}} = \mathfrak{R}_0 \sqrt{\frac{A}{2\pi}} + \frac{x}{\sqrt{2\pi A}}. \tag{18}$$

Values of NI/F' have been determined for the dashed lines of Fig. 11, and are plotted against x in Fig. 12 — the plot marked $\mathfrak{F}/(4\pi\sqrt{F'})$. In agreement with (18) these points determine a straight line. The slope and intercept have been used to determine the values of $\mathfrak{R}_0$ and A listed under "Pull Results" in Table II.

As the reluctance $\mathfrak{R}_E$ in the circuit of Fig. 5 is identical in form with the reluctance $\mathfrak{R}$ of Fig. 4, the pull for Fig. 5 is given by an expression similar to (13), with $\mathfrak{F}$, $\mathfrak{R}_0$, and A replaced by $\mathfrak{F}_E$, $\mathfrak{R}_{02}$, and A_2, respectively. As the two circuits are equivalent at the points of minimum reluctance and the corresponding points of tangency for the pull curves, the pull at these points is the same for the two cases. Thus if the ratio of $\mathfrak{F}_E$ to $\mathfrak{F}$ is determined at the points of tangency, a plot of $\mathfrak{F}_E/(4\pi\sqrt{F'})$ against x should conform to:

$$\frac{\mathfrak{F}_E}{4\pi\sqrt{F'}} = \mathfrak{R}_{02} \sqrt{\frac{A_2}{2}} + \frac{x}{\sqrt{2\pi A_2}}. \tag{19}$$

TABLE II — EVALUATION OF MAGNETIC CIRCUIT CONSTANTS

From Pull Results	From Magnetization Results
$\mathcal{R}_0$: 0.0319 cm^{-1}	0.0300 cm^{-1}
A: 1.46 cm^2	1.34 cm^2
$\mathcal{R}_{02}$: 0.0188 cm^{-1}	0.0181 cm^{-1}
A_2 : 1.85 cm^2	1.66 cm^2
$\mathcal{R}_L$: 0.0630 cm^{-1}	0.0690 cm^{-1}
$\mathcal{R}_{L2}$: 0.0560 cm^{-1}	0.0620 cm^{-1}
$\mathcal{R}_C'$: 0.0072 cm^{-1}	0.0070 cm^{-1}

The points of tangency are indicated on the pull curves of Fig. 11. These values of $\mathfrak{F}$ have been marked on the corresponding curves of Fig. 8, from which can be determined the corresponding values of $\mathfrak{F}_C$, and thus of $\mathfrak{F}_E$. It may be noted in passing that these values of $\mathfrak{F}_C$ are all similar, corresponding to a mean level of about 24 ampere turns, or 30 gilberts. The magnetization results gave values of 4,000 maxwells for φ' (Fig. 3) and of 0.007 cm^{-1} for $\mathcal{R}_C'$ (Table I), giving a value of 28 gilberts for the product. This agrees with the value of $\mathfrak{F}_C$ at the tangent points of the pull curves, showing that these points coincide with the points of minimum reluctance ($\varphi = \varphi'$).

Taking the ratio of $\mathfrak{F}_E$ to $\mathfrak{F}$ as read from the curves of Fig. 8 at the indicated tangent points, the values of $\mathfrak{F}/(4\pi\sqrt{F'})$ in Fig. 12 have been multiplied by the corresponding values of $\mathfrak{F}_E/\mathfrak{F}$, giving the points indicated by triangles in this figure. These conform to a straight line relation, in agreement with (19). The slope and intercept of this linear plot have been used to determine the values of $\mathcal{R}_{02}$ and A_2 listed under "Pull Results" in Table II.

The four quantities determined from the two plots of Fig. 12 appear in the three independent equivalence equations (7), and these three equations may therefore be used to determine the other three quantities: $\mathcal{R}_L$, $\mathcal{R}_{L2}$, and $\mathcal{R}_C'$. The resulting values of these latter quantities are listed under "Pull Results" in Table II. For comparison, the table inincludes the corresponding values of these quantities obtained from the magnetization measurements, and given previously in Table I. The agreement between the results derived from these different kinds of measurements indicates the validity of the analysis described here.

High Density Pull

As previously noted, it is convenient to express the pull F in the high density region in terms of the ratio F'/F, where F' is the pull computed from (13) for the value of $\mathfrak{F}$ applying. F'/F is readily evaluated from a logarithmic plot of the observed pull versus $\mathfrak{F}$, as illustrated by Fig. 11.

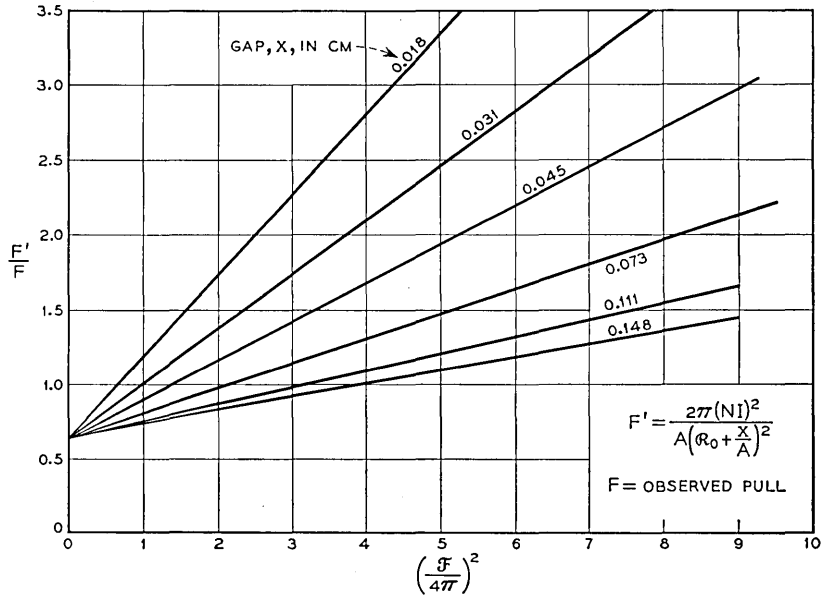


Fig. 13 — Pull characteristics for approaching saturation.

The dashed tangent lines give values of F', so corresponding values of F' and F can be read for various values of $\mathfrak{F}$. Values of F'/F for the results of Fig. 11 have been thus determined, and are plotted in Fig. 13 against $(\mathfrak{F}/4\pi)^2$ for the values of x applying to the individual curves of Fig. 11.

The values of F'/F thus plotted determine a family of straight lines having a common intercept at the origin of $\mathfrak{F}^2$. In this respect these results are representative of the pull characteristics of electromagnets when plotted in this way. As an empirical fact, apart from any analysis, this linearity provides a useful method of plotting pull observations in the high density region, as it facilitates interpolation and the detection of inconsistent observations.

The observed linearity in the relation between F'/F and $\mathfrak{F}^2$ conforms to the relations postulated in (15) and (17), of which the former should apply for core saturation. If (15) applies, the common intercept of 0.65 at $\mathfrak{F}^2 = 0$ in Fig. 13 should equal $1 - (\varphi'/\varphi'')^2$, indicating a value for φ'/φ'' of 0.59. To determine if the relation fully conforms to (15), the relation between the slopes of the lines and the corresponding values of x must be compared with that indicated by this equation.

Let y^2 be the observed slope of the plot of F'/F versus $\mathfrak{F}^2$ for a particu-

lar value of x. If the relation conforms to (15), $y = 1/(\mathfrak{R}'\varphi'')$, where $\mathfrak{R}'$ is given by (3). Writing C_L for the ratio $(\mathfrak{R}_L + \mathfrak{R}_0)/\mathfrak{R}_0$, and u for the ratio $x/(A\mathfrak{R}_0)$,

$$\mathfrak{R}' = \frac{1 + u}{C_L + u}\,\mathfrak{R}_L ,$$

and hence the values of y should conform to the equation:

$$y(1 + u) = \frac{C_L + u}{\mathfrak{R}_L\varphi''} . \tag{20}$$

Using the values of A and $\mathfrak{R}_0$ determined from the pull results in Fig. 12 and tabulated in Table II, values of u have been determined for the values of x applying in Fig. 13, and the corresponding values of y have been determined from the slopes of the corresponding lines. The values of $y(1 + u)$ thus determined are shown plotted against u in Fig. 14. The

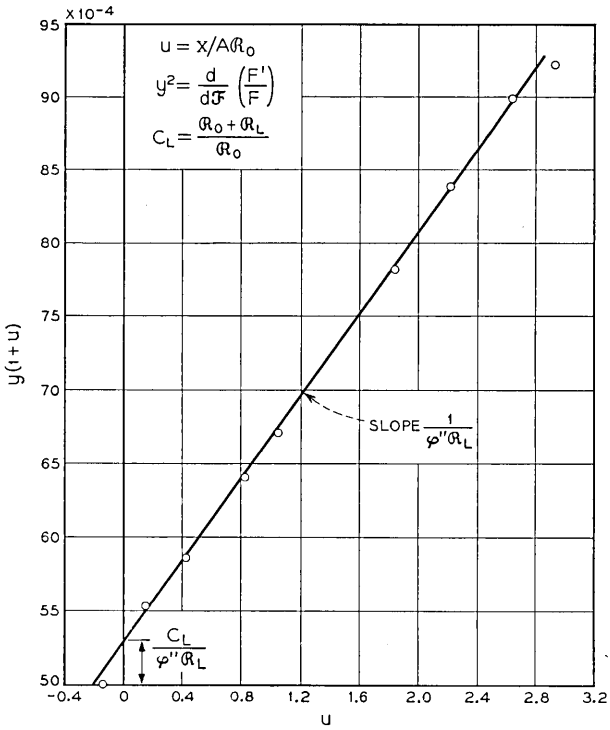


Fig. 14 — Evaluation of saturation flux φ'' and leakage reluctance $\mathfrak{R}_L$ from pull characteristics.

TABLE III — EVALUATION OF MAGNETIC CIRCUIT CONSTANTS

From Pull Measurements	From Magnetization Measurements
Figs. 13 & 14:	Fig. 3:
φ': 4,700 maxwells	φ': 4,000 maxwells
φ'': 8,000 maxwells	φ'': 8,500 maxwells
$\mathfrak{R}_L$: 0.091 cm^{-1}	
Fig. 12:	Fig. 7:
A: 1.46 cm^2	A: 1.34 cm^2
$\mathfrak{R}_0$: 0.0319 cm^{-1}	$\mathfrak{R}_0$: 0.0300 cm^{-1}
$\mathfrak{R}_L$: 0.063 cm^{-1}	$\mathfrak{R}_L$: 0.069 cm^{-1}

resulting plot is linear, in agreement with (20), showing the observed relation between F'/F and $\mathfrak{F}^2$ to conform fully to (15).

The expressions for the slope and intercept given by (20) are indicated in Fig. 14. From the values of the slope and intercept given by the plot, C_L may be evaluated, giving $\mathfrak{R}_L$, as $\mathfrak{R}_0$ is known. With $\mathfrak{R}_L$ thus evaluated, φ'' may be determined from the value of the slope. The values of $\mathfrak{R}_L$ and φ'' thus determined from results of Fig. 14 are listed in Table III, together with the value of φ' given by the ratio φ'/φ'' determined from the common intercept in Fig. 13. Included in this table under "Pull Results" are the values of A, $\mathfrak{R}_0$, and $\mathfrak{R}_L$ determined from the low density data and previously tabulated in Table II. Values of the same quantities as determined from the magnetization measurements are included for comparison.

For both pull and magnetization analysis, empirical relations have been assumed in (6) and (15) for the variation of $\mathfrak{R}$ with φ, agreeing with each other and the known conditions only at their upper and lower limits. In view of this, the agreement in the values of φ' and φ'' between the pull and magnetization results is good, and the deviation in the value of $\mathfrak{R}_L$ found in Fig. 14 from the other values of this quantity is not surprising.

High Density Pull for Armature Saturation

The low density pull for armature saturation may be analysed as in the case of core saturation. The pull data may be plotted against $\mathfrak{F}$ on logarithmic paper as in Fig. 11, and the tangents of slope two giving F' drawn, as in this figure. The values of $\mathfrak{F}/\sqrt{F'}$ may be plotted as in Fig. 12 to determine A and $\mathfrak{R}_0$, and A_2 and $\mathfrak{R}_{02}$ also evaluated if $\mathfrak{F}_E$ has been determined from magnetomotive force gauge measurements. As before, the remaining magnetic circuit constants of Figs. 4 and 5 may be de-

termined: the latter in this case are those corresponding to the equivalent schematic of Fig. 6(b).

For the high density pull, values of F'/F may be determined and plotted against $\mathfrak{F}$ as before. The resulting plot is similar to Fig. 13 in showing linear relations for each value of x, radiating from a common intercept. For the presentation of pull results, therefore, this method of plotting is applicable for both core and armature saturation, and for the case where saturation is approached concurrently in both.

When the slopes y^2 are determined, however, and $y(1 + u)$ plotted against u as in Fig. 14, a straight line is not usually obtained unless saturation is confined to the core. The observed relation for armature saturation is usually a curve which is concave upward, and approaches the horizontal at small values of u (travel x small). This result can be interpreted from the expression (17) for F'/F in the case of saturation confined to the armature. The corresponding expression for y is $1/(\mathcal{R}'_F\varphi''_A)$. It was shown above that in this case $\mathcal{R}'_F$ is approximately equal to $\mathcal{R}_0 + x/A$ or $\mathcal{R}_0(1 + u)$. The corresponding expression for $y(1 + u)$ is therefore the constant term $(1/(\mathcal{R}_0\varphi''_A)$. Thus for saturation confined to the armature, the plot of $y(1 + u)$ versus u is simply a horizontal line, from whose value φ''_A can be determined. In the more usual case of concurrent saturation, this relation is approached at small values of u, where φ_A/φ is relatively large, while at larger values of u, where φ_A/φ is smaller, the relation approaches that for core saturation.

6 DISCUSSION

The procedures for the analysis of pull and magnetization measurements described in this article are primarily intended for the evaluation of the magnetic circuit constants in development studies. They also have some related applications, and the choice of procedure varies with the application and with the measurements available. The following discussion of the applications of this analysis indicates the most convenient procedures for each case.

Presentation of Pull Measurements

Pull measurements are used not only for the guidance of development studies, but as engineering data in the application of relays and other electromagnets. A convenient form of presentation is that used in Fig. 11, where the pull is plotted against ampere turns on logarithmic paper. In preparing such data, interpolation of the measurements and the

recognition of inconsistent observations is facilitated by a supplementary plot of F'/F versus $\mathfrak{F}^2$ (or $(NI)^2$) as in Fig. 13.

Estimation of Pull

Estimates of pull are made in preliminary development either from estimates of the magnetic circuit constants, computed from the dimensions as discussed in the companion article,[1] or from magnetization measurements. The low density pull is given by (13), requiring only the equivalent circuit constants A and $\mathfrak{R}_0$ for its evaluation, and is represented by the F' lines of slope two in a logarithmic plot of F versus NI. Estimates of the high density pull may be prepared as linear plots of F'/F versus $(NI)^2$. As these conform (for core saturation) to (15) they require estimates of φ', φ'', and of $\mathfrak{R}'$, as given by (3).

Estimates of pull may also be required in preparing pull data for engineering use, with respect to the changes in pull associated with dimensional and other variations in a commercial product. Such estimates may be obtained by estimating the effect of the variations on the magnetic circuit constants fitting the observed pull, and computing the pull for the changed constants in the manner described in the preceding paragraph.

Evaluation of Equivalent Magnetic Circuit Constants

The equivalent magnetic circuit constants, which define the field energy and the pull characteristics through the low density range, largely suffice for the prediction of performance. This is illustrated by the use of these constants in the analysis of capability relations in one companion article[1] and in the dynamic relations presented in another.[5]

These constants are most readily and accurately evaluated from magnetization data by the procedure illustrated in Fig. 7 and the associated computations. A check on the values of A and $\mathfrak{R}_0$ is provided by analysing pull data by the procedure illustrated in Fig. 12. When only pull data are available, A and $\mathfrak{R}_0$ can be determined in this way, but the determination of $\mathfrak{R}_L$ requires the measurement of $\mathfrak{F}_E$ by the Ellwood mmf gauge. With $\mathfrak{F}_E$ known, the pull results suffice for the determination not only of $\mathfrak{R}_L$, but of the design circuit constants of Fig. 5 (provided saturation is confined to the core).

Evaluation of the Magnetic Design Circuit Constants

The magnetic design circuit constants are required only for comparison with estimated values, in development studies of the effect on the

performance of the configuration, dimensions and materials of the electromagnet. They are most readily and accurately determined from coil magnetization measurements. It is convenient to use Ellwood mmf gauge measurements to determine the core reluctance, but these may be omitted and the method of Fig. 9 employed, provided saturation is confined to the core. Subject to this same limitation, the design constants may be evaluated from the pull measurements if these are supplemented with mmf gauge measurements.

For armature saturation, the magnetic design circuit constants can only be evaluated from magnetization measurements, which must include armature search coil as well as coil flux determinations, and be supplemented by mmf gauge measurements.

Limits of Application

The validity and usefulness of the procedures described here rest on the agreement of measured quantities with the relations used to analyse them. As all the relations used lead to linear plots, the extent of agreement in any specific case is apparent from the plot obtained. So far as the presentation of pull results is concerned, this is the only question of validity involved.

The relations found for the magnetization results are used to estimate the field energy and the pull. To the extent the magnetization results relate to the flux linkages of the coil, the conclusions drawn from relations fitting those results are valid. This follows from the energy balance considerations discussed in Section 2. The supplementary measurements, such as those with the mmf gauge or an armature search coil, are in principle only convenient means for determining the relations to which the coil magnetization conforms. The relations found by these means are only valid to the extent of such agreement.

The conformity of magnetization relations to the expressions used here, corresponding to the magnetic circuit schematics, is closest for electromagnets in which the reluctance of the iron parts is small compared with that of the joints and air gaps. This condition is satisfied for most ordinary relays and similar electromagnets in the low density range, where the expressions given here most closely apply. The expressions used for the high density range do not give as close agreement, but provide a satisfactory basis for engineering analysis, particularly when saturation is confined to the core. The treatment is less satisfactory for structures that deviate from these conditions, such as those with armature saturation, or those with long cores of small cross section, where the

iron reluctance is a major component through the full range of operation.

ACKNOWLEDGMENT

The procedures described in this article incorporate contributions proposed or developed by H. N. Wagar, M. A. Logan, Mrs. K. R. Randall, and others.

REFERENCES

1. R. L. Peek, Jr., and H. N. Wagar, Magnetic Design of Relays, page 23 of this issue.
2. A. C. Keller, New General Purpose Relay, B.S.T.J., **31**, p. 1023, Nov., 1952.
3. H. N. Wagar, Relay Measuring Equipment, page 3 of this issue.
4. W. B. Ellwood, A New Magnetomotive Force Gauge, Rev. Sci. Instr., **17**, p. 109, 1946.
5. Estimation and Control of the Operate Time of Relays: Part I — Theory, R. L. Peek, Jr., page 109 of this issue. Part II — Applications, M. A. Logan, page 144 of this issue.
6. Kennelly, Trans. A.I.E.E., **8**, p. 485, 1891.

Estimation and Control of the Operate Time of Relays

Part I — Theory

By R. L. PEEK, Jr.

(Manuscript received August 31, 1953)

The dynamic equations applying to the operation and release of electromagnets are derived in a form in which the armature position and the flux linkages of the coil are the variables to be determined as functions of time. The effective magnetomotive force and pull are taken as given by the static magnetization relations. This formulation is valid if the dynamic field pattern is substantially that of the static field, and it is shown that this condition is satisfied if the effective conductance of the eddy current paths is small compared with that of the coil or other linking circuits. This is usually the case in operation, but is not the case in normal release.

Approximate solutions are obtained for both fast and slow operation, giving the time as related to the mechanical work done, the mass and travel of the moving parts, and the steady state power input to the coil. Expressions are derived for optimum coil design and pole face area, and design requirements for fast operation are discussed.

An approximate expression is derived for the release waiting time, the initial period of field decay prior to armature motion. The effect on this decay of contact protection in the coil circuit is discussed. The slow release case is treated in another article in this issue. An analysis is given of the armature motion in release, and the results of an analog computer solution to this problem reported. These show the effect of design parameters, particularly the armature mass, on the velocity attained in release motion, and hence on the amplitude of armature rebound.

INTRODUCTION

The operate and release times of most telephone relays lie in the range from 1 to 100 millisecs. (0.001 to 0.100 secs.). To use these relays in telephone switching, involving complex patterns of sequential switch

and relay operation, it is necessary: (1) to be able to estimate the operate and release times of any relay for which this information is needed in circuit studies, and (2) to develop relays capable of meeting specific timing requirements, both fast and slow.

The essential basis for such design and estimation is a dynamic theory of the operation of electromagnets, of sufficient accuracy for engineering use. The theory presented here is applicable to electromagnets in general, although the applications discussed are those relating to relays. The theory presented is approximate, partly because of the difficulty of providing a more exact treatment, and partly because simplicity and generality are more important for engineering purposes than accuracy in estimation, which is in any case subject to correction by measured results. The theoretical relations are used alone only in preliminary estimation and in the initial stages of development, as discussed in Part I of this article. In advanced development and in the modification and application of existing structures, as discussed in Part II, the theory is used as a guide in the correlation and extrapolation of observed performance.

The dynamics of electromagnets involve the concurrent and interrelated phenomena of field development and armature motion, and are accordingly governed by the two differential force equations respectively applying, each containing a coupling term expressing their reaction on each other. These basic equations are formally identical with those of electromechanical transducers, such as loud-speakers, but the treatment has little else in common. The operation of an electromagnet is the transient change from one state of equilibrium to another, as distinguished from the sustained low amplitude oscillations of the transducer case.

The coupling terms represent the effect of field energy changes on the coil voltage and on the force causing armature motion respectively. In the steady state, the field energy is a function of magnetomotive force and of armature position alone. In a transient state, eddy currents in the magnetic members affect both field energy and the pattern of the field. In developing an approximate theory, it is assumed that these effects are confined to the total effective mmf, and that the pattern of the field is the same as that of the static field: i.e., that the field energy associated with any portion of the structure, such as an air gap, is fixed by the flux linkages of the coil and by the armature position. The limitations on the analysis imposed by this simplifying assumption are discussed at the end of the next section.

1 The Dynamic Equations

The notation of this article conforms to the list given on page 257.

THE ELECTRICAL EQUATION

The voltage equation for a coil of N turns linking a field of strength φ may be written in the form:

$$(i - I)R + N\frac{d\varphi}{dt} = 0, \tag{1}$$

where i is the instantaneous current, R is the circuit resistance, and IR is the constant voltage applied. Writing $\mathfrak{F}_i$ for the instantaneous mmf $4\pi Ni$, and $\mathfrak{F}_{Si}$ for the steady state mmf $4\pi NI$, the equation becomes:

$$\mathfrak{F}_i - \mathfrak{F}_{Si} + 4\pi G_i\frac{d\varphi}{dt} = 0,$$

where $G_i = N^2/R$, the coil constant, or equivalent single turn conductance. If the field φ is linked by a number of circuits, a similar voltage equation applies to each such circuit. By addition of these expressions, there is obtained:

$$\mathfrak{F} - \mathfrak{F}_S + 4\pi G\frac{d\varphi}{dt} = 0,$$

where $\mathfrak{F} = \sum\mathfrak{F}_i$, the total effective mmf, $\mathfrak{F}_S = \sum\mathfrak{F}_{Si}$, the total applied mmf, and $G = \sum G_i$, the total equivalent single turn conductance. If the dynamic field has the same pattern as the static field, the instantaneous mmf $\mathfrak{F}$ must equal $\mathfrak{R}\varphi$, where $\mathfrak{R}$ is the reluctance $\mathfrak{F}/\varphi$ observed in static magnetization measurements. The preceding equation may therefore be written in the form:

$$\mathfrak{R}\varphi - \mathfrak{F}_S + 4\pi G\frac{d\varphi}{dt} = 0. \tag{2}$$

This relation controls the time rate of development of the flux φ. As the coil is the only circuit linking the field that has an externally applied voltage, $\mathfrak{F}_S$ is simply the steady state applied mmf $4\pi NI$. The conductance G includes not only the coil conductance G_C, or N^2/R, but terms for all conductive paths linking the field, including the eddy current paths. The effective conductance of the eddy current paths is denoted G_E. When a short circuited winding or sleeve is used, as in slow release relays, its conductance is denoted G_S. Thus in most cases of interest $G = G_C + G_E$; when a sleeve is used, $G = G_C + G_E + G_S$.

The static magnetization relations between $\mathfrak{F}$ and φ vary with the position of the armature, which is defined by its displacement x from the operated position. Hence the reluctance $\mathfrak{R}$ is, in general, a function of

both x and φ. Thus equation (2) is a differential equation in which φ, x, and $d\varphi/dt$ are the variables. With the armature at rest, x is constant, and (2) may be solved to give the initial field development in operate, or decay in release. When the armature is moving, the variation of φ and x is governed jointly by (2) and the mechanical force equation.

THE MECHANICAL EQUATION

The forces acting to accelerate the armature and associated moving parts, of effective mass m, are the magnetic pull F and the forces exerted by the contact and other springs. The total spring force may be written as dV/dx, where V is the potential energy stored in the springs, expressed as a function of x. The force equation may be written in the form:

$$F + m\frac{d^2x}{dt^2} + \frac{dV}{dx} = 0. \tag{3}$$

As x measures the gap opening, the velocity dx/dt is positive as the gap opens, and negative as it closes. The pull is directed toward the closed position and therefore tends to algebraically decrease the velocity. The spring force tends to open the gap and to algebraically increase the velocity, as V decreases with x, and dV/dx is negative.

On the assumption that the dynamic and static field patterns are the same, the pull F can be evaluated from the static magnetization relations. F is a function of φ and x, the variables of the flux development equation (2). Thus the dynamic performance is governed jointly by (2) and (3), to which the concurrent variation in φ and x with time must conform. To complete the formulation there are required explicit expressions for $\mathfrak{R}$ and F in terms of φ and x.

RELUCTANCE AND PULL

As shown in a companion article,[1] the magnetization relations for most electromagnets conform to the simple magnetic circuit of Fig. 1. Through the region of linear magnetization, in which the flux density is below that producing incipient saturation, the reluctances $\mathfrak{R}_0$ and $\mathfrak{R}_L$ and the area A are constants. If the magnetization relations conform to this schematic, the total reluctance $\mathfrak{R}$ is given by:

$$\mathfrak{R} = \frac{\mathfrak{R}_L\left(\mathfrak{R}_0 + \dfrac{x}{A}\right)}{\mathfrak{R}_L + \mathfrak{R}_0 + \dfrac{x}{A}}. \tag{4}$$

The constants $\mathfrak{R}_0$, $\mathfrak{R}_L$, and A of Fig. 1 and equation (4) are called

the equivalent values of the closed gap reluctance, the leakage reluctance, and the pole face area respectively. Procedures for estimating these quantities from the magnet's dimensions and material constants are described in the article cited above,[1] while methods for their experimental evaluation are described in another article in this issue.[2] When equation (4) applies, the pull F is given by:

$$F = \frac{(\mathcal{R}_L \varphi)^2}{8\pi A \left(\mathcal{R}_L + \mathcal{R}_0 + \frac{x}{A}\right)^2}. \tag{5}$$

In normal operation, the flux level lies in the region of linear magnetization, and equations (4) and (5) apply through the greater part of the period of flux development determining the operate time. In what fol-

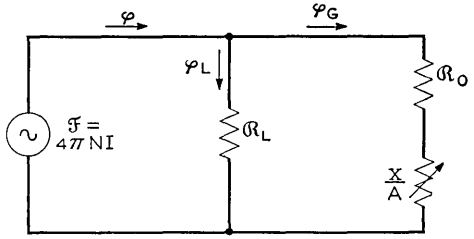


Fig. 1 — Equivalent magnetic circuit of an electromagnet.

lows, some simplification of notation is obtained by writing x_0 for $A\mathcal{R}_0$, and C_L for $(\mathcal{R}_0 + \mathcal{R}_L)/\mathcal{R}_0$, so that these equations become:

$$\mathcal{R} = \mathcal{R}_0 \frac{(C_L - 1)(x_0 + x)}{C_L x_0 + x}, \tag{4A}$$

$$F = \left(\frac{(C_L - 1)x_0}{C_L x_0 + x}\right)^2 \frac{\varphi^2}{8\pi A}. \tag{5A}$$

Thus the dynamic relations are given by (2) and (3), in which $\mathcal{R}$ and F are given by (4) and (5) respectively. The magnetic circuit constants can be evaluated by the methods cited above, and the other constant terms are known or given quantities, with the exception of the eddy current conductance G_E. Evaluation of this term requires determination of the conditions under which such a constant term can adequately represent the effects of eddy currents.

EDDY CURRENT CONDUCTANCE

Fig. 2 shows a simple electromagnet with a cylindrical core of length ℓ and diameter D. If $\mathcal{R}$ is the reluctance, the magnetomotive force $\mathcal{F}_c$ of

the winding current results in a flux $\varphi_C = \mathfrak{F}_C/\mathfrak{R}$, uniformly distributed across the core cross section. The leakage paths cause some longitudinal variation in φ, which can be neglected for the present purpose. If the flux is varying, eddy currents flow in circular paths with a current density j, which varies with the radius r. These give an increment to the total mmf varying from zero at the surface to a maximum at the center, producing a corresponding variation in the density of the total flux. If $\mathfrak{F}_C$ is large compared with the maximum magnetomotive force produced by the eddy currents, however, the density is nearly constant, and its rate of change is nearly constant throughout the core cross section. To the extent that this condition is satisfied, the effect of the eddy currents can be determined as follows.

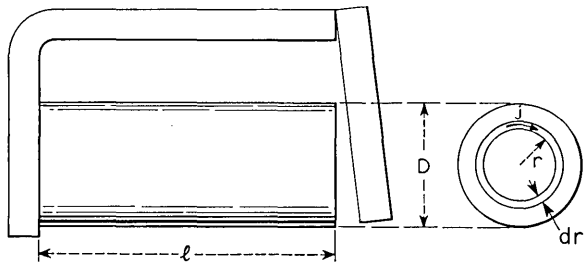


Fig. 2 — Eddy current paths in the core of an electromagnet.

The current in a cylindrical shell of radius r and differential thickness dr is $j\ell\, dr$. This links a part of the field proportional to the area enclosed, or $4r^2\varphi/D^2$, which varies, by hypothesis, at the same proportional rate as the total field. The resistance of the shell is $2\pi r\rho/(\ell\, dr.)$, where ρ is the resistivity of the material. The voltage equation for the shell circuit is therefore:

$$j\ell\, dr\, \frac{2\pi r\rho}{\ell\, dr} + \frac{4r^2}{D^2}\frac{d\varphi}{dt} = 0,$$

and hence:

$$j = -\frac{2r}{\rho\pi D^2}\frac{d\varphi}{dt}.$$

The magnetomotive force of the shell is its current multiplied by 4π. This produces a flux increment $d\varphi$ in the area enclosed inversely proportional to the reluctance of the tubes of induction within this area. This reluctance may be taken as inversely proportional to the area, as would be the case in a closed magnetic circuit of uniform cross section. Then

the flux increment produced by each shell is given by:

$$d\varphi = 4\pi j \ell \, dr \, \frac{4r^2}{D^2 \mathcal{R}}.$$

On substituting the preceding expression for the current density j, and evaluating the integral $\int_0^{D/2} d\varphi$, the total flux φ_E produced by the eddy currents is found to be:

$$\varphi_E = -\frac{\ell}{2\rho \mathcal{R}} \frac{d\varphi}{dt}. \tag{6}$$

This is identical in form with the expression for one of the linking circuits assumed in deriving equation (2), with $\mathcal{R}\varphi_E$ equal to $\mathcal{F}_E$, and $4\pi G_E$ given by $\ell/(2\rho)$. Thus to the extent that the assumption of a uniform flux distribution is valid, G_E is given by:

$$G_E = \frac{\ell}{8\pi\rho}. \tag{7}$$

This expression applies to a round core. A parallel approximation may be obtained for a rectangular rod by considering it as made up of rectangular shells of differential thickness, having their sides in the same ratio as the rod. By treating the perimeter of each shell as corresponding to the circumference of the circular shell in the cylinder case, with the area enclosed corresponding to that enclosed by the circular shell, the following expression is obtained for the conductance of a rectangular rod whose sides are in the ratio k:

$$G_E = \frac{\ell}{8\pi\rho} \frac{\sqrt{\pi k}}{1+k}. \tag{8}$$

This shows that the effect of using a rectangular section is equivalent to increasing the resistivity of the material. A similar effect is obtained by subdividing the core section, as in laminated construction. If the core, for example, were made of two similar round rods, equation (6) would apply to each, and the effective mmf acting on both would be the value of $\mathcal{R}\varphi_E$ given by this expression, with φ in $d\varphi/dt$ equal to half the total flux, so that the resulting expression for G_E would be half that given by (6). This argument can be generalized to show that, aside from the effect of changes in section shape, the effect of subdivision is to reduce G_E in proportion to the number of subdivisions.

The approximate validity of these expressions for G_E rests on the assumption that the eddy current magnetomotive force is minor com-

pared with that of the coil or other external circuit. From the derivation of (2), the component mmf's are proportional to the corresponding terms in G, so the preceding expressions for G_E are valid only if G_E is small compared with $G_C(N^2/R)$. In all but exceptional cases, this condition is satisfied in operation, and in release with a short circuited winding or a slow release sleeve. In these cases, the effect of eddy currents is adequately represented by a constant G_E term in (2). The exact values of G_E applying are in only approximate agreement with those given by (7) and (8), as the derivation of these expressions ignores the variation in flux density along the length of the core, and the eddy current effects in the armature and return path.

EDDY CURRENTS IN RELEASE

In normal release, the winding circuit is open, and the only effective magnetomotive force is that of the eddy currents. Evidently, the field in the outer layers must collapse almost instantly, while that in the center of the core is sustained by eddy currents in paths whose mean conductance, per line linked, is higher than that applying to a uniform field.

The relations applying to a closed path of uniform section can be formulated in differential form, and the solution for a cylindrical section has been given by Wwedensky.[3] For decay from an initially uniform field, the expression for the flux as a function of time is a series of exponential terms with progressively smaller time constants. The first term represents the most persistent part of the flux, and represents a field varying from zero at the surface to a maximum at the center, comprising 69 per cent of the initial field. Its time constant corresponds to an effective value of G_E 35 per cent larger than that given by (7). Somewhat similar relations must apply to an electromagnet, causing a time variation in the pattern of the field, not only radially, as in a closed uniform path, but in the longitudinal variation and in the division of the field between the leakage and armature paths.

An experimental study of flux development and decay in relays has been reported by M. A. Logan.[4] His results agree with this discussion in showing G_E in (2) to be effectively a constant, provided G_E/G_C is less than 0.2, as in the operation of most relays. An empirical expression is given for an effective value of $G_C + G_E$ which provides a correction factor applicable for small values of G_E/G_C. The results for normal release show the field decay to have the general character of Wwedensky's solution, and an empirical expression is given which agrees with observed results. This is primarily of interest in connection with the voltage

induced in the coil and imposed on the contact opening the coil circuit. Because of the changing field pattern, the gap field which determines the pull has a different, though similar, decay rate.

These considerations, as supported by Wwedensky's relations and Logan's results, indicate that the rate of pull decay in normal release is faster, and only roughly of the same order of magnitude, as that estimated on the assumption that (2) applies, with values of G_E similar to those applying in operation.

LIMITATIONS OF THE ANALYSIS

The validity of equations (2) through (5) rests on the assumption that the pattern of the dynamic field is essentially that of the static field to which the magnetization relations apply. The above discussion of the eddy currents indicates that this assumption is valid when G_E is a minor term in G. This condition is approximately satisfied in normal operation. In open circuit release, however, the condition is not satisfied, and equation (2) is only a crude first approximation to the controlling relation.

The use of (4) as an expression for the reluctance $\mathfrak{R}$ rests on the further assumption that the magnetization is linear, a condition only satisfied in the low density region. The initial and controlling stages of operation are usually complete before the field passes out of the low density region, and (4) is therefore applicable to the operate case. A different expression for $\mathfrak{R}$ is required in the release case, as discussed in Section 7.

2 Character of the Operate Solution

GRAPHICAL REPRESENTATION

Some understanding of the relations applying to operation may be obtained from their graphical representation in Figs. 3 and 4. In these two figures the path followed by the variables in dyamic operation is indicated by the dotted lines, with the dots spaced to indicate equal time intervals between them. Fig. 3 shows the φ versus $\mathfrak{F}$ relation, referred to the steady state magnetization curves for various values of x, with $x = 0$ corresponding to the operated position, and $x = x_1$ to the initial unoperated position. Fig. 4 shows the dynamic F versus x relation, together with the load curve (bounding the cross hatched area V) and the steady state pull curve for the applied mmf $\mathfrak{F}_S$.

The flux and pull increase together with the armature at rest at x_1 until the pull equals the back tension at the point 1. In the earlier motion, 1–2, the velocity is small, and the reluctance (from (4)) changes slowly with x so the motion has little effect on the rate of flux develop-

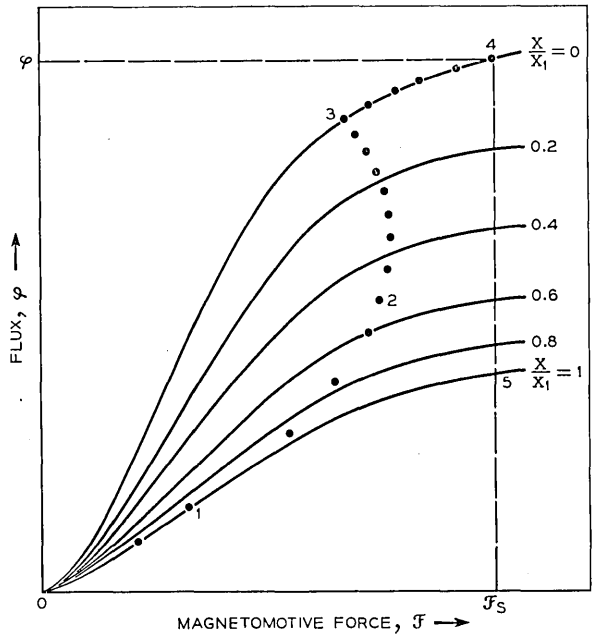


Fig. 3 — Field energy relations in the operation of an electromagnet.

ment. In the later motion, 2–3, the reluctance changes more rapidly with x, and the velocity is high, increasing $d\varphi/dt$ so as to result in a temporary decrease in $\mathfrak{F}$. Operation is complete at 3, with φ and F still below their steady state values at 4, which they then approach exponentially with $x = 0$.

The mechanical work done in operation is represented in Fig. 4 by the area under the dynamic pull curve (dotted line), and in Fig. 3 by the area bounded by 0–1–2–3–0. This, of course, is less than the work that would be done if $\mathfrak{F}$ were equal to its steady state value $\mathfrak{F}_S$ throughout the motion, represented by the area under the $\mathfrak{F}_S$ curve in Fig. 4, and by the loop 0–3–4–5–0 in Fig. 3. In Fig. 4 it can be seen that the work done exceeds the static load V by an amount represented by the shaded area T, corresponding to the kinetic energy of the armature and the parts that move with it. At the end of the stroke this energy is partly dissipated in impact, and partly transferred to vibratory motion of the relay and its parts.

The initial flux development, 0–1 is governed wholly by equation (2). This same relation dominates in controlling the performance in the early travel 1–2, in which the velocity is small, and the reluctance changes slowly. In the later travel, where the velocity is high and the reluctance

changes rapidly, equation (3) is controlling, with F corresponding to a more nearly constant value of φ.

METHODS OF SOLUTION

Formally, the expression for φ given by combining (3) and (5) may be substituted in (2), eliminating φ and giving a third order non-linear differential equation in x and t. No analytical solution to this equation is known, even for the idealized case of a linear load curve. Specific solutions may be obtained by means of computing devices, but there are too many parameters involved to prepare a useful library of specific solutions that would serve as a general treatment.

Approximate analytical solutions can be obtained by a step by step process, starting from the initial stage of flux development with the armature at rest. These solutions are relatively simple in form and application, and provide an analytical statement of the relations between performance and design parameters, from which conclusions can be drawn as to design features and operating conditions favoring a desired

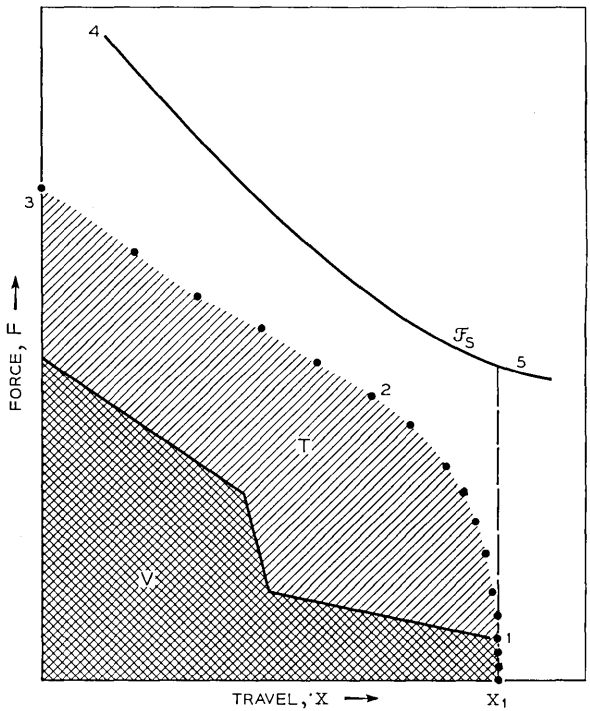


Fig. 4 — Load and pull relations in the operation of an electromagnet.

type of performance. For development purposes, an approximate solution indicating such optimum conditions is preferred to a rigorous solution in a form too complex for such use.

Three such approximate solutions for the operate case are described below. The first of these, the single stage approximation, is an application of the solution to the flux rise equation (2), neglecting the armature motion. It is applicable in cases of slow operation controlled by the rate at which the flux development provides sufficient pull to operate the load. The other two solutions are more generally applicable, particularly for relatively fast operation. One of these, the two-stage approximation, is the simpler in form, while the other, the three stage approximation, is the more accurate.

3 Single Stage Approximation

The initial stage of flux development with the armature at rest ends when the pull equals the back tension. During this stage the flux development is governed wholly by (2), with $x = x_1$, the initial gap, and $\mathcal{R} = \mathcal{R}_1$, as given by (4) for $x = x_1$. Writing $\varphi_1\mathcal{R}_1$ for $\mathcal{F}_S$, equation (2) becomes:

$$t = \frac{4\pi G}{\mathcal{R}_1} \int_0^{\varphi} \frac{d\varphi}{\varphi_1 - \varphi},$$

which on integration gives:

$$t = \frac{4\pi G}{\mathcal{R}_1} \ln \frac{1}{1 - v}, \tag{9}$$

where $v = \varphi/\varphi_1$, the ratio of the flux attained at time t to the steady state flux φ_1 or $\mathcal{F}_S/\mathcal{R}_1$. The mmf ratio $\mathcal{F}/\mathcal{F}_S$ is also equal to v.

Equation (9) applies rigorously only while the armature is at rest. It is, however, also a close approximation for the initial motion, in which the velocity is small and x differs little from x_1. If the rate of flux development is slow ($G/\mathcal{R}_1$ large) the armature moves slowly with the pull only slightly in excess of the load curve. In this case the inertia term $m \cdot d^2x/dt^2$ in (3) is minor, and the pull F nearly equals the static load dV/dx. Thus the operate time is approximately equal to the time required to develop a pull which exceeds the static load at all points in the travel. This pull is attained at the just operate current or ampere turn value, corresponding to the minimum mmf for static operation, $\mathcal{F}_0$. Assuming that the armature moves as the flux and pull develop, the operate time is that required for $\mathcal{F}$ to equal $\mathcal{F}_0$, and is given by equation

(9) for $v = \mathfrak{F}_0/\mathfrak{F}_S$. While this is a crude approximation, it provides satisfactory estimates of the time for slow operation.

EDDY CURRENT CONDUCTANCE DETERMINATION

Equation (9) is used in one experimental method for the evaluation of the eddy current conductance G_E. In this, measurements are made of the time at which motion starts for various values of the coil constant G_C. The latter may be varied by adding series resistance, and the applied voltage adjusted to maintain the current and hence $\mathfrak{F}_S$ constant in successive measurements. The time t for motion to start may be determined from shadowgraph measurements. For a constant back tension, the pull and hence the value of v or $\mathfrak{F}/\mathfrak{F}_S$ when motion starts is constant in these measurements. From (9), t is then directly proportional to $G_C + G_E$, so that a plot of t vs. G_C should be linear, with a negative intercept on the G_C axis numerically equal to G_E. Experimentally, an approximately linear relation is obtained in this way, provided the values of G_C covered are in excess of $5G_E$, in agreement with the discussion of Section 1. Values of G_E thus determined are consistent with those obtained from other measurements, and in approximate agreement with estimates obtained from (7) or (8).

COIL CONSTANT

In equation (9) G_E and $\mathfrak{R}_1$ are determined by the magnetic design. The latter, together with the load, determines $\mathfrak{F}_0$, the just operate mmf, or $4\pi(\mathrm{NI})_0$. As v equals $\mathfrak{F}_0/\mathfrak{F}_S$, or $(NI)_0/(NI)$, the only quantities not fixed by the design and the load requirement are G_C and the steady state ampere turn value NI. As the square of this latter quantity is equal to $G_C \cdot I^2R$, equation (9) may be written in the form:

$$t_0 = \frac{4\pi}{\mathfrak{R}_1}\left(G_E + \frac{(NI)_0^2}{v^2 I^2 R}\right)\ln\frac{1}{1-v}. \tag{10}$$

The steady state power I^2R is either determined by circuit requirements, or chosen with reference to economy of power consumption. With I^2R fixed, the only independent quantity in (10) is v, which is determined by the choice of the coil constant G_C. Neglecting G_E, the time is proportional to

$$\frac{1}{v^2}\ln\frac{1}{1-v},$$

which is shown plotted against v in Fig. 5. This figure includes a curve

giving corresponding values of ln $1/(1 - v)$, which may be used in specific cases to determine the correction corresponding to the term in G_E.

In most cases of slow operation, the operate time is of little importance, and G_C is chosen to make NI greater than $(NI)_0$ in all cases: i.e., after allowing for possible variations in $(NI)_0$ resulting from variations in the load and in the magnetic characteristics. These variations result in a variation in v, and the corresponding variation in time is shown by the curve of Fig. 5. If, however, it is desired to minimize the operate time for a given power input, G_C should be chosen to give the value of v (0.715) corresponding to the minimum of the curve. As this minimum is broad, variations in v, corresponding to those in $(NI)_0$, produce little change in tne operate time.

4 Two Stage Approximation

Unless the rate of flux development is very slow, as assumed in the single stage approximation, the pull attained in the early travel is in excess of the load, and the kinetic energy T is a considerable part of the total work output $V + T$. As a first approximation to this case, the operate time can be computed as though operation occurred in two successive stages: (1) a stage of flux development with the armature at rest in the unoperated position ($x = x_1$), and (2) a stage of motion, in which

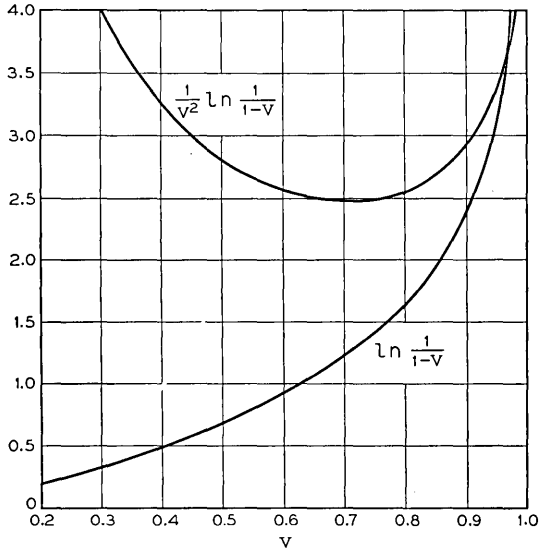


Fig. 5 — Relations for estimating time of flux development.

the flux remains constant at the value attained at the end of the first stage. This approximation necessarily over-estimates the operate time, as it takes the two stages as occurring in sequence, whereas they actually overlap.

Let t_1 be the time of the first stage, and t_2 the time of the second stage, where t_2 is the time for the motion from x_1 to some other armature position x_2, measuring the completion of operation. The first stage ends when the flux has attained some value $v\varphi_1$, where $\varphi_1 = \mathfrak{F}_S/\mathfrak{R}_1$, the steady state flux for the open gap, and v remains to be determined. Then t_1 is given by (9) for this value of v.

By hypothesis, the flux remains constant at $v\varphi_1$ throughout the second stage, and the corresponding pull is then given by (5A) for $\varphi = v\varphi_1$. The work $V + T$ is the integral of $F \cdot dx$ from x_1 to x_2, given by the integral of (5A). Hence:

$$V + T = \frac{(C_L - 1)v^2\varphi_1^2}{8\pi A} \frac{x_0^2(x_1 - x_2)}{(C_L x_0 + x_1)(C_L x_0 + x_2)}. \tag{11}$$

In this equation x_0/A may be replaced by the expression for $\mathfrak{R}_0$ given by (4A) for $x = x_1$, and $G_C \cdot I^2R$ substituted for $(NI)^2$ in the resulting equation, which then becomes:

$$v^2 = \frac{\mathfrak{R}_1 C_W}{2\pi G_C I^2 R}(V + T), \tag{12}$$

where C_W is given by:

$$C_W = \frac{(x_0 + x_1)(C_L x_0 + x_2)}{(C_L - 1)x_0(x_1 - x_2)}. \tag{13}$$

The time t_2 of the second stage depends upon the difference between the pull and load curves. As a representative condition, this may be taken as approximately constant. For such uniform acceleration of the effective mass m through the distance $x_1 - x_2$, the kinetic energy T equals $2m(x_1 - x_2)^2/t_2^2$, giving an expression for t_2 in terms of T. As the time t_1 of the first stage is given by (9), the total operate time t_0, or $t_1 + t_2$ is given by:

$$t_0 = \frac{4\pi(G_E + G_C)}{\mathfrak{R}_1} \ln \frac{1}{1 - v} + \left(\frac{2m(x_1 - x_2)^2}{T}\right)^{\frac{1}{2}}. \tag{14}$$

In this approximate expression for the operate time, v and T are as yet undetermined quantities, related by equation (12), in which G_C is the value of N^2/R for the coil used. The two stage operation assumed in deriving this expression corresponds to the existence of a restraint

which holds the armature at rest until it attains the flux level $v\varphi_1$. The approximation will most nearly approach the true relation if the effect of this restraint is minimized, and v taken at the value which minimizes t_0. This minimum value can be determined by trial, using the curves of Fig. 5 to compute t_0 for several values of v, and taking the minimum from the resulting plot of t_0 versus v.

In development studies, the operate time of interest is usually not that for an assigned winding, but that for the winding giving minimum time for a given power input. Substituting in (14) the expression for G_C given by (12), there is obtained the equation:

$$t_0 = \left(t_E + \frac{2C_W}{v^2}\,\frac{V+T}{I^2R}\right)\ln\frac{1}{1-v} + \left(\frac{2m(x_1-x_2)^2}{T}\right)^{\frac{1}{2}},$$

where t_E is written for $4\pi G_E/\mathcal{R}_1$, the eddy current time constant. As G_C remains to be determined, this expression contains two independent variables, v and T, which are to be so chosen as to make t_0 a minimum. Equating the partial derivative with respect to T to zero, there is obtained the following expression for the value of T for which the time is a minimum:

$$T = 2m(x_1-x_2)^2\left(\frac{v^2I^2R}{4C_W\ln\dfrac{1}{1-v}}\right)^2.$$

On substituting this expression for T, the preceding equation becomes:

$$t_0 = \left(t_E + \frac{2C_W}{v^2}\,\frac{V}{I^2R}\right)\ln\frac{1}{1-v} + \left(\frac{27C_W}{v^2}\ln\frac{1}{1-v}\,\frac{m(x_1-x_2)^2}{I^2R}\right)^{\frac{1}{3}}. \quad (15)$$

If the value of v which minimizes t_0 is determined, the resulting value of t_0 is the minimum operate time attainable by optimum coil design. In particular, if t_E is negligible, this minimum corresponds, from Fig. 5, to $v = 0.715$, for which

$$\frac{1}{v^2}\ln\frac{1}{1-v} = 2.5.$$

The corresponding value of G_C, the coil constant value making t_0 a minimum, is given by (12). In this use of equation (12), v and T are taken at the optimum values obtained as indicated above.

OPTIMUM COIL CONSTANT

As in the case of the single stage approximation, the optimum coil constant corresponds to a value for v of 0.715 when t_E is negligible. From

the curves of Fig. 5, it can be seen that if t_E is not negligible, the optimum value of v is less than 0.715, and will be smaller the larger the ratio of the first term in (15), that in t_E, to the other two terms. From (12), a reduction in v corresponds to an increase in NI, and hence in the coil constant G_C (I^2R being given). Thus the optimum coil constant for fast operation is larger for electromagnets with long cores of low resistivity material, for which G_E and t_E are large, than for those with short cores of high resistivity material, for which these quantities are small.

The physical significance of the optimum coil constant can be deduced from the relations shown graphically in Fig. 4. For a given value of the power input I^2R, an increase in G_C increases both the steady state mmf $\mathfrak{F}_S$ and the time constant of flux development $4\pi G_C/\mathfrak{R}_1$, increasing the upper limit to the attainable pull while reducing the rate at which this limit is approached. If the coil constant were so small that the area under the dynamic pull curve equalled the static work load V, the operate time would be infinite. On the other hand, an infinitely large coil constant would correspond to infinite inductance and an infinite operate time. Between these extremes lies the optimum value of the coil constant, giving minimum operate time. This is a broad optimum, near which a change in $\mathfrak{F}_S$ is compensated by a corresponding change in the rate of flux development, so that the realized dynamic pull curve is not affected by a small change in G_C.

In addition to the operate time, the value of G_C affects the final velocity of the armature, and thus its kinetic energy in impact with the core at the end of the stroke. This energy is dissipated in the relay and spring vibration associated with contact chatter. For values of G_C above the optimum, the higher inductance reduces the pull in the early travel, while the higher value of $\mathfrak{F}_S$ increases it in the later travel. The net effect is to increase both the operate time and the final velocity. Values of G_C above the optimum are therefore disadvantageous not only in slower operation, but in increased impact energy tending to cause contact chatter.

FACTORS CONTROLLING SPEED

Aside from the term in t_E, equation (15) shows that the minimum operate time, corresponding to the optimum value of v, is determined by the static load V, the inertia load measured by $m(x_1 - x_2)^2$, the steady state power I^2R, and the constant C_W. The latter is given by (13), and is the only quantity in (15), aside from t_E, which depends upon the magnetic design. In the range of values applying in practice, C_W is determined primarily by the leakage factor C_L, measuring the ratio of leakage to useful flux. In most practical cases, x_2 is small (zero for complete

operation), C_L equals or exceeds 4 ($\mathcal{R}_L$ equals or exceeds 3 $\mathcal{R}_0$), and x_1/x_0 lies in the range from 1 to 3. For $4 < C_L < 10$ and $1 < x_1/x_0 < 3$, $1.5 < C_W < 2.6$. In practice therefore C_W has a value close to 2.0 for most electromagnets.

The term in V, the second term in (15), varies inversely as the power I^2R, while the third, or inertia term, varies as the cube root of the power. Hence the second term tends to dominate at low values of I^2R, and the third term tends to dominate at high values. For a low power input therefore the operate time is controlled by the load V, and varies inversely as the power input, while for a high power input, the operate tine is controlled by the inertia term, $m(x_1 - x_2)^2$, and varies inversely as the cube root of the power. In the load controlled case, the time varies directly as the load; in the mass controlled case it varies as the cube root of the mass and as the two-thirds power of the travel $x_1 - x_2$. The eddy current term is very nearly a constant increment to the other two terms, and is therefore relatively more important, the faster the operation.

PRELIMINARY TIMING ESTIMATES

Within the limits of accuracy to which this two stage approximation applies, the effect of the magnet design upon the operate time appears only in t_E and C_W. If the former is neglected, the optimum value of v is 0.715, and

$$\frac{1}{v^2} \ln \frac{1}{1 - v} = 2.5 .$$

Taking C_W as having the representative value of 2.0, (15) reduces to:

$$t_0 = \frac{10V}{I^2R} + \left(\frac{135m(x_1 - x_2)^2}{I^2R}\right)^{\frac{1}{3}}. \tag{16}$$

This simple approximation provides rough estimates of the operate times attainable with any electromagnet for a given load and power input. As an illustration, consider a typical relay spring load involving a travel of 40 mil-in (0.1 cm) with a level of spring force of 100 gm (10^5 dynes), so that $V = 10^4$ ergs. Let the effective mass m of the moving parts be 10 gm. For a steady state power input of 0.5 watts (5×10^6 ergs sec.), the two terms of (16) have values of 0.020 sec and 0.014 sec, respectively, so that t_0 equals 0.034 sec. For an input of 5 watts, the two terms become 0.002 sec and 0.007 sec, respectively, and t_0 equals 0.009 sec. The neglected t_E term of (15) might amount to an increment of 0.005 sec for a solid core of magnetic iron, but would be proportionally smaller for higher

resistivity material. In either case, this increment would be of little consequence in the low power case, but would materially affect the operate time in the high power case.

5 Three Stage Approximation

The following analysis provides greater accuracy in estimating operate time than the two stage approximation, together with a more accurate representation of the relations between performance and the controlling variables. It differs from the two stage approximation in treating the initial motion as a separate stage of operation. The three stages of operation thus become: (1) Increase of the flux to the value at which the pull equals the back tension, with the armature at rest, (2) flux development and concurrent armature acceleration, assuming equations (2) and (3) to apply, with the approximation that the variations of $\mathfrak{R}$ and $\mathfrak{F}$ with x are ignored, and (3) the later motion, which is treated in the same manner as the second stage of the two stage approximation.

Formally, the second stage should be restricted to a very small part of the total armature motion, in order to minimize the change in $\mathfrak{R}$ which is ignored. Relatively minor error, however, is introduced in ignoring the variation in $\mathfrak{R}$ through as much as half the total travel. The spring force is taken as constant through the second stage. In the relay case, this is approximately true for the initial travel prior to the actuation of the contacts, which results in an abrupt increase in the spring load. Thus the second stage can be taken to extend to the travel at which contact actuation first occurs. If all contacts are actuated near the same point in the travel, the end of the second stage coincides with complete operation, and the third stage need not be considered. If the contacts are spread out, the third stage coincides with the stagger time, or time between operation of the first and last contacts. The operate time therefore consists either entirely or principally of the time for the first two stages.

As before let:

x_1 be the initial (open) gap, for which $\mathfrak{R} = \mathfrak{R}_1$,

F_1 be the back tension (constant load in the first two stages),

t_C be the initial coil time constant, $\dfrac{4\pi}{\mathfrak{R}_1} \cdot \dfrac{N^2}{\mathrm{R}}$,

t_E be the initial eddy current time constant, $\dfrac{4\pi}{\mathfrak{R}_1} G_E$,

$\mathfrak{F}_S$ be the steady state mmf, $4\pi NI$, and φ_1 the corresponding flux for $x = x_1$, or $\mathfrak{F}_S/\mathfrak{R}_1$.

FIRST STAGE

The pull for $x = x_1$ is given by (5A), which may be written:

$$F = \frac{(v\varphi_1)^2}{8\pi A_1}, \tag{17}$$

where A_1 is given by:

$$A_1 = \left(\frac{C_L x_0 + x_1}{x_0(C_L - 1)}\right)^2 A. \tag{18}$$

In the first stage the flux increases to $v_1\varphi_1$, where v_1 is given by (17) for $F = F_1$. Then the time t_1 of the first stage is given by (9) for $v = v_1$, or by:

$$t_1 = (t_C + t_E) \ln \frac{1}{1 - v_1}. \tag{19}$$

The solution for the first stage therefore determines v_1 and t_1.

SECOND STAGE

It is convenient to write the equations for the second stage in terms of the ratio z defined by:

$$z = \frac{t}{t_C + t_E},$$

where t is now measured from the start of the second stage. Neglecting the change in reluctance with armature motion, the ratio φ/φ_1 or v is given by (9), except that for $v = v_1$ at $t = 0$ the relation must be written:

$$v = 1 - (1 - v_1)e^{-z}. \tag{20}$$

$$\text{As } \varphi_1 = \mathfrak{F}_S/\mathfrak{R}_1, \quad \text{or} \quad 4\pi NI/\mathfrak{R}_1, \quad \text{and} \quad t_C = \frac{4\pi}{\mathfrak{R}_1} \cdot \frac{N^2}{R},$$

$$\varphi_1^2 = \frac{4\pi}{\mathfrak{R}_1} I^2 R \cdot t_C.$$

Taking the pull in the second stage as given by the expression applying for $x = x_1$, and substituting the preceding expressions for v and φ_1, there is obtained:

$$F = \frac{I^2 R t_C}{2\mathfrak{R}_1 A_1} (1 - (1 - v_1)e^{-z})^2. \tag{21}$$

For the second stage dV/dx is constant at $-F_1$ throughout, and (3) reduces to $m\, d^2x/dt^2 = F - F_1$, or:

$$\frac{d^2x}{dt^2} = \frac{I^2Rt_C}{2m\mathcal{R}_1A_1}((1 - (1 - v_1)e^{-z})^2 - v_1^2).$$

Substituting $(t_C + t_E)\ z$ for t, this equation may be integrated with respect to t to obtain the following expressions for the velocity $\dot{x}$ and the travel x in the second stage:

$$\dot{x} = \frac{I^2Rt_C(t_C + t_E)}{2m\mathcal{R}_1A_1} f_1(v_1, z), \tag{22}$$

$$x_1 - x = \frac{I^2Rt_C(t_C + t_E)^2}{2m\mathcal{R}_1A_1} f_2(v_1, z). \tag{23}$$

The functions $f_1(v_1, z)$ and $f_2(v_1, z)$ are the integrals with respect to z of the bracketed term in the preceding equation. These functions can be evaluated from the curves of Fig. 6, which shows them plotted against z for several values of v_1.

For a specific case, the values of v_1 and t_1 will have been determined from the relations for the first stage. Let x_2 be the travel at the end of the second stage. Then the value of $f_2(v_1, z)$ for $x = x_2$ is given by (23), and z_2, the corresponding value of z, can be read from the curve of Fig. 6 for the value of v_1 applying. Then the time for the second stage, t_2, is given by $z_2(t_C + t_E)$. For $z = z_2$, (20) gives v_2, the value of v at the end of the second stage. This in turn gives the corresponding value of the flux $v_2\varphi_1$, and, from (21), the pull F at the end of this stage. The velocity at the end of this stage is given by (22), with $f_1(v_1, z)$ read from the curves of Fig. 6 for $z = z_2$.

THIRD STAGE

The total operate time is $t_1 + t_2 + t_3$, where t_3 is the time for the third stage. In the relay case, this is the stagger time between the operation of the first and last contacts. A first, and frequently adequate approximation to t_3, is given by assuming the velocity in the third stage to be constant at the value attained at the end of the second stage. For a more exact determination, particularly in determining the final velocity for complete operation, it may be assumed that the flux in the third stage is constant at the value $v_2\varphi_1$ attained at the end of the second stage. The mechanical output $V + T$, in the third stage is then given by equation (11), with v taken as v_2, and x_1 and x_2 replaced, respectively, by x_2 and x_3, the latter denoting the travel at the end of the third stage (zero for complete operation). Knowing the spring load V between x_2 and x_3, t_3 and the final velocity can be computed from the increase in kinetic energy T on the assumption of uniform acceleration.

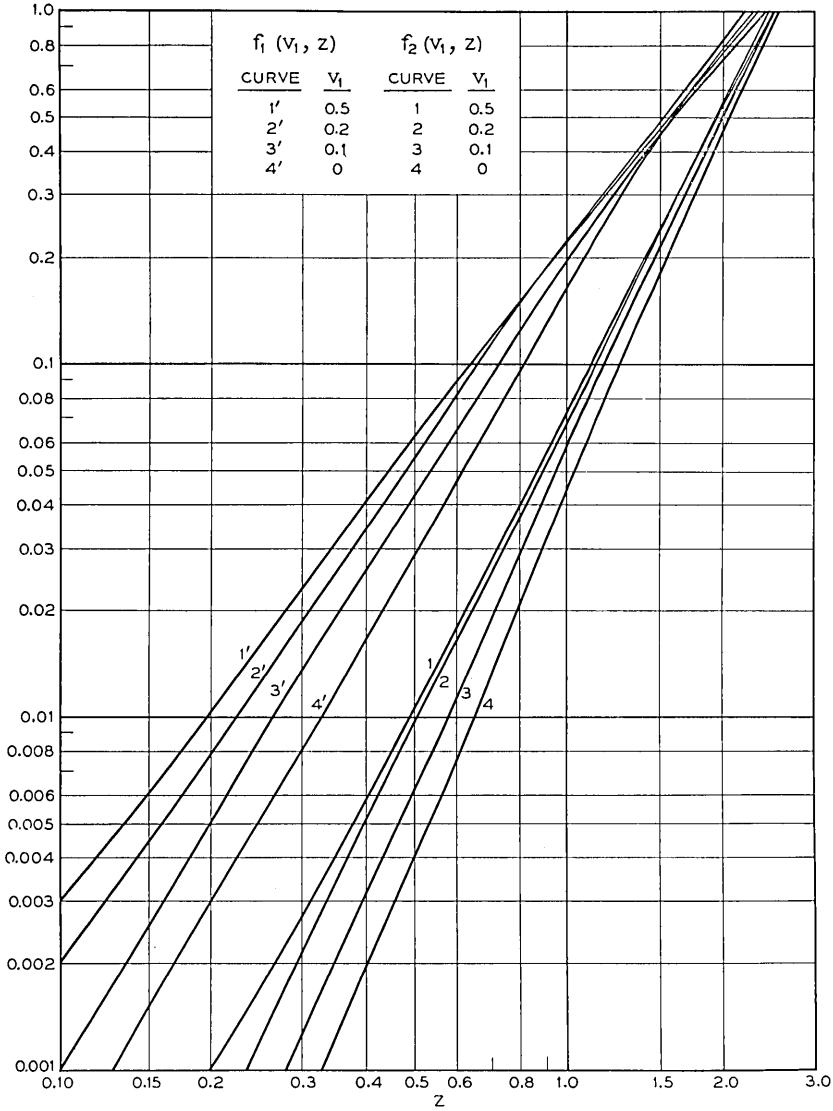


Fig. 6 — Relations for initial motion in operation.

FACTORS AFFECTING OPERATE PERFORMANCE

The dominant term in the operate time is the time t_2 for the second stage, which also determines the velocity and pull level in the third stage, and hence the final velocity. As $t_2 = z_2(t_C + t_E)$, equation (23) gives the following expression for t_2 :

$$t_2^3 = \frac{2mA_1\mathcal{R}_1(x_1 - x_2)}{I^2R} \frac{t_C + t_E}{t_C} \frac{z_2^3}{f_2(v_1, z_2)}. \tag{24}$$

As can be seen in Fig. 6, $f_2(v_1, z)$ is, to a rough first approximation, proportional to z^3. To the extent that this approximation holds, t_2 is proportional to the cube root of $mA_1\mathcal{R}_1(x_1 - x_2)/(I^2R)$, as for the last term in (15) and (16), except that $(x_1 - x_2)^2$ is replaced by $A_1\mathcal{R}_1(x_1 - x_2)$. From (4A) and (18), $A_1\mathcal{R}_1$ is given by:

$$A_1\mathcal{R}_1 = \frac{(x_0 + x_1)(C_Lx_0 + x_1)}{(C - 1)x_0}.$$

Here $x_0 = A\mathcal{R}_0$, and t_2 is a minimum for that pole face area A for which $A_1\mathcal{R}_1$ is a minimum. This condition, as determined by equating the derivative of the preceding expression with respect to x_0 to zero, is given by:

$$x_1^2 = C_Lx_0^2,$$

or $x_1/A = \sqrt{C_L} \cdot \mathcal{R}_0$, where A is the optimum pole face area for fast operation. This is a smaller pole face area than that for maximum sensitivity, for which x_1/A should equal $\mathcal{R}_0$, as shown in the companion article[1] cited above. If the pole face area satisfies the preceding condition, the expression for A_1 reduces to:

$$A_1 = \frac{\sqrt{C_L} + 1}{\sqrt{C_L} - 1} x_1,$$

and the term $A_1\mathcal{R}_1$ in (24) reduces to $x_1(x_1 - x_2)$ multiplied by the factor $(\sqrt{C_L} + 1)/(\sqrt{C_L} - 1)$. This leakage factor, which increases with the ratio of leakage to useful flux, is then the only term in (24) which varies with the design of the electromagnet.

If $f_2(v_1, z)$ were strictly proportional to z^3, equation (24) would be nearly independent of the coil constant N^2/R, to which t_C is proportional. As $f_2(v_1, z)$ departs from this cube law relation, and also varies with v_1, t_2 is not independent of the coil constant. The optimum coil constant is that which minimizes $t_1 + t_2$. In any specific case, a succession of values may be assumed for G_C, and t_1 and t_2 evaluated by the relations given above. The resulting relation between $t_1 + t_2$ and G_C is similar in character to that described above in connection with the two stage approximation.

6 Design for Fast Operation

The approximations discussed above provide a means for estimating the operate time attainable in specific cases, and for selecting certain

design variables to obtain maximum speed. In particular, it has been shown that for a given relay, with a specified load and power input, a winding can be selected to minimize the operate time. Aside from this, the relations show that for a given relay, the operate time depends only upon the load and travel and the power input, varying inversely as the latter at low levels of power input, and inversely as its cube root at high levels.

Further conclusions can be drawn from these relations with reference to development studies of relays and other electromagnets. The spring load and travel may be considered as fixed requirements, so far as the magnet design is concerned. The available power may be fixed by circuit considerations, or it may be related to the design by a requirement that the winding dissipate this power in the holding interval, a condition that imposes a minimum size on the coil and the magnet structure. Subject to this and some other limitations, there is a design choice of the dimensions and configuration which determine the magnetic circuit constants, the mass of the armature, and the eddy current conductance.

The preceding discussion has shown that, with an optimum choice of pole face area, the magnetic characteristics affect the time only with respect to the leakage factor, the ratio of leakage to useful flux. This factor may be reduced by using a "tight" magnetic circuit, but if this is done the factor tends to vary directly as the length of the magnetic path and inversely as the separation of the core and return members in relation to their cross sections. The leakage may be minimized by using a square outline for the magnetic path. The optimum speed magnet then has a specific configuration in which all dimensions are fixed in relation to the cross section of the magnetic members. This dimension then determines the mass of the armature.

For this optimum configuration, the power, the spring load, the mass, and the travel determine a level of pull which minimizes the operate time. This pull requires a certain armature cross section if saturation is to be avoided, as the effective pole face area is fixed in relation to the cross section. If the resulting armature mass is minor compared with the mass of the load, the operate time attainable varies with the cube root of the applied power, as in the cases discussed above.

With increased power, however, the optimum pull and armature cross section increase, and must eventually reach a level where the armature mass becomes the dominant portion of the mass term. As this condition is approached, any increase in power is offset by an increase in armature mass, such that a lower limit is imposed on the operate time proportional to the travel, corresponding to an upper limit to the average armature

velocity attainable with a neutral electromagnet. This upper limit is of the order of 100 cm/sec. Thus, for example, an operate time less than 0.25 millisec cannot be attained with an armature travel of 10 mil-in, no matter how large the steady state power applied.

7 Release Waiting Time

Like operation, release is made up of an initial stage of flux change with the armature at rest, followed by a stage of armature motion, and the total time is the sum of the waiting time and the motion time. In release, the waiting time is usually larger than the motion time.

There are three distinct circuit conditions under which release occurs. These are:

Normal, or unprotected release, in which the coil circuit is open, and the only magnetomotive force maintaining the field is that of the eddy currents.

Protected release, in which the coil circuit is closed through a protective shunt, usually comprising a condenser and a resistance in series. The magnetomotive force comprises that of the eddy currents and that of the coil circuit transient.

Slow release, in which a sleeve or short circuited winding is used to maintain the field and delay release. The magnetomotive force is predominantly that of the sleeve or winding current.

The slow release case is the only one of the three for which the flux decay relation is accurately represented by equation (2). In protected release, the coil current transient is controlled by the condenser, as discussed below. In normal release, the variation in the field linked by different eddy current paths results in the changes in the field pattern discussed in Section 1. As noted there, equation (2) applies to this case only as a very crude first approximation.

In all three cases, the relation between φ and $\mathfrak{F}$ applying is that for decreasing magnetization, as illustrated by the curve for $x = 0$ in Fig. 7. Unlike the linear relation applying in operate, corresponding to the constant reluctance given by (4), the decreasing magnetization curve has a hyperbolic character, and is asymptotic to the saturation flux φ''. Residual magnetism results in a residual flux φ_0, the intercept of the decreasing magnetization curve on the φ axis. An analytical treatment of the decreasing magnetization curve is given in a companion article,[5] where it is shown to provide a satisfactory basis for predicting the release time of slow release relays, the third case above. The other two cases, to which the following discussion is confined, involve both non-linear mag-

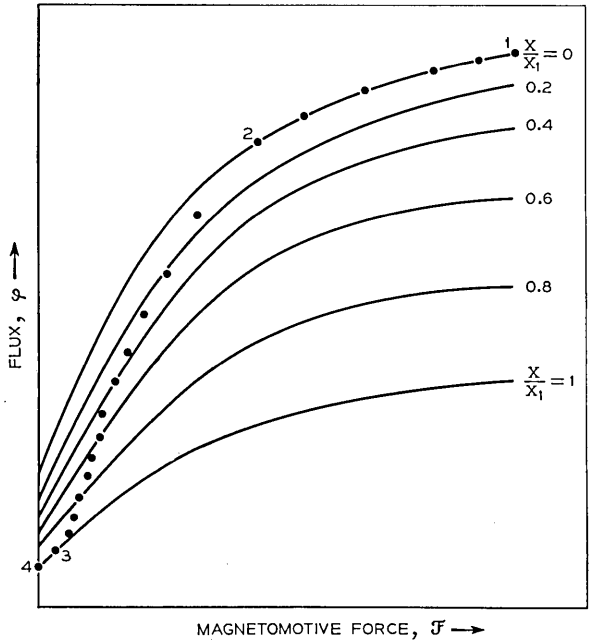


Fig. 7 — Field energy relations in release.

netization and a variable magnetic field pattern. In an approximation neglecting the latter effect, no advantage is derived from an accurate formulation of the former, and a linear approximation to the demagnetization curve may be employed.

NORMAL UNPROTECTED RELEASE

As indicated above, a rough approximation to the decay of the gap flux, and hence of the pull, may be obtained by assuming that this flux decays in conformity with equation (2), with $G = G_E$, where the eddy current conductance G_E has the same value as in operate. As a linear approximation to the demagnetization curve, the reluctance $\mathcal{R}$ can be taken as the closed gap reluctance $\mathcal{R}(0)$. Allowance may be made for residual magnetism by postulating a steady state magnetomotive force $\mathcal{F}_0 = \mathcal{R}(0)\varphi_0$. Then integration of (2) gives the following expression for the time t at which the flux has decayed to the value φ from its initial value φ_1 :

$$t = t_E \ln \frac{\varphi_1 - \varphi_0}{\varphi - \varphi_0}, \tag{25}$$

where $t_E = 4\pi G_E/\mathcal{R}(0)$. Unless φ/φ_0 is near unity, φ_0 may be omitted from this expression. If this is done, the expression may be re-written in terms of pull F in the form:

$$t = \frac{t_E}{2} \ln \frac{F_1}{F}, \tag{26}$$

where F is the pull at time t, and F_1 the initial pull. This expression follows from (25), for $\varphi_0 = 0$, because of the proportionality between F and φ^2. If F is the spring load or tension for the operated position, and F_1 the steady state pull there, (26) is an approximate expression for the release time.

PROTECTED RELEASE

The commonly employed method of contact protection is to use a condenser-resistance shunt connected either across the coil, as in Fig. 8, or across the contact. Except for the steady state voltage of the condenser, the circuit relations are the same for the two cases. As a first approximation, eddy currents may be represented by the current in a short circuited secondary, as in Fig. 8. The effect of such a secondary is determined by the value of t_E representing the ratio of its inductance to its resistance, as this determines the ratio of its contribution to the flux to the time rate of change of the total field.

If perfect coupling and linear magnetization are assumed to apply, the circuit equations for Fig. 8 can be solved and expressions obtained for the flux, current and condenser voltage as functions of time after the opening of the contact. These expressions are similar to those for the simple C-R-L circuit without a secondary, except that they involve the time constant t_E as well as CR and L/R. The discharge may be either exponential or a damped oscillation, but while the discharge of the simple C-R-L circuit is always oscillatory for small values of C, the discharge in the circuit of Fig. 8 is only oscillatory for an intermediate

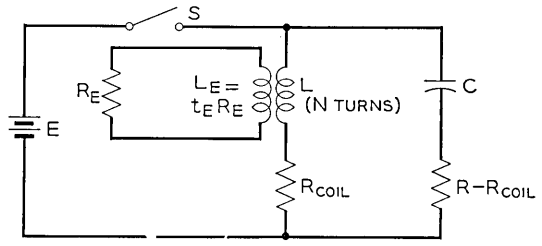


Fig. 8 — Coil circuit with capacitative contact protection.

range of capacity values. For large values of C, the discharge is damped by the coil resistance, and for small values of C it is damped by the currents in the secondary.

The latter case is essentially the condition applying for the small capacity values used for contact protection. Initially, the flux decay is opposed by the magnetomotive force resulting from both the eddy currents and the coil current. The latter discharges the condenser and develops a charge of opposite polarity. The initial rate of flux decay is therefore similar to that for exponential decay with a conductance equal to $G_E + N^2/R$, where R includes the external resistor as well as the coil resistance. The effect of the condenser charge is equivalent to a continuously increasing coil resistance, so that the decay approaches the rate that would apply for exponential decay with $G = G_E$, and the decay rate is then further increased by the reverse magnetomotive force resulting from the coil current reversal in the subsequent discharge of the condenser.

If CR is less than t_E, the effect of the coil circuit is to change the character of the flux versus time relation, without materially changing the time scale of the discharge. The initial decay is delayed, while the later decay is accelerated. Hence the release time is increased for a heavy load, corresponding to a high value of φ for release, while it may be slightly decreased for a light load, corresponding to a small value of φ. The utility of the coil-condenser circuit for contact protection results from the low initial rate of flux decay, which holds the induced voltage across the contact to a relatively low value in the initial stage of contact opening, when the contact separation is small.

Except initially, the predominant magnetomotive force is that of the eddy currents, and this results in a changing distribution of field intensity, as in the case of simple release. Hence the analysis of Fig. 8, as described above, is not applicable quantitatively, and is therefore not given here in detail. Qualitatively, the relations are similar, the protected release has a flux time relation similar in time scale to that for simple release, with a lower initial rate of decay, and a higher rate for the later stage. An illustration of this effect is included in the article by M. A. Logan[4] cited above. No analytical treatment is available for determining the differences in release time between simple and protected release, at least for values of CR of the same order at t_E, as in the protection networks commonly employed.

8 Release Motion

It was stated in Section 7 that the waiting time in release is usually larger than the motion time. This, however, is not necessarily or in-

variably the case, and interest attaches to the conditions under which the motion time may become relatively large. Another important aspect of the release motion is the final velocity attained when the armature strikes the backstop. This impact results in a rebound which may, in the relay case, result in re-actuation of the contacts: rebound chatter. Rebound may be reduced by appropriate mechanical design, as discussed in an article by E. E. Sumner,[8] but its amplitude is always, other things being equal, proportional to the kinetic energy of the armature in impact on the backstop.

Fig. 9 shows the force travel relations in release, corresponding to the operate relations of Fig. 4. The solid line, as in Fig. 4, represents the spring load, which has an operated value F_0 at the point marked 2. This corresponds to the similarly marked point in Fig. 7, which shows the corresponding φ versus $\mathfrak{F}$ relations. The field decays along the demagnetization curve for $x = 0$ to the point 2, where load and pull are in equilibrium. Further decay results in a pull less than the spring load, and hence in a net accelerating force producing armature motion. The flux continues to decrease as indicated in Fig. 7, following the path 2–3–4. The pull follows the similarly marked curve in Fig. 9, related to the flux-travel path by equation (5).

The net accelerating energy is therefore represented by the area marked T, lying between the pull curve and the spring load. Thus only this portion of the energy V stored in the spring load appears as kinetic energy of the armature. The remainder of V is represented by the area

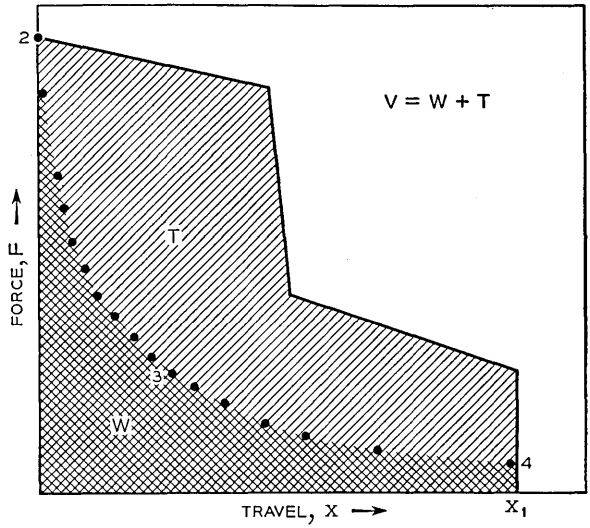


Fig. 9 — Load and pull relations in release.

marked W, lying below the release pull curve. This energy may be termed the magnetic drag: it is restored to the magnetic field and dissipated in the eddy current paths or in other circuits linking the field. Experimental studies have shown that the kinetic energy T of the armature in backstop impact may vary from 20 per cent to 90 per cent of V. Such a wide variation in armature energy has a proportional effect on rebound amplitude. Within certain limitations, the following analysis establishes the relations controlling the ratio W/V, and also serves for the evaluation of the release motion time.

EQUATIONS APPLYING

As the motion starts after an initial period of field decay, the relations applying are given approximately by (2) and (3), with $\mathcal{R}$ and F given by (4A) and (5A), and $G = G_E$. The use of (4A) and (5A) is justified by the fact that the lower portion of the demagnetization curve, which applies to the later stages of flux decay, is approximately linear. The assumption that (2) applies, with $G = G_E$, is justified by the fact, discussed in Section 1, that the initial rapid decay of the field in the outer layers of the core is followed by an exponential decay of the greater part (about 70 per cent) of the initial field. Thus if t_E is written for $4\pi G_E/\mathcal{R}(0)$ in (2), this expression applies in the latter stages of flux decay when the armature is at rest at $x = 0$. In this case, the value of t_E is a constant, even if the values of G_E and $\mathcal{R}(0)$ applying differ somewhat from those applying in the operate case.

Taking these equations to apply, it is desired to determine the solution for the initial conditions applying to the release case. This cannot be obtained by the approximations used for the operate case, as the simplifying condition of a nearly constant reluctance does not apply: the change of reluctance with x is a maximum for $x = 0$. The procedure that has been employed is the use of the analogue computer described in articles by A. A. Currie[6] and E. Lakatos[7] to obtain solutions for a restricted category of cases. The equations may be reduced to a form adapted to such solution as follows:

An expression for $\mathcal{R}/\mathcal{R}(0)$ obtained from (4A) is substituted in (2). Writing t_E for $4\pi G_E/\mathcal{R}(0)$, there is obtained:

$$\frac{C_L(1+u)}{C_L u}\,\varphi^2 + \frac{t_E}{2}\,\frac{d\varphi^2}{dt} = 0,$$

where u is written for x/x_0, or $x/(A\mathcal{R}_0)$. By substituting in (3) the expression for F given by (5A) there is obtained an expression for φ^2.

On substituting this in the preceding equation, there is obtained:

$$(1 + u)(C_L + u)\left(\frac{dV}{dx} - m\frac{d^2x}{dt^2}\right) + \frac{t_E}{2C_L}\frac{d}{dt}\left((C_L + u)^2\left(\frac{dV}{dx} - m\frac{d^2x}{dt^2}\right)\right) = 0.$$

To reduce this expression to dimensionless form, the following substitutions are made:

f is written for $\frac{dV}{dx} \Big/ F_0$, where F_0 is the operated load, so that $f = 1$ at the start of the release motion.

τ is written for $2C_L\, t/t_E$, expressing the time t as a multiple of $t_E/(2C_L)$.

t_M^2 is written for $2mA\mathcal{R}_0/F_0$. Thus t_M is the time for travel of the mass through the distance $A\mathcal{R}_0$ for a constant accelerating force F_0.

K is written for $2C_L\, t_M^2/t_E^2$.

The preceding equation then becomes:

$$(1 + u)(C_L + u)\left(f - K\frac{d^2u}{d\tau^2}\right) + \frac{d}{d\tau}\left((C_L + u)^2\left(f - K\frac{d^2u}{d\tau^2}\right)\right) = 0. \tag{27}$$

This is the form of the release motion equation to which analogue computer solutions were obtained. It is a third order differential equation giving the travel, expressed as a multiple u of $A\mathcal{R}_0$, as a function of the time expressed as a multiple τ of $t_E/(2C_L)$. The boundary conditions correspond to zero initial values of travel, velocity, and acceleration, so that for $\tau = 0$, $u = du/d\tau = d^2u/d\tau^2 = 0$.

ANALOGUE COMPUTER SOLUTION

The analogue computer gives solutions to specific cases, and a specific case of (27) is defined by the values of C_L and t_M/t_E applying, and by the form of the relation between f and u, defining the shape of the spring load. To confine the cases considered to a manageable category, they were limited to those for which $f = 1$ for all values of u: the case of a constant spring load.

The constant C_L, the ratio $(\mathcal{R}_0 + \mathcal{R}_L)/\mathcal{R}_0$, is an inverse measure of the relative magnitude of the leakage field. As noted in Section 2, its value is usually in excess of 4. Solutions were obtained for two cases: $C_L = 3$ and $C_L = 5$. For $f = 1$, and a fixed value of C_L, the remaining param-

eter defining a specific case is t_M/t_E, or the corresponding value of K $(2C_L t_M^2/t_E^2)$. For the two selected values of C_L, solutions were obtained for the following values of K: 0.05, 0.1, 0.3, 1.0 and 10.

The computer set-up for this problem is indicated in Fig. 9 of the article by E. Lakatos cited above.[7] Solutions were obtained to (27) for $f = 1$ and the values of C_L and K cited above, for the given boundary conditions. The solutions were obtained in the form of machine drawn curves giving, u versus τ and $du/d\tau$ versus τ, over the range $0 < u < 3$. In accordance with the character of the net accelerating force, all these curves were smooth and monotonic.

For the conditions covered by the solution, these results give the motion time for a given travel directly in the form of the relation between τ and u. It is convenient to express the values of τ in terms of the corresponding values of t/t_M. The results for $C_L = 5$ are shown in Fig. 10 as curves of t/t_M versus u for various values of t_M/t_E.

The other result of interest is the ratio of the magnetic drag W to the spring load energy V, or of $(V - T)/V$, where T is the kinetic energy at the end of a given travel: the total travel in an actual case. T equals $m(dx/dt)^2/2$, or $2m(C_L A\mathcal{R}_0 \cdot du/d\tau)^2/t_E^2$. For $f = 1$, V equals $F_0 x$, or $F_0 u A\mathcal{R}_0$. It follows that T/V is given by:

$$\frac{T}{V} = \frac{mA\mathcal{R}_0}{2F_0 u}\left(\frac{2C_L}{t_E}\right)^2\left(\frac{du}{d\tau}\right)^2,$$

and therefore by:

$$\frac{T}{V} = \frac{K}{2u}\left(\frac{du}{d\tau}\right)^2. \tag{28}$$

Corresponding values of u and $du/d\tau$ were read from the computer results and substituted in (28) to determine T/V and thus W/V. The resulting values of W/V for the case $C_L = 5$ are shown in Fig. 11 plotted against t_M/t_E for various values of u.

DISCUSSION OF COMPUTER RESULTS

The significance of the results shown in Figs. 10 and 11 can be more readily grasped by reference to representative values of the parameters involved. For relay electromagnets, representative values of $\mathcal{R}_0$ and A are 0.04 cm^{-1} and 1 cm^2, respectively, for which $A\mathcal{R}_0 = 0.04$ cm, or 16 mil-in. This distance is, for this case, the travel for which $u = 1$. For this value of $A\mathcal{R}_0$, an effective mass m of 10 gm, and an operated load F_0 of 2×10^5 dynes (200 gm wt), $t_M = 2$ millisec. If the load were

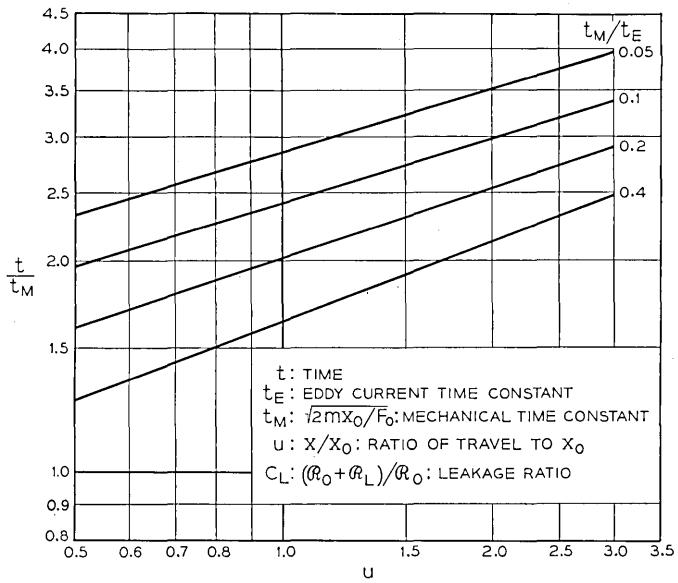


Fig. 10 — Release motion time relations.

doubled and the mass halved, t_M would be halved. These values of t_M are of the same order but smaller than the values of t_E generally applying, corresponding to the range in t_M/t_E covered by the figures. Similarly, the total travel usually lies in the range covered by the values of u, corresponding to travels up to 48 mils for the above value of $A\mathcal{R}_0$.

The curves of Fig. 11 show that the fraction W/V of the spring load energy absorbed by magnetic drag varies from 15 to 80 per cent over the range covered, and consequently that the kinetic energy at the end of the stroke, which determines the rebound amplitude, varies from 85 per cent to 20 per cent of V, or by a factor of 4 to 1.

The way in which W/V varies with u and with t_M/t_E is readily understood from physical considerations. Referring to Fig. 9, it is apparent that for a given time rate of field decay, the faster the rate of armature motion, the higher is the restraining pull and the less the kinetic energy T. For a short travel, or low value of u, the restraining pull will be larger than for a high value of u: hence W/V decreases as u is increased. As t_M measures the time for a travel of $A\mathcal{R}_0$, or $u = 1$, for a force F_0, t_M/t_E is a measure of the motion time relative to the time of field decay. Hence a large value of t_M/t_E corresponds to a condition where the motion is slow relative to the field decay, giving little restraining pull. Thus W/V decreases as t_M/t_E increases.

The effect of the magnetic drag on the motion time is apparent in the curves of Fig. 10. The motion time for a given value of u increases as t_M/t_E decreases, corresponding to increased drag. Thus a reduction in rebound by increased magnetic drag is, of course, accompanied by a longer motion time. It is of interest to note that the increasing acceleration results in an armature displacement which varies as the cube of the time, as in the relation applying in operate.

The results shown are for the case $C_L = 5$. For $C_L = 3$, corresponding to higher leakage, the results are similar, but the drag ratio W/V values are from 5 to 10 per cent lower than those for $C_L = 5$, and the values of t/t_M are correspondingly smaller. Thus the value of C_L applying has only a secondary effect on the release motion.

While these results are formally limited to the special case of a constant retractile force, the conclusions are of broader application. A load curve that decreases as the gap opens modifies the solution for the same operated load F_0 only in reducing the net accelerating force, increasing the drag ratio and the motion time. The residual magnetism, neglected in this discussion, has a similar, and minor, effect. The results show C_L to have little effect on the motion, which is governed primarily by the values of u and t_M/t_E applying.

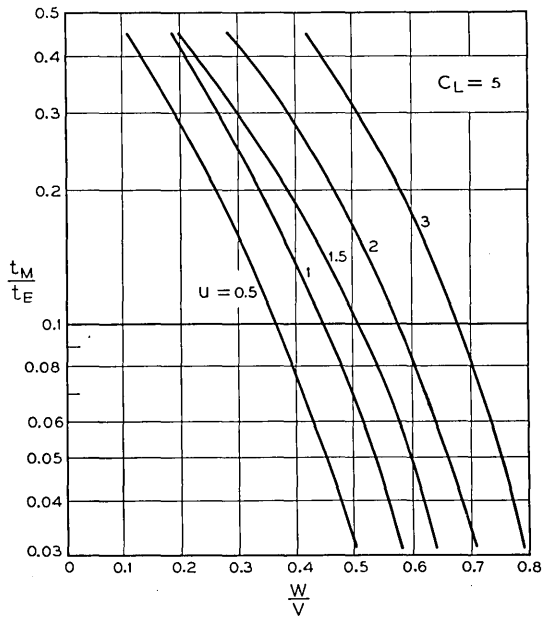


Fig. 11 — Energy absorbed by magnetic drag in release motion.

As minimizing rebound is of major importance in relay design, the relations controlling the kinetic energy to which it is proportional are of design interest. Most of the quantities appearing in the above relations, such as the spring load and the travel, are fixed by operating requirements or other design considerations. There is, however, considerable freedom in the design relations fixing the effective mass of the armature, particularly in selecting its axis of rotation. The preceding relations show that the magnetic drag ratio can be increased, and rebound reduced, by the decrease in t_M resulting from a decrease in effective mass. While such a decrease increases the ratio t/t_M, it results in a net decrease in the value of t, the motion time. A low effective armature mass is therefore advantageous both for fast release and for a high drag ratio for the reduction of rebound.

REFERENCES

1. R. L. Peek, Jr., and H. N. Wagar, Magnetic Design of Relays, page 23 of this issue.
2. R. L. Peek, Jr., Analysis of Measured Magnetization and Pull Characteristics, page 78 of this issue.
3. B. Wwedensky, Ann. d. Phys., **64,** p. 609, 1921.
4. M. A. Logan, Dynamic Measurements of Electromagnetic Devices, B. S. T. J., p. 1413, Nov., 1953.
5. R. L. Peek, Jr., Principles of Slow Release Relay Design, page 187 of this issue.
6. A. A. Currie, The General Purpose Analog Computer, Bell Lab. Record, **29,** pp. 101–108, Mar., 1951.
7. E. Lakatos, Problem Solving with the Analog Computer, Bell Lab. Record, **29,** pp. 109–114, Mar., 1951.
8. E. E. Sumner, Relay Armature Rebound Analysis, B. S. T. J., **31,** pp. 172–200, Jan., 1952.

Estimation and Control of the Operate Time of Relays

Part II — Design of Optimum Windings

By M. A. LOGAN

(Manuscript received September 4, 1953)

For each relay structure, there exists a best winding, once the circuit power has been specified. The best winding is that one which operates the relay in the least time. Methods for determining the best winding are developed.

On the basis of the operating time, the relay behavior is classed as mass or load controlled. The design method chosen depends upon this classification.

The design of windings for series connected relays is based on a method of determining equivalent single relays, the behavior of each corresponding to one of the series relays. This method is generalized to allow for different magnetic structures for the several relays. Each relay winding is then designed in turn, using the design data for its own type of structure.

The best winding design is not given directly by an explicit formula. Rather, methods are developed for determining the operate time for any winding. Then by choosing a range of windings, the best one is selected by interpolation.

INTRODUCTION

The selection of a relay for a circuit application, particularly so in common control systems, involves in part a determination of its operating and releasing time. In Part I of this article expressions are derived for the relations between these times, the design parameters of a relay, and the conditions of operation. These expressions are approximate, and have been developed with primary reference to the selection of favorable characteristics in design. The timing estimates they provide are of sufficient accuracy for the comparison of design alternatives. Once a basic design has been selected, the choice of windings and the prediction of the limiting times occurring in specific applications requires

a more detailed and exact procedure than can be obtained entirely from the application of these approximate relations. The information furthermore has to be quite versatile, to permit all new type variations to be included, such as winding, number of contacts and armature travel. It must, moreover, be in a form that permits determination of the upper and lower limits to the time observed in each specific type of relay as actually used, as in applications, interest attaches not only to the nominal conditions, but to all variations that may arise in actual manufacture and use.

Relay operation is complex, and in principle the nonlinear differential equations which describe it can be solved exactly only by computers. Even if this is done, the results are subject to any uncertainty that exists as to the exact form of the relations that apply. The representation of non-linear magnetic material properties, the discontinuous load-travel characteristics, the eddy current effects, etc., do not make such an approach attractive.

The relay, however, is a perfect analog of itself. With the magnetic structure set, controlled models can be built to include dimensional and magnetic material variations. With these, exact solutions to a variety of conditions can be determined. With these data available, approximate solutions can be used for interpolation and extrapolation, determining the effect of small variations from the tested conditions.

This part of the article will exhibit the form of data presentation in two classes, mass and load controlled operation. It includes the theory used in selection of the forms, and the correction methods used for estimation of variations from the standards. The initial part will be concerned with a single relay operating in a local circuit. The latter part will consider series and series-parallel operation of similar structures but not necessarily identical windings. This latter problem is solved by determination of an equivalent relay in a local circuit for each of the several relays. The earlier analysis then can be applied to each in turn.

The order of analysis can be reversed. That is, given required operating times, windings can be determined which will provide these times.

In this article, a best winding is (1) that winding which, for the specified applied power, results in the minimum operate time or (2) that winding which, for a specified operating time, requires the minimum power. A unique solution exists.

Existence of Best Winding

Fig. 1 shows measured operate time of a relay for two different dc power conditions, versus number of turns in the winding. That a best

winding, in the definition of this article exists, as exhibited by the minima on Fig. 1, was shown in Part I to be a consequence of equation (12). This best winding and its determination is the basic subject of this article.

OPERATE TIME — SINGLE RELAYS

As in Part I, the operating time of a relay is considered as made up of three stages: (1) the waiting time while the armature remains at the backstop and the pull builds up to equality with the back tension, (2) the motion time beginning at the end of the waiting time and continuing while the armature moves from the backstop to the position of the earliest contact, and (3) the stagger time during which the armature actuates all of the remaining contacts.

The mass controlled case, as a practical matter, simplifies to a determination of only the first two without regard for the spring load, with the displacement of the armature taken to the latest contact in the array. The armature pull builds up to values in excess of the load during its early motion and the velocity is so high during the stagger time that this small interval can be included as part of the motion time. This treatment is essentially that of the three stage approximation of Part I.

The load controlled case simplifies to a determination of the time required for the pull to build up to the maximum load, including the back tension, with most of the relatively slow motion taking place during this pull buildup. The remainder of the motion time is accounted for as an empirically determined correction factor in the expression for time of pull buildup. This treatment corresponds to the single stage approximation of Part I, with a correction term to account for the additional motion time.

Fig. 1 shows graphically the basic characteristics of the two types of operation. This chart displays the operating time of the same relay under the conditions of 1.0 or 5.7 watts, and three contact load conditions: 12, 18, and 24 contact pairs. For each of the six conditions, a curve is drawn showing the effect of the number of winding turns. It is clear that (1) there is a best number of turns for each case, (2) for the 5.7 watts the best number of turns is the same for all three loads, (3) for the 5.7 watts case the difference in operating time is only 0.1 millisec in 5.6 millisec, between 12 and 24 contact pairs, (4) for the 1.0 watt case the best number of turns increases with the number of contact pairs, and (5) for the 1.0 watt case the operate time is double for 24 compared to 12 contact pairs. These strikingly different behaviors form the basis for division of relays

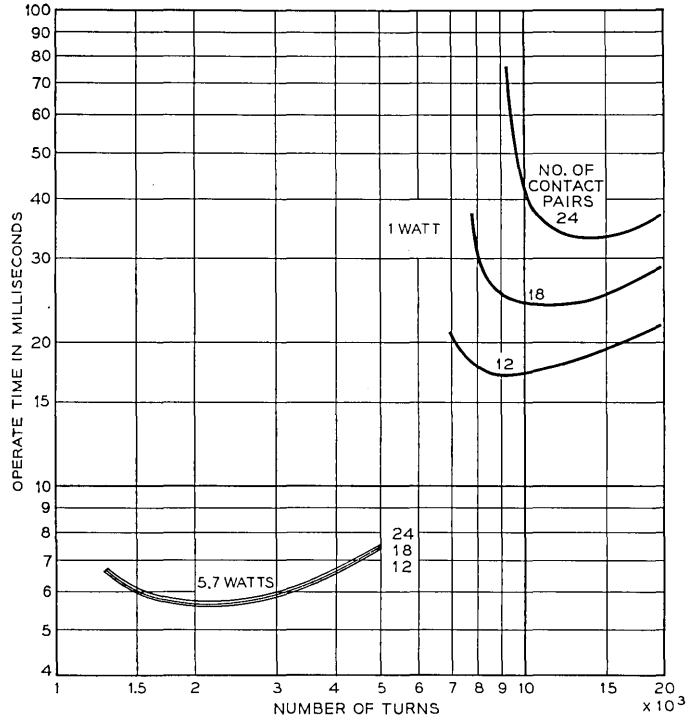


Fig. 1 — Chart showing typical mass and load controlled operation.

into two classes called mass and load controlled, and furthermore empirically establish which of the three stages of operate time predominates. This distinction corresponds to that made in Part I in the discussion of equation (12). For the lower power cases the term involving the spring load predominates, for the higher power cases the mass term predominates. As this equation shows, there is necessarily a transition range where the operation time merges from one class into the other. Both types of charts are extended to include this range. A rule of thumb for an estimate of whether a time under discussion is in the mass or load controlled case is that if the time is less than

$$t_M = \sqrt{\frac{27m(x_1 - x_3)}{F_3}} \tag{1}$$

it is mass controlled; otherwise it is load controlled. The derivation of this bound will be discussed when mass controlled operation is considered.

The single relay design data are general enough to include resistors

external to the relay winding. Wherever a resistor or power term appears, the sum of all resistances or powers is indicated.

Maximum Versus Average Operate Time Presentation

In what follows, load controlled operation data are given in terms of maximum times, whereas mass controlled data are average. Either choice could be used for both. The data for converting to average or maximum respectively will be described. However, the choices made here are consistent with the normal use of the data.

Mass controlled, sometimes called speed, relays operate with relatively large power and ordinarily are used in common control circuits where many events occur in succession. The total number of similar common equipments necessary in an office is related to the control circuit holding time. For a system design, the cost of power is balanced against the cost of equipments to arrive at a minimum office cost.[1] Now the maximum holding time per call is never the sum of the maximum possible times of each of the several relays operating in succession. It rather is more nearly the sum of the averages, increased by considerably less than the common maximum to average ratio of one of them. For example, if n relays are assumed to have a distribution approaching normal, an analytical expression for the probable maximum is:

$$\delta t_{\text{av}} + \left(1 + \frac{\frac{t_{\max}}{t_{\text{av}}} - 1}{\sqrt{n}}\right) \sum t_{\text{av}} , \tag{2}$$

where δt_{av} is an allowance for short time deviations of the manufactured product from the long time average. Thus for mass controlled relays the most directly applicable type of data is in the form of average time, and maximum to average time ratio.

For load controlled relays just the opposite is true. Here the operating times are relatively long, either because the power drain is to be kept to a minimum, as in a long holding time circuit, or it is an event which takes place while several successive mass controlled events occur. For either case, it is a single event and only its maximum duration is desired. Here then the most directly applicable type of data is maximum. In addition to these, other data for minimum times are needed for studies of timing when two parallel circuit paths occur.

Because how far the relay armature has to move is the outstanding variable in mass controlled relay timing, these charts are prepared for each nominal distance which can be chosen. When there is no concern as

to relative actuation time of the several contacts on one relay, the smallest armature travel is ordinarily provided. When a definite sequence is needed for circuit reasons, the actuating means is arranged to guarantee operation of certain contacts before the others. This necessitates the provision of a greater armature motion.

The principles used in preparing these other types of charts are identical to those used for the two types which are developed specifically in this article.

Local Circuit Load Controlled Operation

Power Given

The best winding for a load controlled relay is not here given explicitly by a formula, but rather is found indirectly by developing a method for determining the operate time for any winding. Then by assuming a range of these, the best one can be selected by interpolation of these derived data. In Part I, the waiting line for a linear system was shown as equation (9) to be:

$$t_1 = \frac{4\pi}{\mathfrak{R}_3}(G_C + G_E + G_S) \ln \frac{1}{1 - I_0/I}, \tag{3}$$

with I_0/I substituted for ν. Without motion time, saturation, eddy currents, or non-linear effects, Part I of this article also shows that the best winding is that for which the turns are selected so that $NI_0/NI = 0.715$.

Taking the other variables explicitly into account complicates the determination and is at best only approximate. Instead they are included through this interpolation approach. In a previous article[2] a better representation for the core eddy current constant was shown to be:

$$G'_E = G_E e^{-G_E/(G_C+G_S)} = \text{effective core eddy current constant.}$$

Substituting this and explicitly indicating a correction term for the small motion time, the form best suited for the present discussion becomes:

$$t_0 = (1 + t_2/t_1)L_1(G_C + G_S + G'_E) \ln \frac{1}{1 - q}. \tag{4}$$

Except for the correction term for the motion time, t_2/t_1, this is still a linear equation with all factors known. Omitting the motion time correction, it can be solved either by numerical substitution or by a nomogram, once a value for L_1 has been chosen. A conservative value is the average one turn inductance, for the late contact critical load point where $x = x_3$. A more accurate value is determined through use of both

the inductance at the open gap and at the critical load gap. These are weighted in proportion to the ampere turns developed at each location. The ampere turns necessary to overcome the back tension is one factor, and the other is the difference between this first value, and the total required for operation at the critical load gap. This weighting yields an effective inductance value intermediate between the two extremes for each problem, but a new chart does not have to be prepared. An operate time using the chart is first determined. This time is then adjusted by the ratio of the effective inductance applying and the inductance used for making up the chart. This is exact, as the inductance term appears only as a direct multiplier.

A typical nomogram is shown in Fig. 2. It already includes the motion time correction, whose determination will be discussed. The dashed line

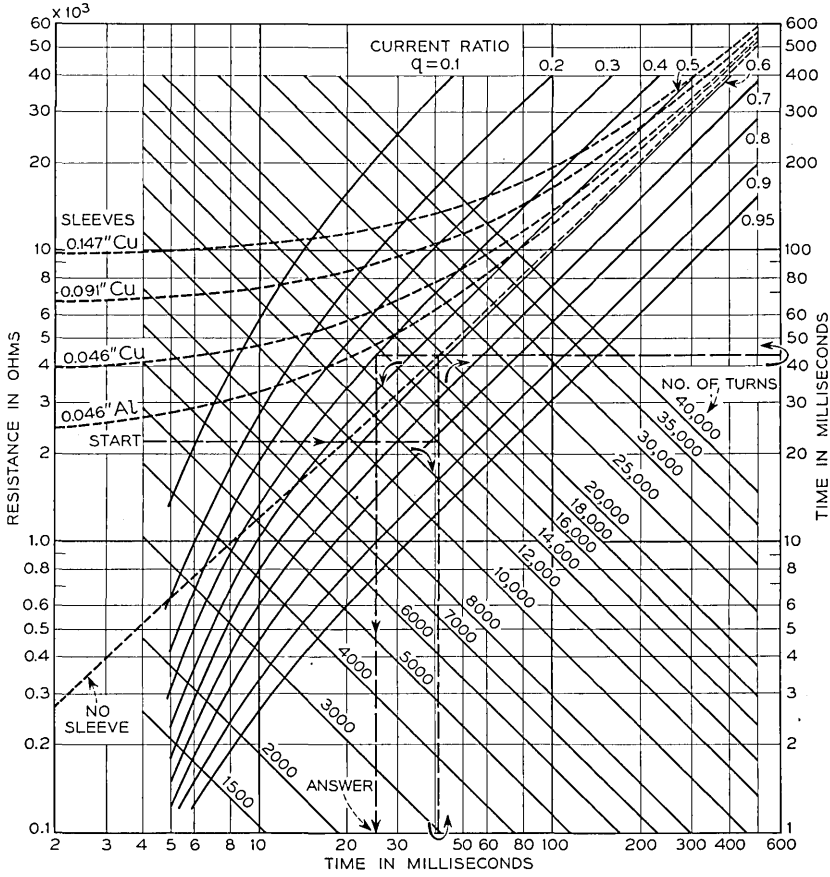


Fig. 2. — Nomogram for solution of load controlled operate time.

shows the successive steps taken in determining the numerical value for a specific case.

The dc circuit resistance is determined by the known circuit voltage and specified power. Entering the chart on the left ordinate at this resistance and proceeding horizontally to the right, intersections with increasing number of turns lines, sloping downward to the right, are found, and the appropriate one is chosen. Dropping vertically to the abscissa, the winding time constant $t_C = L_1 G_C$ is determined.

To this is next numerically added the known effective core eddy current and a sleeve (if any) time constant by returning vertically to an intersection with the appropriate core indicated as "no sleeve," or core plus sleeve, curve. Proceeding horizontally to the right hand ordinate scale from this intersection, the time constant multiplier of the ln term in equation (4) is determined.

The multiplication of these two factors is accomplished by proceeding to the left along this same horizontal line, to an intersection with the proper q line, sloping downward to the left. Vertically below this last intersection is the operate time.

Initially, tentative q lines are drawn, omitting the motion time correction. These lines are straight with a positive 45° slope. Then with an actual relay whose just operate current has been measured, operate time measurements are made, keeping the final current, and hence q, fixed at several values in turn. This is done by adjusting an external resistance and battery voltage over a wide range, effectively changing N^2/R. These measured data are plotted, following the same steps through the nomogram, except the last intersection is with the measured time vertical, rather than the known q. This provides several empirically determined q lines. These are used as templates, to progressively alter the shapes of the tentative straight q lines drawn earlier and shift them to the right. This adjustment then introduces the motion time correction.

It has been found empirically that for large operating times, the motion time correction factor has a value of about 0.1 There is no definite division between mass and load controlled operation, but as the total time decreases, travel time becomes more important. The correction factor increases to a value of about 0.5, in the transition range. For completely mass controlled operation, there is no q effect, so the q lines must all eventually converge.

As stipulated above, this chart applies to the current, flux and pull range where the first approximation magnetic constants of the structure are applicable. For this reason, the tests for the motion time corrections are made under conditions meeting these restrictions.

These curves should be corrected for magnetic saturation if there is a

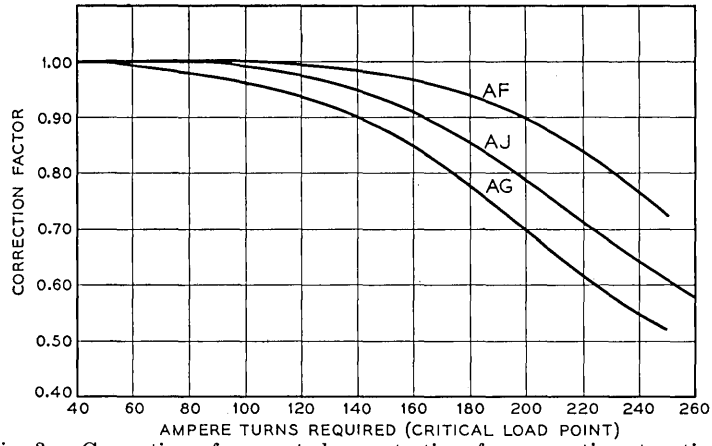


Fig. 3 — Correction of computed operate time for magnetic saturation.

large load at the critical load point and the ampere turns needed are in the saturation region for the core or armature. In the saturation region the current builds up faster than the linear time constants indicate, and, therefore, the indicated operate times are too large. Correction factors are determined by graphical integration of the integral form of the flux rise equation expressed as a ratio to the linear relationship:

$$\text{Saturation Correction Factor} = \frac{\int_0^{\varphi_3} \frac{d\varphi}{NI - Ni}}{-L_1 \ln (1 - NI_0/NI)}. \tag{5}$$

These corrections are plotted as a function of the just operate ampere turns NI_0, with the final NI as a parameter for each curve. Actually, because the correction factor is only of the order of 20 per cent maximum, assuming the final winding ampere turns are well into the saturation region, it is found that a single curve for any one type of relay fits all the computed points to an accuracy of a few per cent, and generally is used. Such composite curves are shown in Fig. 3 for the three types of wire spring relays.

A method has now been established for determination of the operate time of a relay with two restrictions (1) that the final ampere turns will operate the relay and (2) that the relay is in the load controlled class. An indication of the latter is whether the operate time determined is in the time region where the q curves are decreasing in curvature. For Fig. 2, 10 milliseconds is taken as the lower bound for load controlled relays.

Now the determination of an optimum winding for a particular

problem can be finished. Starting with the power given and choosing an arbitrary number of winding turns, the operate time is determined as described above. If it falls into the load controlled class, then a different number of turns is next assumed and a second time determined. This procedure is repeated until a time curve versus number of turns similar to Fig. 1 can be plotted, including a minimum. The best winding i this minimum. A different curve and optimum number of turns will apply to each contact load assumed. Three computed curves corresponding to the measured 1-watt curves are shown on Fig. 4. These were determined using effective inductance values described earlier.

Finite wire sizes permit only certain number of turns to be physically realizable when the resistance has been specified, without splicing two gauges of wire. The nearest gauge on the coarser wire size side is chosen, resulting in slightly too many turns. Note that the curves rise less steeply on the high turn side and the time penalty therefore is less than if too few turns were supplied.

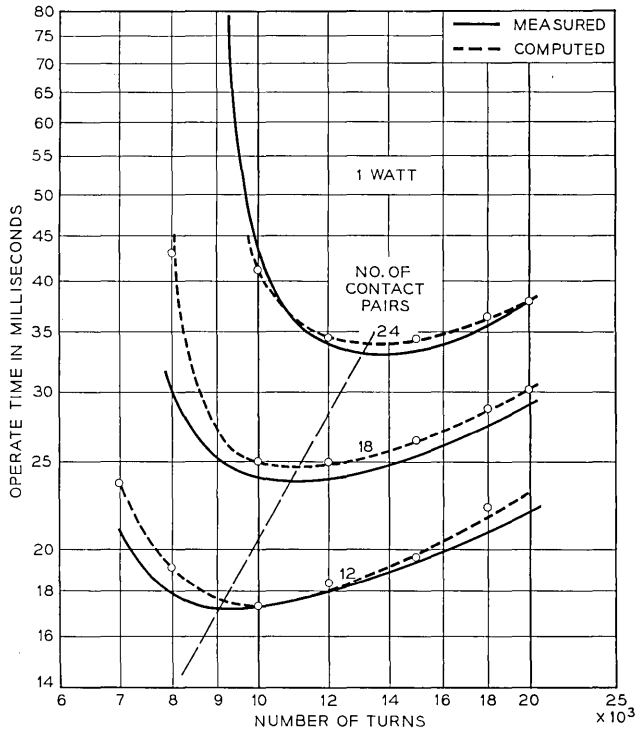


Fig. 4 — Typical computed maximum operate time curves for load controlled relays.

Maximum Time Given

For a specified maximum time, the above process is repeated for several assumed circuit powers until the specified time is bracketed. Then by interpolation of the optimum times indicated, the minimum power, maximum resistance and optimum turns are determined. As an example, Fig. 4 can be used to demonstrate the method. Assume that the three curves were computed for different circuit powers, rather than for different contact spring loads. A line is drawn through the minima. The intersection of this line with the required operate time determines the number of turns. For instance, if 30 millisec were required, the turns would be 12,500. The circuit power at this same intersection can be interpolated for, using the known circuit powers associated with the three curves.

It is not economical to have a different winding for each spring combination. For this reason, a winding is designed for the maximum spring load and then used for smaller loads. The operate time will always be less with the smaller loads.

Measurements of Time Curves

Before considering mass controlled operation, the simulation of windings will be discussed. In the above description for establishing the q curves, it was pointed out that by adjusting an external series resistor and the battery voltage to maintain constant final current, the coil constant N^2/R was altered without changing q. This can be further extended to permit simulation of any winding for test purposes providing only that the experimental coil fills the winding volume as much or more than the coil to be simulated. For this purpose, a special test winding

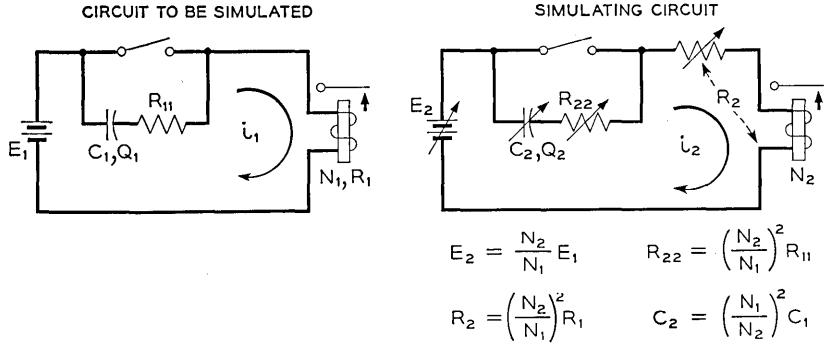


Fig. 5 — Simulation of winding circuit for timing tests.

always is used which completely fills the winding volume. Hence any winding which can be designed also can be simulated.

The conditions which must be fulfilled for perfect simulation derive from Lenz's Law. It is essentially an impedance transformation technique keeping the magnetic flux invariant, with the assumption that a winding can be considered as a lumped rather than a distributed network. This is equivalent to stating that at any instant the current flow is the same in every turn and there is no propagation time involved. This is true for times involved in electro-magnets.

Fig. 5 shows a circuit to be simulated, in which all the components with subscripts 1 have been given. The simulating circuit has only the number of turns N_2, of the test winding, given. The other four elements must be determined. After switch closure, the exact differential equations applying are

$$\begin{aligned} N_1 \frac{d\varphi_1}{dt} &= E_1 - i_1 R_1 , \\ & \qquad t = 0; \qquad i_1 = i_2 = 0. \\ N_2 \frac{d\varphi_2}{dt} &= E_2 - i_2 R_2 . \end{aligned} \tag{6}$$

Now for equality of magnetic flux, the two rates of flux change must be identical at all times including the first instant. Inserting the initial boundary conditions and equating the two rates, we have

$$\frac{E_1}{N_1} = \frac{E_2}{N_2}. \tag{7}$$

At infinite time, the same magnetomotive force must apply to both circuits for equality of final flux. Equating these, and cancelling the 4π factor,

$$N_1 I_1 = N_2 I_2 . \tag{8}$$

Noting that

$$I_1 = \frac{E_1}{R_1}, \tag{9}$$

and

$$I_2 = \frac{E_2}{R_2},$$

we have, after using (7) and rearranging,

$$\frac{N_1^2}{R_1} = \frac{N_2^2}{R_2}, \tag{10}$$

which states that the coil constants must be equal. After steady state has been reached and the switches opened, the two differential equations are:

$$E_1 = i_1(R_{11} + R_1) + N_1 \frac{d\varphi_1}{dt} + \frac{1}{C_1}\int_0^t i_1\, dt,$$

$$E_2 = i_2(R_{22} + R_2) + N_2 \frac{d\varphi_2}{dt} + \frac{1}{C_2}\int_0^t i_2\, dt, \tag{11}$$

$$\text{at} \qquad t = 0, \qquad i_1 = \frac{E_1}{R_1}, \qquad i_2 = \frac{E_2}{R_2}, \qquad Q_1 = Q_2 = 0.$$

Multiplying the first equation by N_1C_1 and the second by N_2C_2, and equating term by term for equality at all times and equal magnetomotive forces, two additional equations result:

$$\begin{aligned} (R_{11} + R_1)\, C_1 &= (R_{22} + R_2)\, C_2\,, \\ N_1^2 C_1 &= N_2^2 C_2\,. \end{aligned} \tag{12}$$

The four defining equations, with some further rearrangements using (10), are shown in Fig. 5 as the relations applying for equality of magnetic flux in the two circuits. As the mechanical behavior of a magnetic structure is completely determined by the magnetic flux, it follows that the mechanical performances will be identical and timing measurements made with the simulating circuit will represent the actual circuit. This is true whether the relay is mass or load controlled. It includes eddy current effects and whether or not there is motion of the armature.

Local Circuit Mass Controlled Operation

Typical mass controlled operate time curves are shown in Fig. 6. These are for two different armature travels, indicated as short and intermediate. The curves are characterized by the circuit power used for each, with the coil constant plotted along the abscissa and the corresponding operate time as the ordinate. As mentioned earlier, these curves are substantially independent of contact spring load, and are plotted for average conditions, including an averaging of the time for the first and the last contact to be actuated.

It will be noted that the best coil constant is not independent of the circuit power, decreasing continuously as the power is increased. Also by increasing the circuit power from 2.3 to 23 watts, the operate time is decreased by a factor of a little less than 3, which is nearer the square

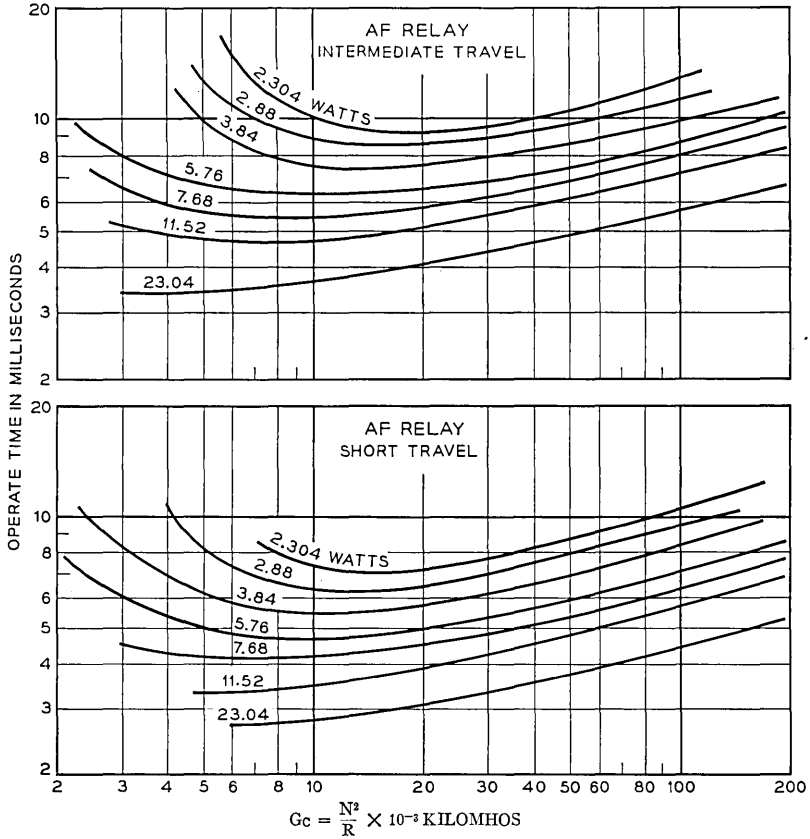


Fig. 6 — Mass controlled operate time tests.

root than the cube root of the power ratio, developed in Part I, for the mass controlled case. This is partly attributable to the fact that waiting time is included, and partly to the fact that the 2.3 watts case is in the transition range between mass and load controlled operation.

For the operating region where coil constants are larger than the best, the time curves are parallel and increasing. Considering any one vertical line representing a particular winding, increasing the power by increasing the battery voltage will always decrease the operate time. For curves plotted as in Fig. 1, where the battery voltage is kept constant and the time curves are plotted with winding turns as the independent variable, parallel curves again obtain. Thus an increase in circuit power obtained by keeping the voltage constant and reducing the circuit resistance always will reduce the operate time. Conversely, adding any series

impedance, not including a capacitance, will always increase the operate time. This will be made more evident when the operate times of relays with their windings in series are considered.

The curves can be used in either of two ways (1) given the relay circuit for which the applied power and coil constant can be computed, the average operate time can be found from the chart, or (2) given a required average operate time, the necessary circuit power and coil constant can be found from the chart.

These curves can be plotted in this form to exhibit the best winding directly because of the independence to contact spring load. For the load controlled case, the contact spring load was an essential parameter. For the mass controlled case, the armature travel becomes the outstanding parameter to be considered, but there are only two or three of these. All the other factors except contact load also enter and need to be evaluated for two reasons, (1) to provide an estimate of the range in operate time to be expected and (2) to adjust experimental measured time data to average. For any experimental setup, it is seldom possible to provide a structure which is average in every respect. For any one structure, however, all the factors known to affect its performance can be measured. Comparison of these to the manufacturing specifications locates the experimental setup in the universe of all relays as regards each of the factors. It is thus necessary to develop representations relating each of these factors to the operate time, which will suffice for the two uses named above. The development of these relationships will be the subject of the following sections.

Waiting Time

The waiting time, whether an electromagnet is mass or load controlled, is given by the same form of equation used for the total operate time of a load controlled relay:

$$t_1 = L_1(G_C + G'_E) \ln \frac{1}{1 - q_1}. \tag{13}$$

where now q_1 is determined by the armature back tension F_1 ; L_1 applies to the open gap; and no sleeve conductance is present. For the present purpose, it is desirable to rewrite this equation in terms of the fundamental parameters of the relay. For the open gap case the magnetic material is operated in its linear region and the open pole face gap provides additional linearity. For these reasons the expression is quite accurate. The sketches of Fig. 7 show the factors to be used. The value of

L_1 ,[4] the inductance for one winding turn is:

$$L_1 = 4\pi\left(\frac{1}{\mathcal{R}_L} + \frac{1}{\mathcal{R}_0 + x_1/A}\right). \tag{14}$$

Also

$$NI = \sqrt{G_C W}. \tag{15}$$

From the network of Fig. 7,

$$\varphi_G = \frac{4\pi Ni}{\mathcal{R}_0 + x_1/A}. \tag{16}$$

The armature force developed is

$$-F = \frac{\varphi_G^2}{8\pi A}. \tag{17}$$

When the magnetic pull has reached F_1 , the waiting time is over, determining φ_G , which in turn determines NI_0 , and hence q.

Substituting (14), (15), (16), and (17) in (13) we have

$$t_1 = 4\pi\left[\frac{1}{\mathcal{R}_L} + \frac{1}{\mathcal{R}_0 + x_1/A}\right](G_C + G'_E) \ln \frac{1}{1 - (A\mathcal{R}_0 + x_1)\sqrt{\dfrac{F_1}{2\pi A G_C W}}} \tag{18}$$

as the desired expression for the waiting time in terms of the fundamental constants. Use will be made of this later when an expression for motion time has been developed.

Motion Time

The motion time immediately follows the waiting time and is the time required for the armature to move from the backstop position x_1 , to the location of the last contact x_3 . From Fig. 1, it is permissible to omit any consideration of contact spring load, and consider the motion to be controlled entirely by the armature mass and magnetic pull developed after the waiting time is over. The determination of motion time is simplified because the initial velocity and net force on the armature are both zero, the latter following from the definition of the end of t_1 . In any problem involving motion, there must be established the initial velocity, position, and a suitable expression for the ensuing force. With

these factors, the differential equation of motion can be solved, and the time for a given travel determined.

Thus there is now needed an expression for the armature force developed by the pole face gap flux as a function of time. Briefly, the basis for the method to be described is the assumption developed in Part I, that the winding flux continues to rise during the initial motion interval in the same way it would have risen if the armature had not moved. With this assumption, an expression for the motion can easily be derived. With this expression it then can be shown that the armature does spend most of its time in the close vicinity of the backstop, justifying the initial assumption.

The present approach differs from the more general one of Part I. It uses the results there developed regarding the behavior of mass controlled operation, to simplify the initial motion equation and obtain an expression for motion time in a form better suited for the present purpose.

Armature Force Rise and the $\dot{F}$ Concept

As an introduction to the method which will be used, consider the general character of flux build-up in a relay. It will have a shape somewhat similar to an exponential curve. Now with the armature at the backstop, from the first approximation magnetic network of Fig. 7, a fixed portion of this winding flux will pass through the pole face gap, with the result that the pole face gap flux curve will also have the same general character. Such a measured curve is shown in Fig. 8, as well as the curve when the armature is free to move.

Now the armature force developed is proportional to the square of

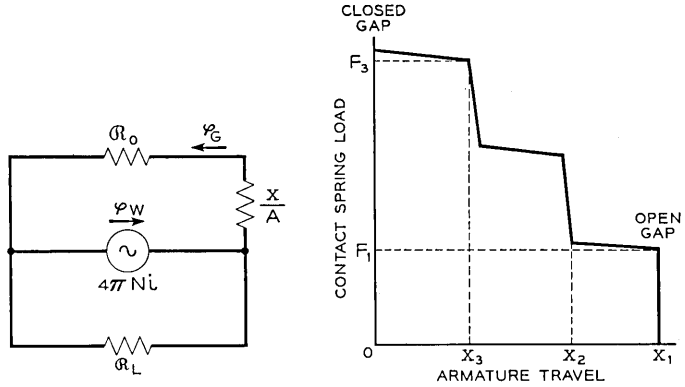


Fig. 7 — Schematics of nomenclature applying to operate time.

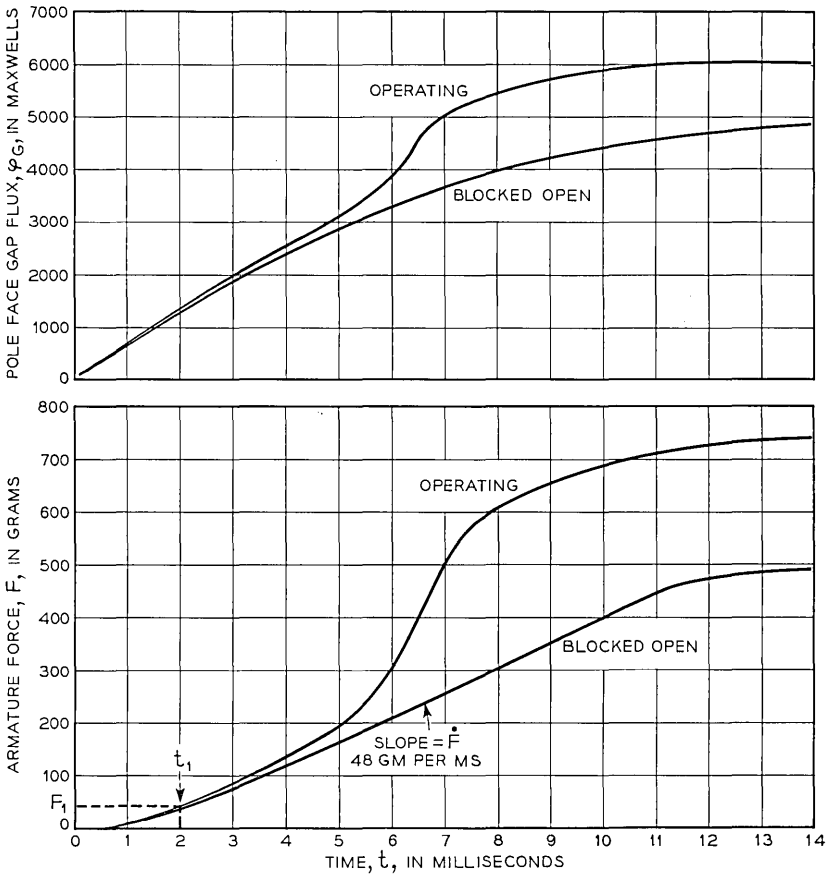


Fig. 8 — General character of winding flux and armature force rise.

the pole face gap flux as given by (17). Hence, the initial force rise will be parabolic, as shown by the lower portion of the curves of Fig. 8. After a long enough time however, the flux will have reached its maximum and the finally developed force will be constant with time. This is shown by the force curves approaching horizontal lines to the right in the same force diagram. As the force curve is continuous, there also must be an inflection point in the curve, and the entire curve has the general shape as shown.

The waiting time t_1 occurs while the magnetic force builds up to F_1, as indicated, amounting to 2 millisec for the example shown. This time generally includes all of the parabolic part of the curve. Following the waiting time, the initial force build-up necessarily is almost linear be-

cause of the inflection of the curve. Two other factors, caused by the ensuing motion, act to provide even more force than the open gap curve indicated by Fig. 8. The first is that a smaller gap takes a larger portion of the winding flux than assumed in the diagram, and the second is that the winding flux rises more rapidly and to a higher value than the open gap curve assumed. The operating force curve shown includes these effects. The relay operates in about 7 millisec and for 6 of the 7 millisec, the two force curves differ very little.

Hence, for a relatively long interval after the start of the motion time, a very simple force relationship holds, namely that the force is directly proportional to time, and the proportionality factor is the slope of the straight portion of the curve, designated as $\dot{F}$. For the example given, this amounts to about 48 grams per millisecond. By using Newton's Law, an expression for the motion time can be written.

$$m \frac{d^2(x_1 - x)}{dt^2} = \dot{F}t, \tag{19}$$

from which, with the initial velocity zero

$$m(x_1 - x) = \frac{\dot{F}t^3}{6}. \tag{20}$$

Rearranging, the expression for motion time becomes:

$$t_2 = \sqrt[3]{\frac{6m(x_1 - x_3)}{\dot{F}}}. \tag{21}$$

The factor $\dot{F}$ will now be considered.

Derivation of Expression for $\dot{F}$

Experimental—The factor $\dot{F}$ can be determined graphically. This brings in the second order effects such as quality of iron, cross section, residual magnetism, fit of parts, etc. For this purpose three working curves characteristic of the structure, all at the open gap, are first prepared for the particular winding:

(a) The dynamic flux rise curve,[2]
(b) The static flux curve versus ampere turns, and
(c) The static pull curve versus ampere turns.

Choosing a time on curve (a), the corresponding instantaneous flux transferred to curve (b), determines the equivalent magnetomotive force. This transferred to curve (c) yields the instantaneous magnetic force. Repeating this for other times in the range of interest, an armature force

curve like Fig. 8 is established. The waiting time is read directly, corresponding to F_1. The slope of the ensuing linear force range determines $\dot{F}_1$. This, with equation (21) completes the determination of mass controlled motion time. Of course, the final pull developed has to exceed the operated load. This check is made from another pull curve taken at the armature gap x_3, using the known final ampere turns and the load.

Analytical—For our present purposes, an analytical relation, expressed in the fundamental constants, similar to that for the waiting time is needed. Its derivation follows:

The solution which will be developed is based on linear circuit theory. This necessarily implies exponential flux rise, which is not exactly true. However, the relation is dimensionally correct and accurate to better than first order. Then the use which will be made is to determine the motion time of an electromagnet as the parameters are varied one at a time. These are plotted as ratios to one of them, chosen as a reference. By this means the ratio curves become accurate to better than second order and provide excellent correction factors for actual measured data.

For a linear circuit, the pole face gap flux will increase, after the winding circuit is closed to a battery, with the same time constant as the winding:

$$\varphi_G = \frac{4\pi NI}{\Re_0 + \frac{x}{A}} (1 - e^{-t/T}), \tag{22}$$

where $T = L_1 (G_C + G'_E)$ and $I = E/R$. Then the pull:

$$-F = \frac{\varphi_G^2}{8\pi A} = \frac{1}{8\pi A} \frac{(4\pi NI)^2}{\left(\Re_0 + \frac{x}{A}\right)^2} (1 - e^{-t/T})^2, \tag{23}$$

$$\frac{dF}{dt} = \frac{4\pi (NI)^2}{AT\left(\Re_0 + \frac{x}{A}\right)^2} [e^{-t/T}(1 - e^{-t/T})]. \tag{24}$$

The bracket term is of the form $a(1 - a)$, $0 < a < 1$, which has a maximum value of 0.25, and from $0.2 < t/T < 1.5$ is between 0.15 and 0.25. This range corresponds to the maximum slope of the armature force versus time plot in Fig. 8. Arbitrarily 0.2 is chosen and the expression for the maximum rate of force rise becomes:

$$\dot{F} = \frac{4\pi (NI)^2}{5AT\left(\Re_0 + \frac{x}{A}\right)^2}. \tag{25}$$

Substituting for T and rearranging:

$$\dot{F} = \frac{(NI)^2}{5(G_C + G'_E)(A\mathcal{R}_0 + x_1)\left(1 + \frac{\mathcal{R}_0}{\mathcal{R}_L} + \frac{x}{A\mathcal{R}_L}\right)}. \tag{26}$$

Substituting (26) and (15) into (21), the final expression for motion time becomes:

$$t_2 = \sqrt[3]{\frac{30m}{W}(x_1 - x_3)\left(1 + \frac{G'_E}{G_C}\right)(A\mathcal{R}_0 + x_1)\left(1 + \frac{\mathcal{R}_0}{\mathcal{R}_L} + \frac{x_1}{A\mathcal{R}_L}\right)}. \tag{27}$$

This, with equation (18) for the waiting time, forms the basis for the variation effects which now will be determined. Its region of accuracy is for windings with turns exceeding the best, as it presupposes the relay will always operate. It thus is applicable in the region where the time curves are parallel. For studies of lower numbers of turns, the three stage approximation of Part I does not have this limitation.

Operate Time Variations

For a particular electromagnetic structure the factors in equations (18) and (27) can be measured. These measured factors can then be compared to the manufacturing specification and estimates made of average values for each. With these average values, the waiting time, the motion time and then the sum, which is the operate time, can be computed. This is the reference condition.

Each factor is then varied over its appropriate range, one at a time, and the computation repeated. The ratio of these variation times to the reference condition can then be plotted. Fig. 9 shows the chart for the wire spring type relay. For this chart, the range for each factor was taken as 3 without regard to the actual range. Similar charts have been prepared for other types of electromagnets, including quite a size difference. These ratio charts all agree remarkably well when plotted in this way. For this reason, for early estimates of any structure in the mass controlled class, this chart is entirely adequate for estimates of variations.

The basic assumption made in this method of determination is that for small variations, the interactions are negligible and a separated solution of products, one for each factor, is applicable. Checks made by varying two at a time confirm that essentially the presumption is fulfilled.

The G_C curve does not exhibit a minimum, as do the measured curves of Fig. 6. This results from the simplifying stipulation made that the

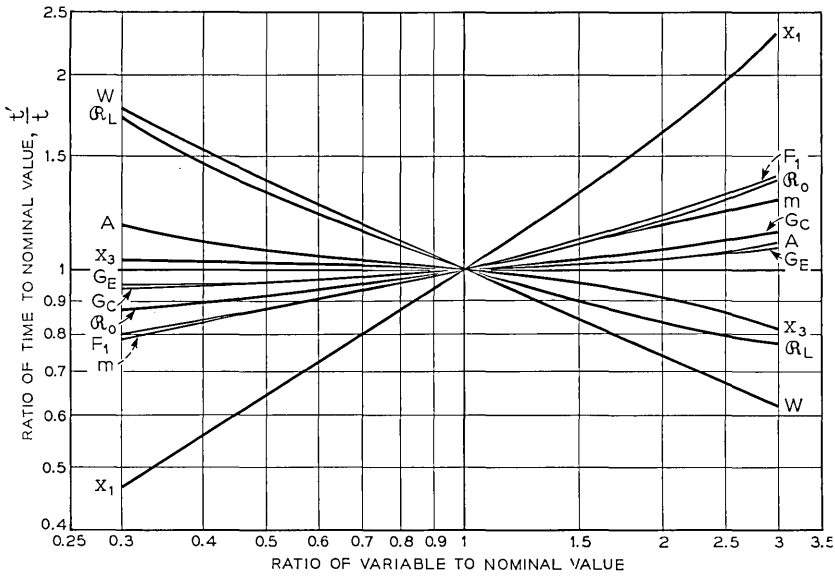


Fig. 9 — Mass controlled operate time variations.

relay always must operate. However, the purpose of these variation curves is to adjust measured data. The coil constant is the independent variable in any measurement and can always be set exactly. Hence it does not require an adjustment.

The chart shows clearly that the single most important factor is the armature travel. Following are power W, and leakage reluctance $\mathcal{R}_L$. The least sensitive is the pole face area A, because it has been optimized in the design. Taking the slopes at the (1, 1) point, the operate time, expressed in a separated variable form, becomes:

$$t_0 \approx \frac{x_1^{0.66} F_1^{0.24} \mathcal{R}_0^{0.2} m^{0.21} G_C^{.084} G_E^{.056} A^0}{W^{0.45} \mathcal{R}_L^{0.33} x_3^{.084}} . \tag{28}$$

The exponents are a measure of sensitivity — the nearer they approach zero, the less the sensitivity. In designing electromagnets for speed, every effort should be made to keep the armature travel to a minimum as it is outstanding in its effect on time.

For our present purpose, the chart is used to adjust measured operate time data to the average value. The factors to be corrected are F_1, m, x_1, x_3, A, $\mathcal{R}_0$, $\mathcal{R}_L$, i.e., all the pertinent geometric and load factors shown in Fig. 7. This completes the description of the method used for establishing the average mass controlled operate time curves.

Maximum and Minimum Values

The next to last part of this phase of the article concerns estimates of the range of the mass controlled maximum operate time of a particular relay code. This involves determining the range of each of the variables from the manufacturing specifications. Then with the variation chart the range in the time, due to each component, is evaluated, as a ratio. The actual range of course is not the product of these, as all variables never will be adverse simultaneously, but it is greater than the largest single individual contribution. For general purposes, a root sum square addition is used. For the wire spring relay this amounts to a range of about ±30 per cent.

Preliminary Estimate of Type of Operation

Earlier, a rule of thumb expression for an estimate of whether an electromagnet is in the mass or load controlled class was given. It is easily derived by use of the $\dot{F}$ concept, and the observation from computing the variation charts, that the waiting time generally is about one-half of the motion time when the electromagnet is mass controlled.

When the motion time begins, the force starts to build up according to equation (19). However, multiplying this initial rate by the motion time need not result in a force value equal to the load for operation to be complete. Four factors act to this end: (1) during the waiting time the back tension F_1 already has been overcome; (2) when the armature gap closes, the armature takes a greater portion of the winding flux; (3) as the armature moves in the winding flux builds up more rapidly; and (4) the kinetic energy can cause the armature to crash through a small distance where the load projects through the rising dynamic pull curve.

Calling t'_M the longest average motion time which is mass controlled, and arbitrarily setting

$$\dot{F}_1 t'_M = \frac{F_3}{2}$$

in view of the foregoing, an estimation of the limiting average $\dot{F}_1$ is determined. Substituting this in (20), the maximum average mass controlled motion time is:

$$t'_M = \sqrt{\frac{12m(x_3 - x_1)}{F_3}},$$

which, when multiplied by ⅔ to allow for waiting time, and dropping

the prime to indicate total time, becomes

$$t_M = \sqrt{\frac{27m(x_3 - x_1)}{F_3}} . \tag{29}$$

This equation was given as equation (1) in the early part of this article.

For the wire spring relay, this expression has a value of 7.5 millisec. Increasing this by 30 per cent for a maximum value, an estimate of 10 millisec results, agreeing with the earlier lower estimate of maximum time for load controlled operation.

Selection of Winding for Mass Controlled Operation

One more group of factors needs consideration before a winding is selected. These are (1) the range in dc resistance of the windings, (2) the winding temperature as determined by the duty cycle, and (3) the range in the battery voltage. The number of turns of the winding is ordinarily not considered as a variable once it is chosen because of the automatic machine method of winding. An examination of Fig. 6 shows that if the turns are too few, a greater time penalty obtains than if there are too many. Also, decreasing the circuit power, increases the best coil constant. These two considerations indicate that the best coil constant should be chosen under worst circuit conditions. For any other condition the operate time will be reduced. Further, the range between worst circuit and average time will be a minimum.

The procedure for choosing a winding is to determine the dc resistance of a maximum resistance winding at the operating temperature set by minimum battery voltage and the maximum duty cycle. This sets the worst circuit power, and by use of Fig. 6, the best number of turns. In no case is a winding specified with fewer turns than will supply sufficient ampere turns to operate the worst relay with the maximum load. In some cases of low power, this sets the number of turns. For some cases of intermediate power, heating requires the maximum winding surface area, also resulting in excess turns. The average resistance with its variation, all at a standardized temperature at 68°F, completes the design.

Summary of Single Relay Local Circuit Operation

An analytical determination of the operate time of a single relay cannot be obtained in closed form because it requires the solution of two simultaneous, non-linear, non-homogenous, differential equations, without adding the complications of representing the magnetic saturation

and eddy current effects in some convenient form. Approximations for solutions have been developed which predict with good accuracy the order of operate time attainable for a design, but not of sufficient accuracy to exactly determine the best winding for a particular case. Thus once a magnetic structure has been established, the operate time data for a single relay necessarily have to be determined empirically by measurements using controlled samples.

Actually this procedure is quicker and easier, and besides it gives the correct answer, including all the non-linear effects, such as eddy currents, saturation and motion, no matter how complicated. Complete data for the range of all relays and windings using the given structure can be determined using one such relay with a full winding, through the use of impedance transformation techniques and estimation of small variation effects, using the approximate solutions. By using the approximate solutions only for corrections, the errors become of second or smaller order. Thus single relay operate time data can be determined accurately and presented in a form permitting either analysis or synthesis of performance.

The design of a best winding for load controlled operation, is accomplished by a variation of the method of successive approximations. Appropriate battery voltage, winding temperature and resistance of course are considered as part of the solution of the problem. A range of windings is chosen and operate time data established for each winding. If necessary, the range is extended in the appropriate direction until a minimum operate time is included. This minimum determines the best winding.

For mass controlled operation, the contact spring load is immaterial, and the data can be presented directly. As part of this type of study, the relative importance of the several parameters affecting the operate time has been determined. These show at a glance whether (1) a change will have a significant effect or (2) what change or changes are necessary to effect a necessary reduction in operate time.

OPERATE TIME — SERIES RELAYS

Series relays are two or more relays whose windings all are connected in series, and energized by the same current, controlled by a single contact. The impedance of each one enters into the manner in which the common current will increase, after contact closure. If the procedure used for single relays were followed, there would be a double infinity of combinations to portray, or else experimentally study each combination

when proposed. This can be avoided, with no loss in generality, by transforming each relay into an equivalent single relay in a local circuit. Then the foregoing methods for single relays can be applied to each in turn.

This procedure is the common device of breaking up a complicated problem into parts, each of which can be solved by familiar methods. Two assumptions are made. The first is that each winding is a lumped two terminal network. At any instant the same current is in every turn of each relay and there is no propagation time involved. This holds for the times involved in electromagnets. The second assumption is that, when the relays have different operate times, the current reduction caused by the motional impedance of the first one to operate does not significantly extend the operate times of the later ones. In the following transformations, the winding turns, currents, and hence magnetomotive forces, are kept constant.

Identical Relays

If the two relays are identical they have some impedance $Z(p)$ which is the same for both. Part of this is the dc resistance and the other part is the ac effect, proportional to the turns squared. By the extension of Ohm's Law to ac circuits, when a potential source is applied to the two identical devices in series, exactly one-half the source appears across either device at any time after application, forming a virtual constant potential point. Thus if a battery of E_0 volts is applied, exactly one-half the battery appears across each, including the effect of eddy currents. Now if the voltage across a coil is known, then the response is uniquely determined, knowing just the relay characteristics and the voltage, disregarding the mechanism of how the voltage is applied. For this situation the voltage is in a most convenient form, represented by exactly one-half the battery. The operate time can easily be determined for either relay with this information, as the effects of eddy currents and motion are included in the data. The coil constant is already known and the power is one-half the total power.

General Case

This procedure can be generalized to include any division of dc resistance, different turns, and different magnetic structures providing, for the latter case, that eddy currents can be ignored. The justification for this will be considered later.

The basic problem is: given the battery voltage E_0, relay No. 1 of N_1 turns, a 1 turn inductance L_1, and resistance R_1, in series with relay No.

TABLE I — EQUATION 30

Relay	Turns N	Resistance R_E	Voltage E_E	Power W_E	Coil Constant
No. 1	N_1	$\dfrac{R_1 + R_2}{1 + \frac{L_2}{L_1}\left(\frac{N_2}{N_1}\right)^2}$	$\dfrac{E_0}{1 + \frac{L_2}{L_1}\left(\frac{N_2}{N_1}\right)^2}$	$\dfrac{E_0^2}{(R_1 + R_2)\left[1 + \frac{L_2}{L_1}\left(\frac{N_2}{N_1}\right)^2\right]}$	$\dfrac{N_1^2}{R_1 + R_2}\left[1 + \frac{L_2}{L_1}\left(\frac{N_2}{N_1}\right)^2\right]$
No. 2	N_2	$\dfrac{R_1 + R_2}{1 + \frac{L_1}{L_2}\left(\frac{N_1}{N_2}\right)^2}$	$\dfrac{E_0}{1 + \frac{L_1}{L_2}\left(\frac{N_1}{N_2}\right)^2}$	$\dfrac{E_0^2}{(R_1 + R_2)\left[1 + \frac{L_1}{L_2}\left(\frac{N_1}{N_2}\right)^2\right]}$	$\dfrac{N_2^2}{R_1 + R_2}\left[1 + \frac{L_1}{L_2}\left(\frac{N_1}{N_2}\right)^2\right]$

2 of N_2 turns, a 1 turn inductance L_2 and a resistance R_2. What are the two equivalent relays, each in a local circuit and what are their virtual applied voltages? The first step is to reassign the total resistance to obtain two equivalent windings having the same time constant N^2L_1/R. Then determine the two virtual voltages by division of the total in proportion to the impedances. These and their dependent power and coil constant relationships for the case of two relays are tabulated in Table I. The procedure can be extended to any number of dissimilar series structures.

For the case of all identical magnetic structures, the 1 turn inductances L_1, L_2, etc., are all equal and their ratios become unity, simplifying the expressions. Note that the coil constants are not equal unless the structures are the same magnetically, but that the time constants always are equal. Further, the total power is constant and equal to that of the original circuit.

After determining the effective powers and coil constants, the operate time for each can be read from the applicable single relay charts such as Fig. 6.

Two Like Parallel Relays in Series with a Third Relay, Identical Structures

The equivalent relay method can be extended to include the case of two like parallel relays in series with a third relay, as shown in Fig. 10. The first observation to make is that, because of symmetry, the current flow in the two parallel relays is identical. No winding current change would be made if, for instance, all the dc resistance of the two parallel relays were removed and half of either were connected in series with the single relay. Thus again, the resistances can be assigned as necessary to result in equal winding constants and total power. The same net dc resistance gives a second condition:

$$\begin{aligned} \frac{N_1^2}{R_{1E}} &= \frac{N_2^2}{R_{2E}}, \\ R_1 + \frac{R_2}{2} &= R_{1E} + \frac{R_{2E}}{2}. \end{aligned} \tag{31}$$

Solving

$$R_{1E} = \frac{2R_1 + R_2}{2 + \left(\frac{N_2}{N_1}\right)^2},$$
$$R_{2E} = 2\left(R_1 + \frac{R_2}{2} - R_{1E}\right). \tag{32}$$

Because the equivalent coil constants have the same ratio as the equivalent resistances, the voltage division will be:

$$E_{1E} = \frac{R_{1E}E_0}{R_1 + \frac{R_2}{2}},$$
$$E_{2E} = E_0 - E_{1E}. \tag{33}$$

With these, the equivalent coil constants and powers can be computed and charts such as Fig. 6 used as before.

Selection of Optimum Coils, Identical Structures

Neither Winding Known, Operate Times Given

In the above discussions it was assumed that the windings were known and that the operate time data were sought. The procedure can be reversed, starting with a desired operate time, or times, and then choosing optimum coils. From time curves such as Fig. 6, the watts corresponding to the desired times can be obtained provided they are read from the same coil constant vertical because the same current flows through both and hence the effective coil constants necessarily must be equal.

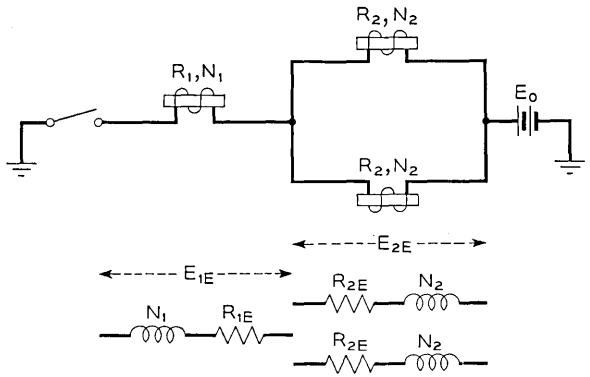


Fig. 10 — Transformation of series-parallel relay circuit.

The particular vertical N^2/R to be used is the one passing through the higher of the two given times, which lies at the minimum point on a watts curve. Satisfying the smaller wattage relay assures sufficient ampere-turns for both windings, and also gives a smaller time loss compared to optimum for the other higher speed relay, as the higher power curves are flatter for coil constants above the optimum. Thus the specified times determine equivalent watts for each coil, and the actual total circuit watts are exactly the sum of the equivalent watts. Hence:

$$W_{\text{total}} = \frac{E_0^2}{R_1 + R_2}. \tag{34}$$

This sets the total resistance $(R_1 + R_2)$. Also, knowing the current is the same through the two coils, and that the actual division of the dc resistance has no effect on time, the ratio of the two resistances can initially be chosen the same as the ratio of the two effective powers:

$$\frac{W_1}{W_2} = \frac{R_1}{R_2}. \tag{35}$$

Finally, knowing that the effective coil constants for the two relays are equal, the desired turns are determined using the already selected coil constant, G_C. Solving these three equations, the individual relays are:

$$R_1 = \frac{W_1 E_0^2}{W_{\text{total}}^2}, \quad R_2 = \frac{W_2 E_0^2}{W_{\text{total}}^2}, \quad N_1 = \sqrt{R_1 G_C}, \quad N_2 = \sqrt{R_2 G_C}. \tag{36}$$

The resistance values R_1 and R_2 can be used as shown above or divided in any other way as long as the total is unchanged. The numbers of turns, however have to be kept fixed.

A particularly simple relationship exists when it is desired to have equal times. In this case the turns on the two coils are equal.

One Winding Known

As is often the case in actual use, a winding must be chosen to be used in series with another known winding and have best operate times. The method described here can be applied to any relay structure, but the numerical values in the analysis are applicable to the wire spring relays only. The first step is to choose the desired resistance for the coil. This is usually set by the heating and power limits, knowing that the higher the total power is, the less the operate time. Then the choice of turns for the coil depends on which of the two coils needs speed the most.

Assume first that we want speed on the coil to be designed, say relay No. 1, rather than the known relay in the circuit, say relay No. 2. Then we want the effective power for this relay as high as possible provided that the coil constant is not too far from optimum. From Table I it has already been noted that the two effective coil constants are always equal when the structures are identical. Also the sum of the two effective powers is exactly equal to the total power; that is, as the effective power to the first relay increases by increasing winding turns, that to the other correspondingly decreases. We see that for relay No. 1, the effective power increases as N_1 increases, but also that the effective coil constant increases. Thus an optimum turns value can be found, where on one side the low effective power slows up the operate time, and on the other the high coil constant does. Fig. 11 shows these optimum values. The solid curves for relay No. 1 are plots of time versus the turns ratio N_2/N_1 with total power, $E^2/(R_1 + R_2)$, and the "series coil constant" of relay No. 2, $N_2^2/(R_1 + R_2)$, as parameters. The optimum turns values for relay No. 1 show up clearly on this curve, and are seen to vary with both parameters.

This relation of optimum turns to the two parameters is shown on Fig. 12, where optimum N_2/N_1 values are plotted against total power with the relay No. 2 series coil constant as the parameter. The added series relay always has the most turns when it is designed for least time. Where speed on relay No. 1 is the only concern, the optimum turns can be chosen directly and easily from Fig. 12. Fig. 11, however is of more general use since it actually gives the times and also shows the time values for the second relay (the dotted curves). Thus the turns can be chosen to approach optimum speed on either relay or to choose a compromise value.

Now for relay No. 2 with total power $E^2/(R_1 + R_2)$, and the second relay series coil constant $N_2^2/(R_1 + R_2)$ as parameters, the effective power decreases and the effective G increases when N_1 is increased, both increasing the time. In other words, the given relay will always be slowed down by any added series relay winding.

The minimum N_1 value is limited by sufficient ampere-turns to operate the first relay. As shown by the dotted curves of Fig. 11, which are the time versus N_2/N_1 curves for relay No. 2, the gain in speed is slight as N_2/N_1 is increased beyond about 2. Thus, although a compromise must always be chosen for speed on relay No. 2, the loss in speed for relay No. 2 is not necessarily great if N_2/N_1 can be chosen near 2. For the case of equal operate times, in every case equality of turns of course applies.

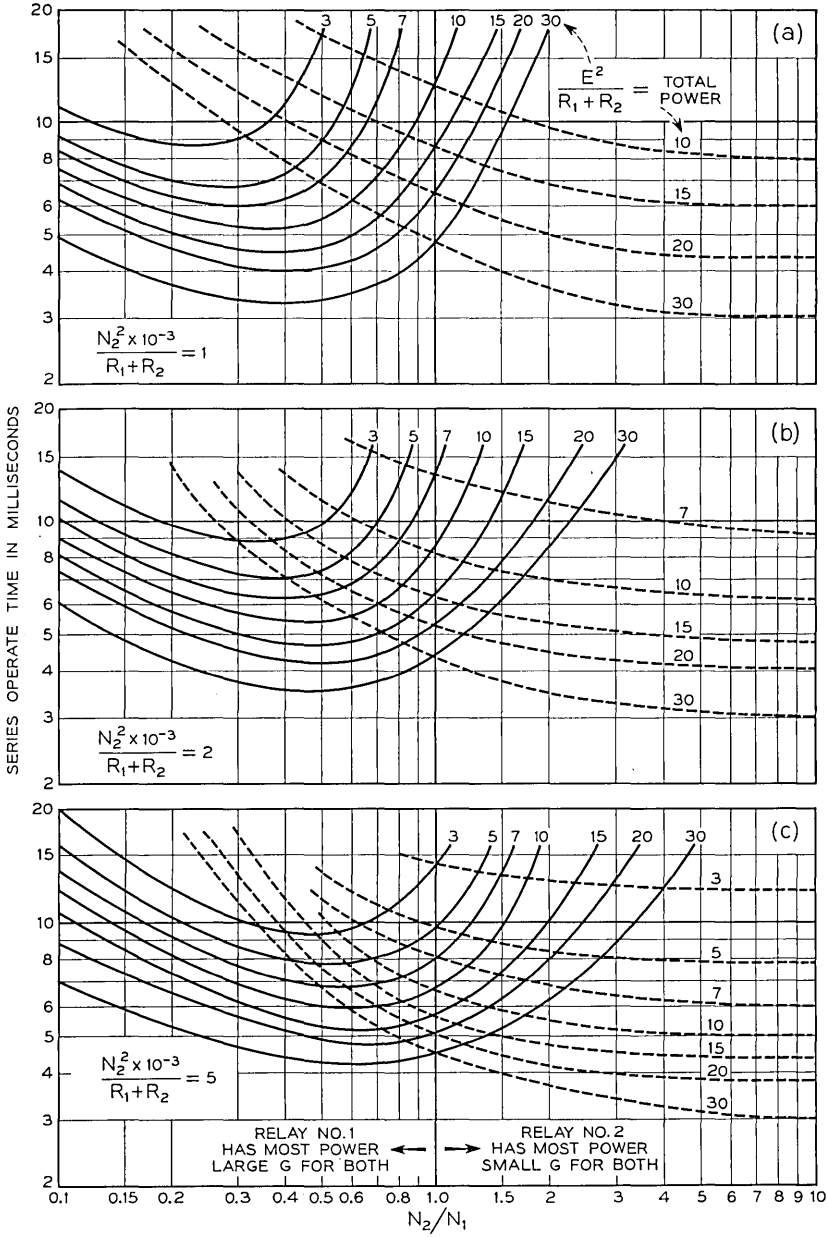
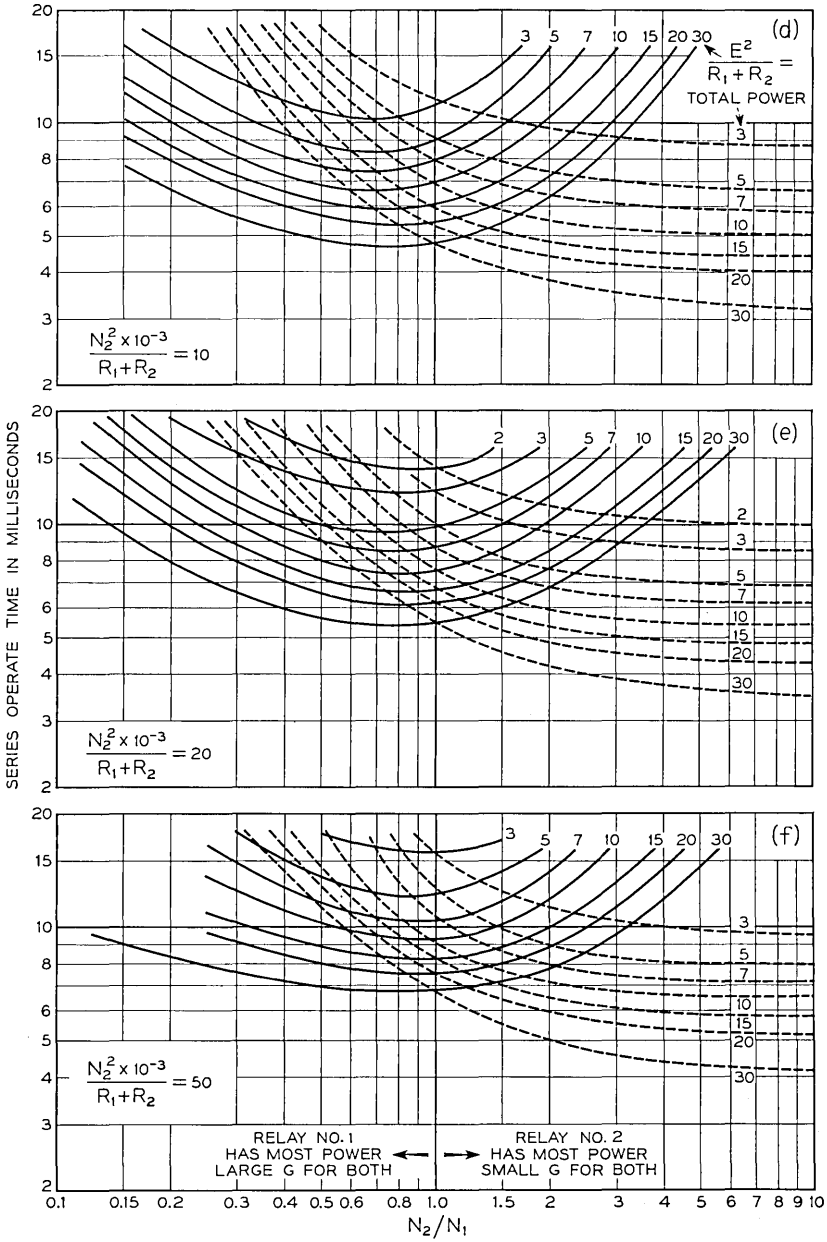


Fig. 11 — Operate time of



series relays versus turns ratio.

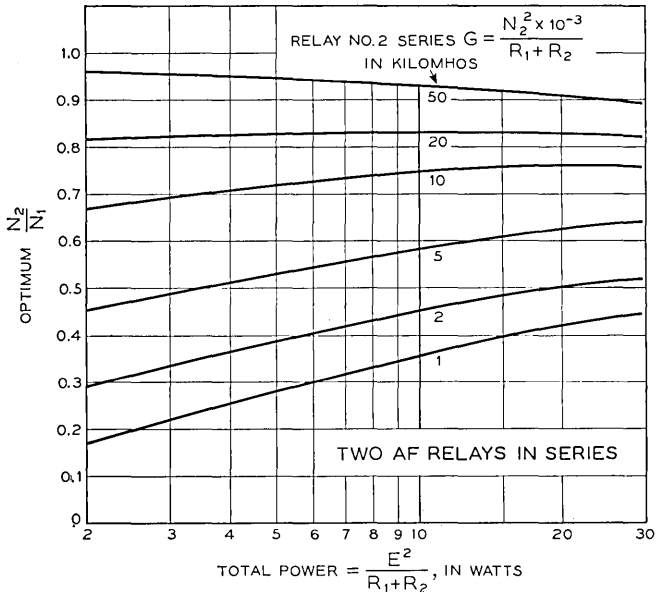


Fig. 12 — Optimum turns ratio for speed on relay No. 1, No. 2 known.

THE OMISSION OF EDDY CURRENT CORRECTIONS

The foregoing discussion of series connected relays developed transformation relationships ignoring eddy current effects. The following discussion will qualitatively show that ignoring the eddy current effects introduces second or smaller order errors.

For very large coil constants and different magnetic structures, the core constants can be ignored. With large coils, the current and flux rise are influenced very little by the core constant.[2] As the coil constant is reduced, the core effect becomes more important, but not significantly so until the speed class is reached. Unfortunately, we are generally only interested in the speed case. In the following, the most affected high speed applications are considered. It is found that the virtual constant potential point is only slightly affected by the eddy currents, and that equivalent single relays for each of the series relays determined using only the winding resistance and turns, with the relay reluctance correction as described above, are sufficiently accurate. The operate time for each relay itself, is affected by the core. This effect of course is included in the measured timing data which are used after equivalents are determined. The present discussion is directed toward the effect of the core eddy currents on the voltage division.

The linear equations used for developing the first approximation operate time equations also represent the behavior of the three element, two terminal network of Fig. 13, shown referred to a one turn admittance form. This exhibits the core eddy current effect as a resistor shunting an ideal inductance, rather than as an infinite line. Whatever value the shunting resistor may have, it affects only the transient response of the network, the part with which we are now concerned.

In what follows, it will be assumed that the transformation relations have already been applied, and the G_c values applying to Fig. 13 are effective values related to the winding turns in accordance with equations (30).

For two such networks in series, the voltage division would be independent of time if the ratios of the shunting to series conductances for each structure were equal. The method developed for similar structures then would apply with no error.

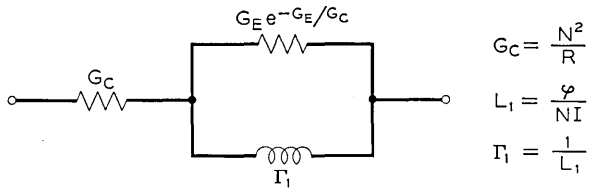


Fig. 13 — Equivalent linear circuit represented by time equation.

The ratio is:

$$\frac{G_E}{G_C} e^{-G_E/G_C} = ae^{-a}.$$

This function has a maximum value of $1/e$; that is, the shunting resistor is at least e times larger than the series resistor. An analysis of the range of core conductances and speed windings in use, shows that the actual resistance ratios are somewhat greater than this and hence the ratios are not quite independent of the structure. However, because the maximum is broad, it reduces the actual range to about 10 per cent, including all windings plus a 2 to 1 G_E change. In turn, this signifies that for a suddenly applied voltage, the initial linear network voltage division would differ from the final by less than 10 per cent. This error decreases with time.

The above discussion applies to linear networks such as Fig. 13. In an actual magnetic structure the initial voltage division is not affected by eddy currents as they have not had time to build up. The voltage

division starts and ends exactly in the ratio of the effective series resistances by virtue of the method used in determining them. For times after circuit closure, a transient voltage error does develop, but it will not approach in magnitude the initial transient error of the linear network. The operate time error will be still smaller. The operate time varies as the two-thirds power of the voltage (power varies as the voltage squared).

These considerations lead to the conclusion that ignoring the eddy currents in setting up the equivalent relays, results at the most in errors of operate time estimates for series connected relays of the order of a few per cent.

RELEASE TIME

The relase time of a relay is not as directly affected by the winding, as is the operate time. For instance, if it is opened by the controlling contact without an *RC* network, the winding current, ignoring arcing at the contacts, abruptly drops to zero. The winding subsequently plays no part in the flux decay. If the winding has a shunting resistor or *RC* circuit, then winding current does flow and some effect is present. A resistor always increases the release time. A favorable *RC* choice can cause a slight decrease in release time. Because of this minor effect, the winding almost always is designed from operate time or sensitivity considerations.

Differing slightly from operate time, the release time is divided into two parts, (1) waiting time plus motion time until actuation of the nearest contact and (2) stagger time. For simplicity the combination (1) above is merely called release waiting time.

The release time of a relay is a more complicated function than is the operate time. The primary cause of this is the closed gap situation, with little stabilizing effect from an air gap. For the earlier operate time studies the air gap largely contributed to the simple exponential relationships. A second effect in release is that the magnetic material is almost always in the non-linear saturated portion of its characteristic. Hence an approximately linear relationship between steady state flux and current cannot be assumed. Finally, for release without an *RC* network, the only current flows in the core, so it completely controls the flux decay. Even with an *RC* network, only a minor decrease in release time is possible. For these reasons, except for rough preliminary estimates of release times during preliminary design, all release data are based on measurements.

Conductance Shunt

In a companion article,[3] a hyperbolic relationship between the flux and current is used to represent the portion of the hysteresis loop of concern in release. This, with a conducting sleeve or shunted winding, gives an excellent representation of the release waiting time. It also provides an understanding of the controlling parameters, even with only eddy currents controlling the release. The form most useful for release consideration is:

$$t = \frac{(\varphi'' - \varphi_0)(G_S + G_C + G_E)}{NI_0} \left(\frac{\ln z}{z - 1} - \frac{1}{z}\right), \tag{37}$$

where

$$z = \frac{\varphi'' - \varphi_0}{\varphi - \varphi_0}.$$

In these experessions, φ is the flux corresponding to the ampere turns NI_0 at which the relay just releases, φ_0 is the residual flux, and φ'' is the asymptotic saturation flux. The conductance terms are the same as for the operate case except G_C now is computed using the dc winding resistance plus any winding shunting resistance. If the winding has no shunt then $G_C = 0$. The operated contact spring load enters through φ and NI_0.

The function of z in brackets has a broad maximum in the region of relay release. It therefore is appropriate to consider the releasing ampere turns as the independent variable and plot the measured release time with the total conductance as a parameter covering the range of interest. As G_E is a characteristic of the structure and not subject to adjustment, the curves are actually labelled in terms of just the sleeve, if any, plus the shunted winding. For the wire spring relay, Fig. 14 shows data in this form. If the function of z were truly constant, the curves would all have a slope of -1 in this plot.

Strictly speaking, the above equation applies only to the time until the magnetic pull has decayed to equality with the operated spring load. Following this is the motion time, during which the pull decays further. For convenience, however, the armature motion time through the distance to the nearest contact is included in the chart for release time as (a) the contact actuation is the means used to measure the time and (b) this much motion always takes place before any contact is actuated.

The further displacement of the armature continues at almost con-

stant velocity. This is because the decreasing magnetic pull (called drag for the release case) has become small and the contact spring load is dropped by the motion. The further forces acting on the armature then are only the back tension and the difference between the rapidly decreasing contact spring load and magnetic drag.

Returning now to equation (37) above, the release time is directly proportional to the sum of the conductances, and to the difference between the saturation and residual fluxes. The conductance variations can be determined from temperature, winding data, sleeve dimensions, and material. The saturation flux is directly proportional to the core

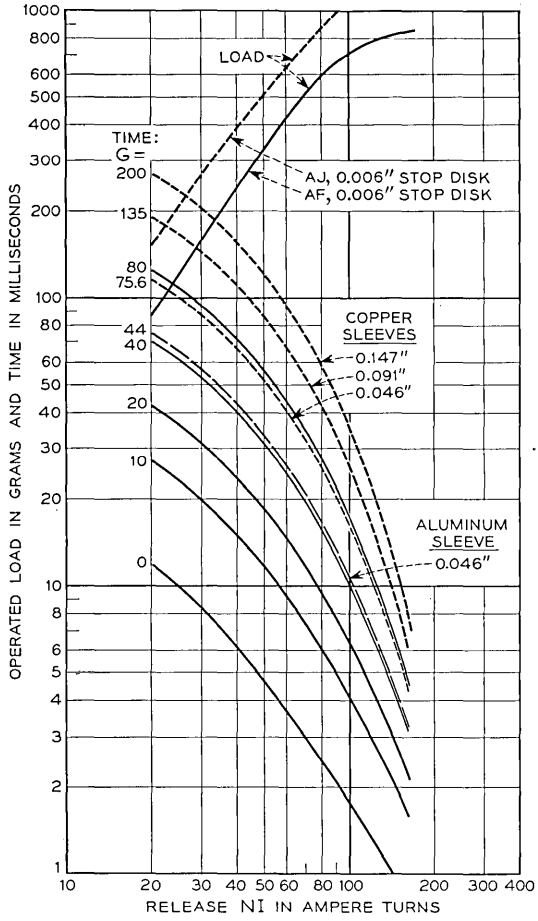


Fig. 14 — Release time with a shunted winding.

cross-section. The residual flux is directly proportional to the coercive force times the length of the magnetic material, and inversely propotional to the closed gap reluctance. The two largest factors affecting the closed gap reluctance are magnetic permeability, and fit of the joints, including variations in stop pin height. By measuring these factors on the test relay and estimating the corresponding values for the desired reference condition, the measured data can be corrected to the reference condition. Then having the release pull curves after magnetic soak for the same reference condition, the release ampere turns for any load under consideration and hence its release waiting time can be determined. These release pull curves are also shown as part of Fig. 14 for the wire spring relay. The ordinate scale is marked for both contact spring load in grams and releasing time in milliseconds. The particular chart shown is for a constant initial number of ampere turns. For other initial values, a correction chart is provided. If not as desired, the time can be adjusted either upward or downward by changing to a different sleeve, shunting resistor, or both. Of course, in no shunting conductance case can the release time be less than the open circuit time.

RC Shunt

The above considerations all related to shunting conductances. These serve the purpose of increasing the release time. The shunting resistor also greatly reduces the transient peak voltage developed when the winding circuit contact is opened. For contact protection reasons, *RC* networks are frequently used across the winding or contact for this same purpose. Such a network also has no power drain when the relay winding is energized. By a suitable choice of capacitance the network can reduce the release time to a value less than the open circuit time. It does this by developing a heavily damped oscillation of winding current. The frequency must be of the order of the reciprocal of the unprotected release time for a release time reduction. This fixes the choice of capacitance to a small range i.e., there is a best capacitor, one for each winding. Smaller values cause the time to increase toward the unshunted case. Larger values also cause an increase but for this case, if too large, an increased time beyond the unshunted case can result.

For speed windings, where timing is important, a network is designed for each. For the slower higher resistance windings, because time is not as important, the closest to the best of the available networks is chosen. This results in a time penalty but furthers standardization.[1]

To evaluate a network, the transformations developed and shown in

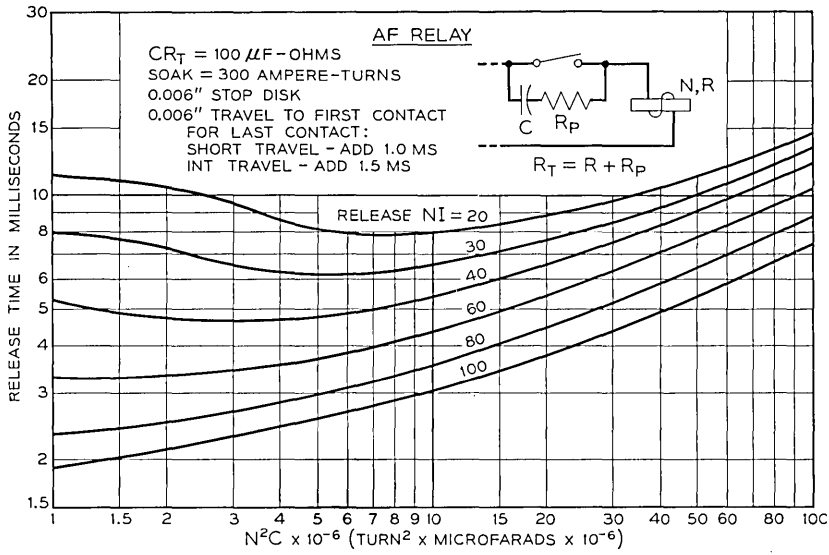


Fig. 15 — Release time as a function of N^2C.

Fig. 5 are used. This permits exact simulation for any winding and network from a time standpoint. For presentation of data, an arbitrary separated product form for the more important variables is used. The particular factors chosen are:

(1) N^2C, a measure of frequency, with NI_0 (RELEASE NI) as a parameter.

(2) RC, a damping term, with NI_0 as a parameter.

(3) Soak NI with N^2/R as a parameter.

Essentially what is assumed is that the release time can be represented well enough as a function of the form:

$$t_1 = f_1(N^2C) \times f_2(RC) \times f_3(NI). \tag{38}$$

Then by holding two factors fixed and varying the third, each factor in turn is evaluated. Generally the accuracy is better than 10 per cent even with this elementary approach. For a few cases the error approaches 15 per cent. Typical charts for these three factors are shown in Figs. 15, 16 and 17.

Release Time for Series Relays and an RC Shunt

With two series relay windings it is usual to provide a single RC network, as shown in Fig. 18. This results in a double infinity of possible combinations if handled directly. However, it can be broken into two

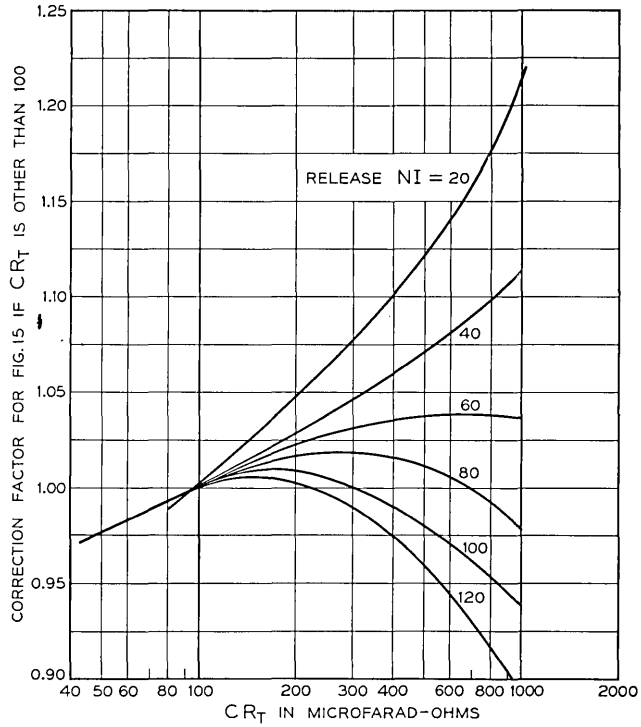


Fig. 16 — Release time correction for *RC* damping.

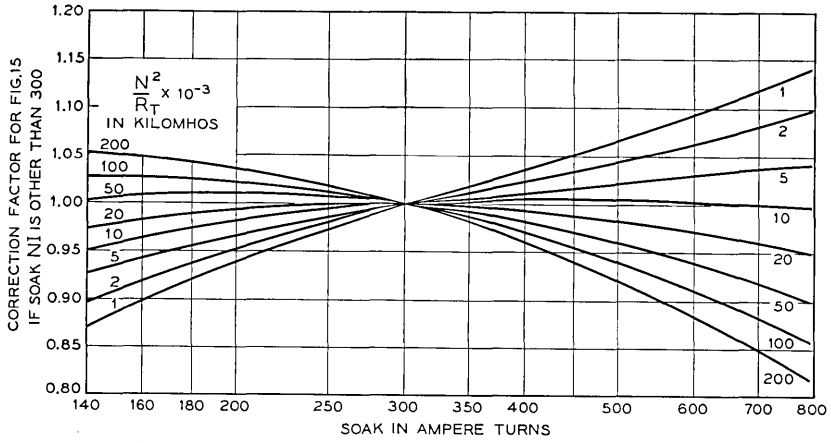


Fig. 17 — Release time correction for magnetic soak.

parts, each of which exactly represents, as individual relays, the two series relays. This is done by first assigning the total of the winding resistances for equal N^2/R values. The two equivalent relays then form a voltage divider, independent of frequency. The RC network next is drawn as two series RC networks as shown in Fig. 19. The procedure now is to determine the component RC values to have the identical voltage divider effect as the relays. Then the two equal voltage points can be connected as shown by the dotted line and no circulating current will flow. The resultant is a three node circuit and each branch can be considered independently, using the method of the preceding section.

There are four unknowns and hence four equations are needed. They are

$$\begin{aligned} R_{11} + R_{22} &= R, \qquad \frac{N_1^2}{R_{11}} = \frac{N_2^2}{R_{22}}, \\ \frac{C_1 C_2}{C_1 + C_2} &= C, \qquad N_1^2 C_1 = N_2^2 C_2 . \end{aligned} \tag{39}$$

The solutions for the four unknowns are shown on Fig. 19.

For experimental measurements, each is transformed again by means of Fig. 5. Note that the time constants all are equal, as they must be, because the same current flows through all elements. However, the initial ampere turns NI are different by virtue of $N_1 \neq N_2$. Thus only one circuit like Fig. 5 need be set up. Measurements then are made with two different voltages to provide the two different ampere turns and the two release times. By choosing the subscript 1 network to represent the actual circuit, then no subscript confusion results in arriving at the simulating circuit of Fig. 5.

Release Time for Similar Parallel Relays and an RC Shunt

Similar parallel relays are split as shown in Fig. 20. Two equal parallel RC networks are first drawn. One is then assigned to each, and then the pairs are divided. The release times are each equal to that of one of the separated circuits.

Summary

The release time of fast electromagnets is influenced much more than the operate time by the fit of the magnetic parts. For release, the small non-magnetic stop disc introduces a relatively small stabilizing air gap compared to the open gap of the operate case. Secondly, the release always starts after an applied magnetomotive force which differs as

between circuits, whereas operation invariably follows an open circuit. Thirdly, with *RC* protection networks connected, a more complicated winding current is present. Finally, the transmission line behavior of the magnetic core material is less masked by the winding, and in fact controls completely for the open circuit case. For these reasons, the

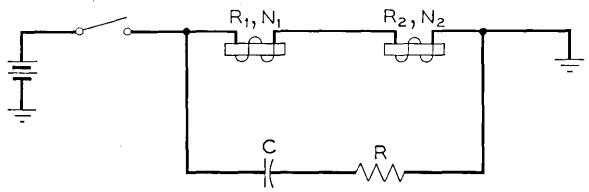


Fig. 18 — Series relays with one *RC* network.

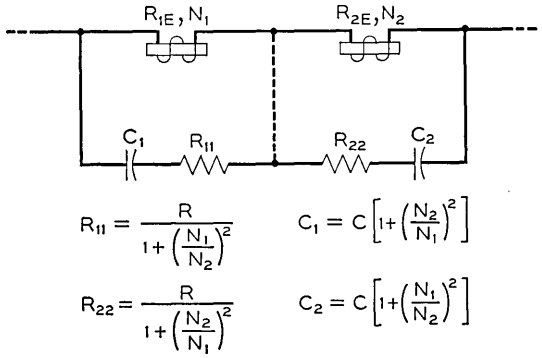


Fig. 19 — Transformation to series *RC* networks.

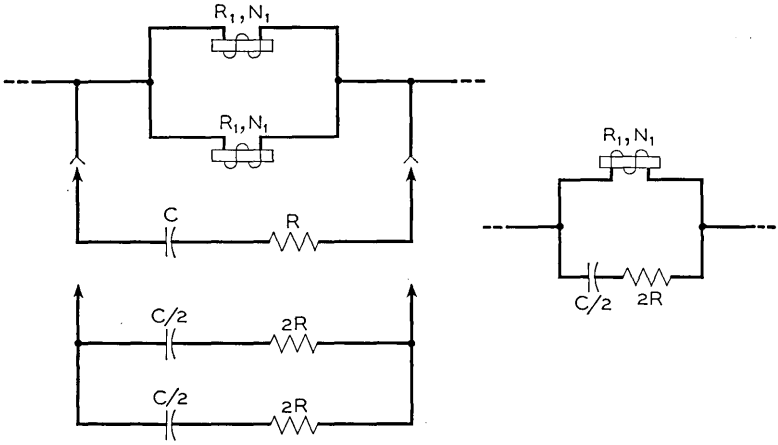


Fig. 20 — Transformation to parallel *RC* networks.

analytical presentation, as noted in Part I, of fast release time data is not now as advanced as the operate time data.

The material presented here describes the present state of the art. For slow releasing relays, the performance can be predicted with accuracy. For fast relays, while the general pattern is known, accurate means for estimating variations have not been developed. The present engineering of releasing relay circuits therefore, depends upon specific measured data in chart form, for each condition.

ACKNOWLEDGMENTS

The analyses leading to the forms of data presentation for the several types of relay timing information are the results of contributions from many people. In particular, the early nomograms for load controlled operation were developed by P. W. Swenson. The node method for separation of series releasing relays shunted by an *RC* network was developed by R. H. Gumley. The graphical method for design of optimum series windings was developed by Mrs. K. R. Randall. To her, I am also indebted for preparation of all the charts of measurements.

REFERENCES

1. H. N. Wagar, Economics of Telephone Relay Applications, page 218 of this issue.
2. M. A. Logan, Dynamic Measurements on Electromagnetic Devices, B.S.T.J., **31**, pp. 1413–1466, Nov., 1953.
3. R. L. Peek, Principles of Slow Release Relay Design, page 187 of this issue.
4. R. L. Peek and H. N. Wagar, Magnetic Design of Relays, page 23 of this issue.

Principles of Slow Release Relay Design

By R. L. PEEK, Jr.

(Manuscript received September 25, 1953)

This article presents an analytical treatment of the relations controlling the release time of slow release relays, in which field decay is delayed by the currents induced in a conducting sleeve or slug. An hyperbolic relation between flux and magnetomotive force, fitting the decreasing magnetization curve, is used for the relation between induced voltage and current in the sleeve circuit in determining the rate of field decay. Methods are given for estimating and measuring the magnetization constants and those appearing in the relation between pull and field flux.

These relations are used as a basis for a discussion of the design of slow release relays and of the adjustment procedures employed to meet timing requirements.

1 INTRODUCTION

Slow release relays are built and adjusted to provide a time delay between the opening of the coil circuit and the release motion which restores the contacts to their unoperated condition. They are used to assure a desired sequence of circuit operation, as, for example, in maintaining a closed path through the slow release relay's contacts during the pulses sent in dialing a digit, and opening this path during the much longer interval between digits. Slow release relays constitute a minor but significant part of the relay population in an automatic central office: about ten per cent of the total.

For economy in manufacture, installation, and use, slow release relays are made as similar to the ordinary or general purpose relays with which they are used as their special requirements permit. Each general structural relay developed for telephone work has had a variant form for slow release use. Thus the *Y* type[1] relay is the slow release form of the *U* type relay[2] widely used in the Bell System, while the *AG* relay is the slow release variant of the recently developed wire spring[3] relay.

The circuit functions of most slow release relays permit considerable variation in release time if the minimum delay specified is assured. This

tolerance in the timing requirements permits the use of more economical practices in the construction and use of slow release relays than would be needed for closer control. As these wide tolerances apply to the great majority of applications, over-all economy is attained by developing slow release relays with reference to them, using other devices for the special applications requiring close timing control.

The magnitude of time delay desired is fixed by the operate and release times of the associated general purpose relays. The latter times lie in the range from 5 to 50 milliseconds, with the majority in the lower part of the range. The delays required to assure sequences of events each requiring time intervals of this order consequently cover the range from 50 to 500 milliseconds (one twentieth to one half a second). Delays of less than 100 milliseconds can generally be provided by ordinary general purpose relays having shorted secondary windings or sleeves (single turn conductors) to delay their release. Special slow release relays are used for delays in excess of 100 milliseconds.

Factors Controlling Release Delay

In terms of circuit operation, release time is the interval from the opening of the coil circuit to the completion of contact actuation during the return motion of the armature. This time is the sum of (a) the time for the magnetic field to decay to the level at which the pull just equals the operated spring load and (b) the motion time for contact actuation. The motion time is never more than a few milliseconds, and is therefore a trivial increment to the long delays of slow release relays. The release time of such relays is therefore, for practical purposes, simply the time of field decay.

When the coil circuit is opened, the field decay induces currents in any circuit linking the field. These currents tend to maintain the field and to delay its decay. A short circuited winding or the single high conductivity turn provided by a copper sleeve gives a high magnetomotive force for a low induced voltage, resulting in a relatively slow decline in field strength. The time for the flux to reach a given level depends upon the conductance of the sleeve, and upon the reluctance of the electromagnet: the ratio of magnetomotive force to flux. The flux level which determines the end of the delay is that for which the pull equals the spring load. The delay is therefore prolonged by a pull characteristic such that the load is held until the flux drops to a minor fraction of its initial value.

Thus the essential features of a slow release relay are a shorted winding

or sleeve, a low reluctance magnetic circuit, and a high level of pull for relatively low values of field strength. In the following analysis of slow release performance, expressions are developed for the time of field decay for the decreasing magnetization characteristic, for the reluctance of the electromagnet, and for its pull characteristics. These expressions permit the estimation of release times, indicate the design conditions to be satisfied to attain a desired level of delay, and indicate the effect on the release time of variations in the spring load, in the dimensions of the electromagnet, and in other design parameters.

The notation used in this article conforms to the list that is given on page 257.

2 FIELD DECAY RELATIONS

If a closed circuit of resistance R and N turns links a magnetic field of flux φ, the voltage equation is:

$$iR + N\frac{d\varphi}{dt} = 0,$$

where i is the circuit current. Multiplying by $4\pi N$ and dividing by R this equation may be written as:

$$\mathfrak{F}_i + 4\pi G_i \frac{d\varphi}{dt} = 0,$$

where $\mathfrak{F}_i$ is the magnetomotive force of this circuit, and G_i is its value of N^2/R, which may be termed the equivalent single turn conductance. If there are several such circuits linking the same magnetic field, a similar expression applies to each, and these may be added to give the equation:

$$\mathfrak{F} + 4\pi G \frac{d\varphi}{dt} = 0, \tag{1}$$

where $\mathfrak{F} = \sum \mathfrak{F}_i$, and $G = \sum G_i$. In the case of a slow release relay, one linking circuit is usually a sleeve, whose conductance may be designated G_S. The applications are identical for a short circuited winding, if the applicable value of N^2/R is substituted for G_S. In either case, G also includes a term G_E representing the net effect of the eddy current paths. As the eddy currents at different distances from the center of the core link different fractions of the total field, this representation by a single term is an approximation. As shown in a companion article,[4] however, the approximation is satisfactory when G_E is a minor part of G, as in the

slow release case. In this same article it is shown that, under these conditions, the relation between the changing values of $\mathfrak{F}$ and φ in (1) is substantially identical with that given by static magnetization measurements.

The static magnetization relation between $\mathfrak{F}$ and φ may therefore be substituted in (1), which may then be integrated to give the relation between φ and t. For linear magnetization, or constant reluctance $\mathfrak{R} = \mathfrak{F}/\varphi$, integration of (1) gives the equation.

$$\frac{\varphi}{\varphi_1} = e^{-t/t_G},$$

where φ_1 is the value of φ for $t = 0$, and $t_G = 4\pi G/\mathfrak{R}$, the time constant for this simple exponential decay. For release, however, the magnetization relation is not linear, but has the character shown in Fig. 1. The φ versus $\mathfrak{F}$ relation is asymptotic to the saturation flux φ'', and has an intercept φ_0 at $\mathfrak{F} = 0$, resulting from the remanence of the magnetic material. The relation between φ and t in release is given by (1) with the relation between φ and $\mathfrak{F}$ that of the decreasing magnetization curve illustrated in Fig. 1.

Graphical Determination of Release Time

Writing $4\pi Ni$ for $\mathfrak{F}$ in (1), the integral form of this equation is:

$$\frac{t}{G} = \int_{\varphi}^{\varphi_1} \frac{d\varphi}{Ni}, \tag{2}$$

where t is the time for the field to decay from an initial value φ_1 to a final value φ. If experimental data for the decreasing magnetization curve are available, the integral of (2) may be evaluated graphically as is indicated in Fig. 2.

For purposes of illustration, the left hand plot shows two decreasing magnetization curves, such as might be obtained with two models of the same relay. The dashed curve 1 corresponds to a higher value of φ_0 and a lower reluctance than apply to the solid curve 2. To evaluate the integral of equation (2), the curves are replotted as at the right to give φ versus $1/(Ni)$. Then the integral of equation (2) is given by the area between the curve and the axis of φ. For curve 1, for example, the area o-d-a' measures the value of t/G for the flux to decay from φ_1 to the value of φ at a.

If the areas are evaluated for a series of points, such as a and c, the resulting values of t/G can be plotted, either against the correspond-

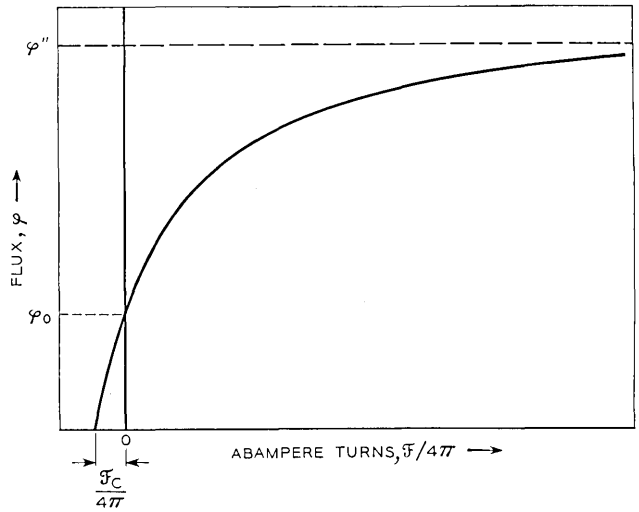


Fig. 1 — Decreasing magnetization of an electromagnet.

ing values of φ, as in Fig. 3, or against the corresponding values of NI, as in Fig. 4. In comparing two models, corresponding to curves 1 and 2 in these figures, the curves of t/G versus φ in Fig. 3 indicate the comparative release times for the same spring load, as the pull versus φ relation is the same for similar structures. Thus in Fig. 3 the times for the pull to drop to a common value for curves 1 and 2 correspond to the points c and b, for example, and are measured by the areas o-e-c' and o-e-b' respectively. Hence curve 1 gives longer times in this comparison than curve 2, as illustrated in Fig. 3.

If, on the other hand, the spring loads for the two models are adjusted so that they release on the same value of NI, the release times to be compared correspond to decay to the points a and b, for example. The corresponding values of t/G are given by the areas o-d-a' and o-e-b'. These differ little from each other, and consequently curves 1 and 2 in Fig. 4 are similar. It is therefore possible to adjust individual relays to give nearly the same release time by adjusting the spring load so that they release on the same value of Ni. Advantage is taken of this fact in the adjustment practices employed with slow release relays, as discussed in a latter section.

This graphical method for the prediction of release times, which was developed by H. N. Wagar,[5] is described here because of its utility in demonstrating the relation between the timing relations and the character of the demagnetization curves. It provides a means for accurately predicting the release times in specific cases, but does not afford as

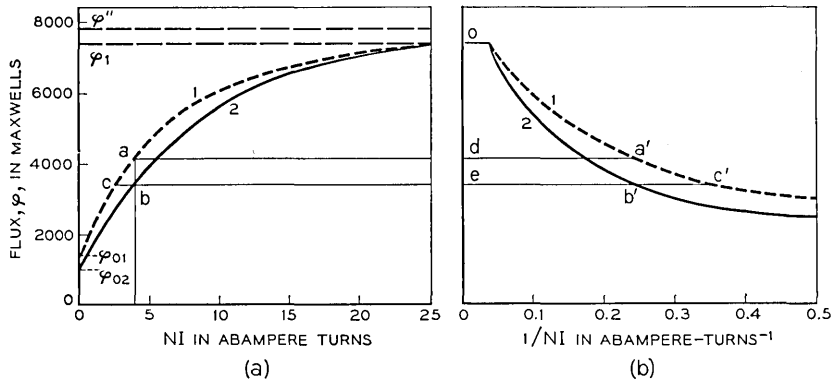


Fig. 2 — Graphical determination of release time.

explicit a statement of the general relations applying as does the approximate analytical treatment outlined below.

Hyperbolic Approximation to Decreasing Magnetization Curve

The decreasing magnetization curve, as illustrated in Fig. 1, has the general character of a rectangular hyperbola, and may therefore be represented approximately by the empirical equation:

$$\frac{\mathfrak{F} + \mathfrak{F}_c}{\varphi} = \frac{\mathfrak{R}''\varphi''}{\varphi'' - \varphi}, \tag{3}$$

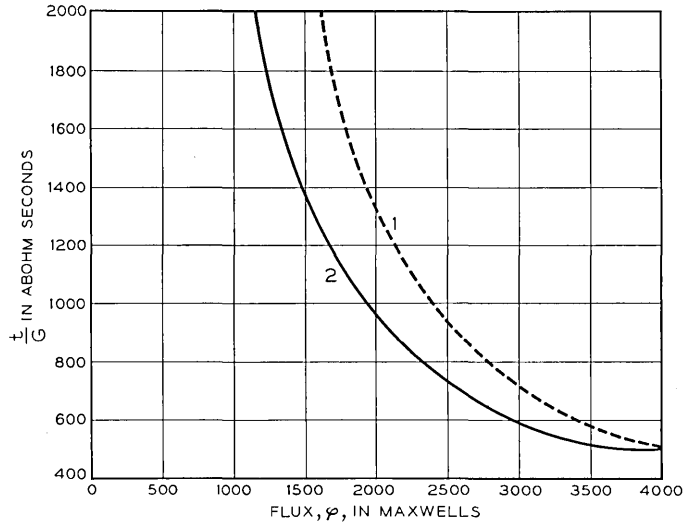


Fig. 3 — Release time versus flux.

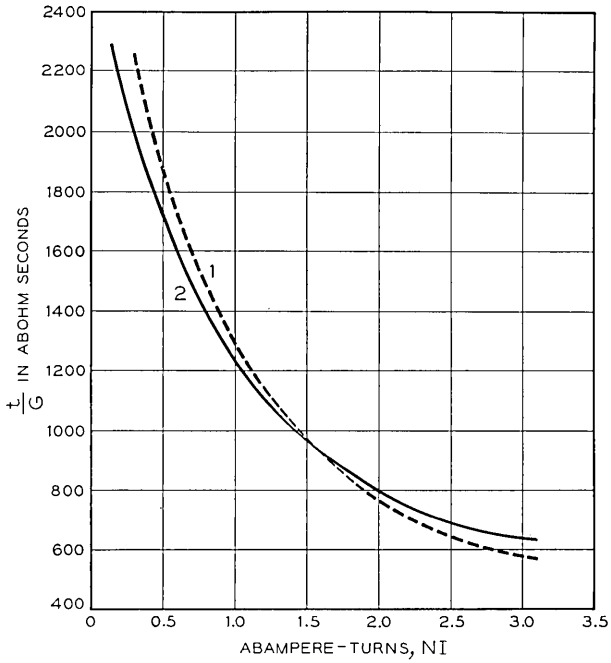


Fig. 4 — Release time versus ampere turns.

for which $\varphi = 0$ for $\mathfrak{F} + \mathfrak{F}_C = 0$, while φ'' is the asymptote approached by φ as $\mathfrak{F}$ is increased. $\mathfrak{R}''$ is the initial incremental reluctance, the value of $d\mathfrak{F}/d\varphi$ for $\varphi = 0$. If φ_0, as in Fig. 1, is the value of φ for $\mathfrak{F} = 0$, the resulting expression for $\mathfrak{F}_C$ may be substituted in (3) and this equation written in the alternative form:

$$\frac{\mathfrak{F}}{\mathfrak{R}''\varphi''} = \frac{\varphi}{\varphi'' - \varphi} - \frac{\varphi_0}{\varphi'' - \varphi_0}. \tag{4}$$

The incremental reluctance $\mathfrak{R}_i$, the value of $d\mathfrak{F}/d\varphi$ at $\varphi = \varphi_0$, is given by:

$$\mathfrak{R}_i = \left(\frac{\varphi''}{\varphi'' - \varphi_0}\right)^2 \mathfrak{R}''. \tag{5}$$

The expression for $\mathfrak{F}/(4\pi)$ given by (4) may be substituted for Ni in (2) to give an expression for the release time t. After clearing fractions, the resulting equation is:

$$\frac{t}{G} = \frac{4\pi(\varphi'' - \varphi_0)^2}{\mathfrak{R}''\varphi''^2} \int_{\varphi}^{\varphi_1} \left(\frac{1}{\varphi'' - \varphi_0} - \frac{1}{\varphi - \varphi_0}\right) d\varphi.$$

From (5), the term outside the integral sign is $4\pi/\mathcal{R}_i$. On integration, the following expression is obtained for the release time t:

$$t = \frac{4\pi G}{\mathcal{R}_i}\left(\frac{\varphi - \varphi_1}{\varphi'' - \varphi_0} + \ln\frac{\varphi_1 - \varphi_0}{\varphi - \varphi_0}\right).$$

In slow release relays, a "soak" or high ampere turn value is applied in operation, and the initial value φ_1 is close to the saturation value φ''. It is therefore a satisfactory approximation to take $\varphi_1 = \varphi''$ in the preceding equation, which then reduces to:

$$t = \frac{4\pi G}{\mathcal{R}_i}\left(\ln z - 1 + \frac{1}{z}\right), \tag{6}$$

where,

$$z = \frac{\varphi'' - \varphi_0}{\varphi - \varphi_0}. \tag{7}$$

The time therefore varies as the bracketed term in (6), which is shown plotted against z in Fig. 5. As z varies inversely as $\varphi - \varphi_0$, the difference between the flux and its ultimate value, (6) gives the release time for the value of φ at which the pull equals the operated load.

To obtain an expression for the release time in terms of the ampere turn value at which release occurs requires an expression for z in terms of $\mathcal{F}$, or $4\pi NI$. This may be obtained by substituting in (4) the expression for $\mathcal{R}''$ given by (5). The resulting equation reduces to:

$$\mathcal{F} = (\varphi'' - \varphi_0)\mathcal{R}_i\frac{\varphi - \varphi_0}{\varphi'' - \varphi} = \frac{(\varphi'' - \varphi_0)\mathcal{R}_i}{z - 1},$$

giving the equation:

$$z = 1 + \frac{(\varphi'' - \varphi_0)\mathcal{R}_i}{\mathcal{F}}. \tag{8}$$

If $\mathcal{F}$ is the value of $4\pi NI$ at which release occurs, the corresponding value of z given by (8) may be substituted in (6) to determine the release time. In this way there have been determined the values of $t\mathcal{R}_i/(4\pi G)$ plotted against $\mathcal{F}/(\mathcal{R}_i(\varphi'' - \varphi_0))$ in Fig. 6. This is a universal curve for the relation between release time and the ampere turn value at which release occurs. The observed relation for any specific case is given by this curve, displaced vertically by the value of $4\pi G/\mathcal{R}_i$ and horizontally by the value of $\mathcal{R}_i(\varphi'' - \varphi_0)$ for the case in question. This is illustrated in Fig. 7, which shows the observed release time versus release ampere turn curves for the Y type relay for two sleeve sizes, corresponding to

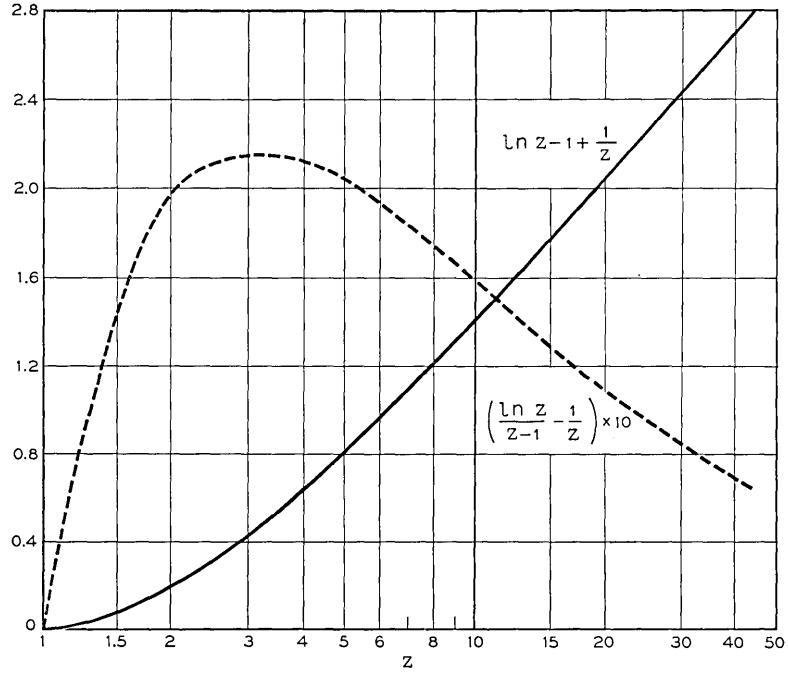


Fig. 5 — Factors for computing release time.

different values of G. For each sleeve size two curves are shown, corresponding to the limits of variation in magnetic characteristics, or in $\mathfrak{R}_i$ and $\varphi'' - \varphi_0$. The dotted curve included for comparison is the relation of Fig. 6.

An important property of the t versus $\mathfrak{F}$ relation can be demonstrated by substituting in (6) the expression for $\mathfrak{R}_i$ given by (8). There is thus obtained the equation:

$$t = \frac{4\pi G(\varphi'' - \varphi_0)}{\mathfrak{F}} \left(\frac{\ln z}{z - 1} - \frac{1}{z} \right). \tag{9}$$

A plot of the function of z appearing in brackets in this equation is included in Fig. 5. In the vicinity of the maximum at $z = 3.09$, this function is nearly a constant, varying little for values of z between 2 and 6. In this range t varies inversely with $\mathfrak{F}$, and the proportionality constant depends only upon G, the sleeve conductance, and the term $\varphi'' - \varphi_0$, which is relatively independent of material and dimensional variations, as shown later. This confirms the conclusion, previously indicated by graphical analysis, that through a considerable range of operation the

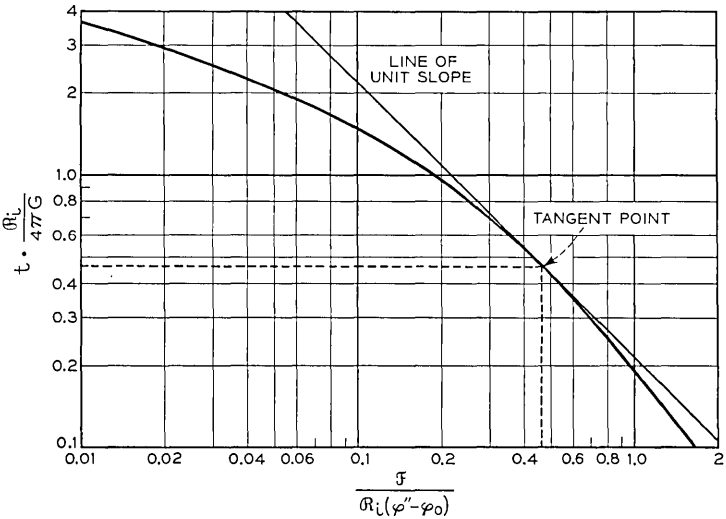


Fig. 6 — General relation between release time and ampere turns.

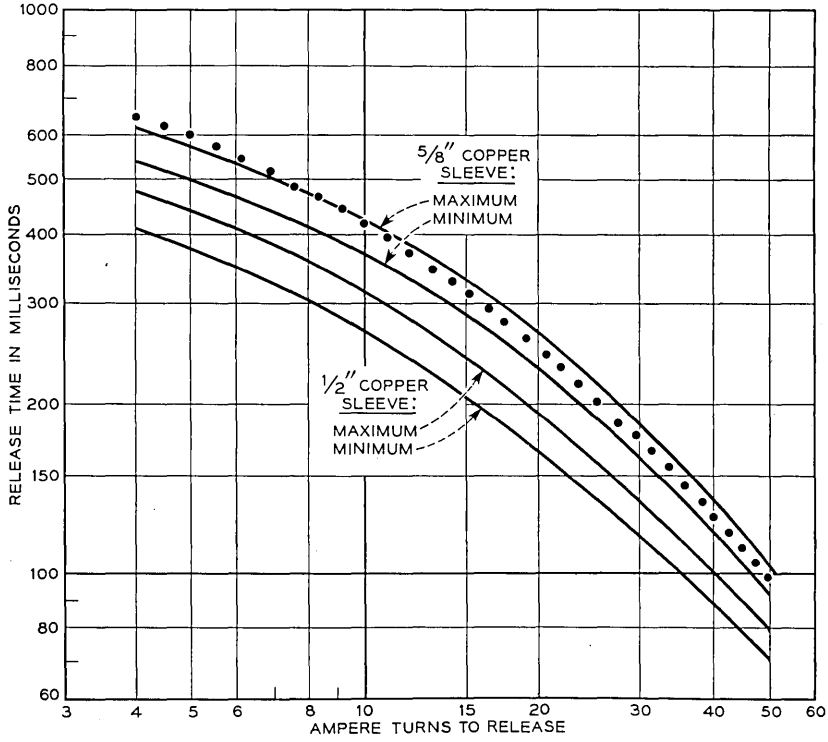


Fig. 7 — Observed release time characteristics.

release time is relatively independent of magnetic variations, provided it is adjusted to release at a specified ampere turn value.

The range in which t is inversely proportional to $\mathfrak{F}$ is that in which the logarithmic plot of Fig. 6 has a slope near unity. The point of tangency with a line of unit slope coincides with the value of z for which the bracketed function of z in (9) is a maximum. By equating the derivative of this to zero, it is found that this maximum occurs for $z = 3.09$, corresponding, from (8), to a value of 0.46 for $\mathfrak{F}/(\mathfrak{R}_i(\varphi'' - \varphi_0))$. From (6), the corresponding value of $t.\mathfrak{R}_i/(4\pi G)$ is 0.46. Thus by drawing a line of unit slope tangent to the observed relation between t and $\mathfrak{F}$, and reading the co-ordinates of the point of tangency, $4\pi G/\mathfrak{R}_i$ and $\mathfrak{R}_i\ (\varphi'' - \varphi_0)$ may be evaluated. If G is known, these suffice to determine the values of $\mathfrak{R}_i$ and $\varphi'' - \varphi_0$.

Thus the release time is given by (6), and may be expressed in terms of the flux φ at which release occurs by means of (7), or in terms of the corresponding value of $\mathfrak{F}$, or $4\pi NI$, by means of (8). To relate the time to the spring load determining release, expressions are required relating the pull to φ or to $\mathfrak{F}$. To relate both time and pull characteristics to the design requires means for evaluating the magnetic constants and G in terms of the dimensions and materials of the design. The magnetic constants and the pull relations are discussed in the two following sections.

3 DECREASING MAGNETIZATION RELATIONS

To determine the decreasing magnetization relation experimentally, the magnet is demagnetized, and a measurement made of the flux developed on applying a full "soak," or high ampere turn value. The decreasing flux is then measured as the applied current is reduced, and finally reversed to determine $\mathfrak{F}_c$, the value required to restore the field to zero. The relation between φ and $\mathfrak{F}$ thus determined has the character shown in Fig. 1. If this curve conforms to the empirical relation given by (3) and (4), it is characterized by three constants: $\mathfrak{R}''$, φ'', and either $\mathfrak{F}_c$, as in (3), or φ_0, as in (4). These may be evaluated from measurements or estimated in preliminary design by the procedures indicated below.

Experimental Determination of Magnetic Constants

Equation (3) may be written in the form:

$$\frac{\mathfrak{F}_c + \mathfrak{F}}{\varphi} = \mathfrak{R}'' + \frac{\mathfrak{F}_c + \mathfrak{F}}{\varphi''}. \tag{10}$$

As $\mathfrak{F}_C$ may be read directly from the measured curve, as indicated in Fig. 1, values of $(\mathfrak{F}_C + \mathfrak{F})/\varphi$ may be computed from corresponding values of φ and $\mathfrak{F}$, and plotted against the corresponding value of $\mathfrak{F}_C + \mathfrak{F}$. One such plot for a slow release relay model is shown in Fig. 8. In agreement with (10), the plot is approximately linear over the range covered, and the slope and intercept may be used to evaluate $\mathfrak{R}''$ and φ'' as indicated in the figure. If co-ordinate paper is used, as in Fig. 8, with radial lines spaced to give a convenient scale of φ, the value of φ'' is that corresponding to the radial line parallel to the plot. The observed linearity of this relation does not extend to much higher values of $\mathfrak{F}$, and the observed asymptote φ'' is fictitious, as the full curve is concave downwards, asymptotic to a lower value of φ than that found in this way. If the observed value of φ_0 does not agree with that computed from values of the other constants, it is preferable to use the observed rather thad the computed value of φ_0.

As shown in the preceding section, values of $\mathfrak{R}_i$ (which is related to $\mathfrak{R}''$ by (5)) and of $\varphi'' - \varphi_0$ may be independently determined from measurements of release time versus release ampere turns. In cases of disagreement, these values are to be preferred to those determined from the decreasing magnetization measurements. The latter are primarily of interest in checking design estimates, and in indicating the effects of variations in dimensions and material properties.

Estimation of Magnetic Constants

In the "tight" magnetic circuit of a slow release relay, the leakage field may be ignored, and the reluctance taken as the sum of the iron reluctance and that of the air gaps, which may be designated $\mathfrak{R}_E$. The estimation of $\mathfrak{R}_E$ is discussed in the following section, in connection with the pull relations. The iron reluctance is substantially that of a permanent magnet, supplying flux to the external circuit of reluctance $\mathfrak{R}_E$. In using (3) to characterize the magnetization relation, the demagnetization in the second quadrant is taken as continuous with the decreasing magnetization in the first quadrant. Thus the value of $\mathfrak{F}_C$ is that which would apply to a permanent magnet of length ℓ, equal to that of the iron path, as given by:

$$\mathfrak{F}_C = H_C \ell, \tag{11}$$

where H_C is the coercive force of the material. For the soft magnetic materials used for slow release relays, H_C is of the order of 0.5 to 1.0 oersteds, so if $\ell = 10$ cm, $\mathfrak{F}_C$ lies in the range from 5 to 10 gilberts (4 to 8 ampere turns).

Similarly, the initial iron reluctance corresponds to the slope of the demagnetization curve at H_C, as given by the initial permeability μ''. Thus the total initial reluctance $\mathfrak{R}''$ is given by the equation:

$$\mathfrak{R}'' = \mathfrak{R}_E + \frac{1}{\mu''} \sum \frac{\ell}{a}, \tag{12}$$

where $\sum \ell/a$ denotes the sum of terms corresponding to the component parts of the iron path, each term giving the length of the part divided by its cross sectional area. As μ'' is of the order of 20,000 for soft magnetic materials, the iron reluctance term is small compared with $\mathfrak{R}_E$, which is typically of the order of 0.010 cm^{-1} or less.

The value of φ'' may be estimated as the saturation flux for the core, or $B''a$, where a is the cross sectional area of the core. B'' is the saturation density, which may be taken as the value of B_M listed in a companion article.[6] Estimates of φ'' thus obtained are lower than those which best fit the decreasing magnetization measurements. No convenient means for correcting this disparity has been found, other than an arbitrary increase by a factor of 20 per cent, based upon experience. In a particular relay design, however, the observed values of φ'' vary directly with the core cross section.

4 PULL RELATIONS

As shown by equation (6), the release time varies inversely as the incremental reluctance $\mathfrak{R}_i$. This is proportional to $\mathfrak{R}''$, in which the dominant

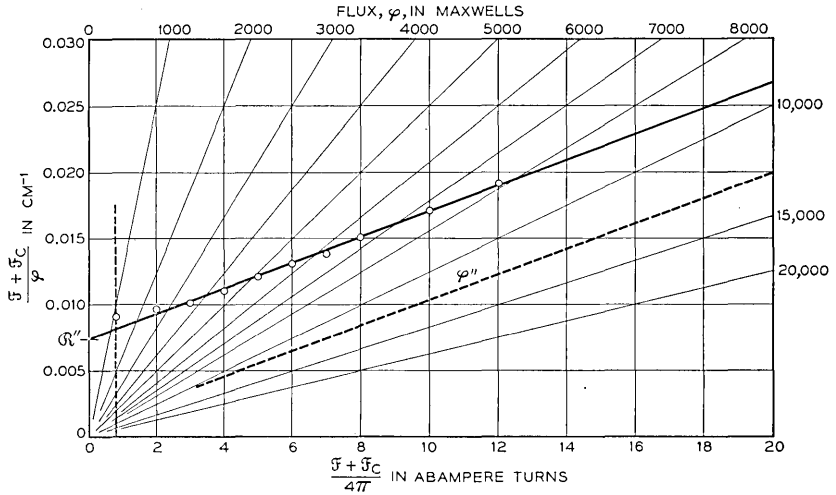


Fig. 8 — Evaluation of magnetic constants from measurements of decreasing magnetization.

term, from (12), is the total air gap reluctance $\mathcal{R}_E$. To obtain long delays, therefore, $\mathcal{R}_E$ must be made small, and consequently sensitive to small dimensional variations. These may be compensated in an initial adjustment, but subsequent changes must be minimized if constancy of performance is to be attained.

Two expedients have been used to provide a small and stable gap reluctance. The older one is the use of a "residual screw," an adjustable non-magnetic member which serves as an armature stop and assures a small air gap at the pole face. With this scheme, the residual screw is used to adjust the relay. The alternative scheme, used in the flat type relays of the Bell System, is to employ a domed pole face on the armature, providing a spherical surface in contact with a mating plane surface on the core. The only effective air gap at the point of contact is that of the chrome finish on the parts. With this scheme, the relay is adjusted by varying the spring tension, and thus the operated load.

General expressions for the pull of electromagnets are discussed in a companion article,[6] where it is shown that the pull F provided by a gap flux φ is given by the equation:

$$F = \frac{\varphi^2}{8\pi} \frac{d\mathcal{R}_G}{dx}, \tag{13}$$

where x is the dimension in the direction of the pull, and $\mathcal{R}_G$ is the gap reluctance. In the usual case, $\mathcal{R}_G$ varies linearly with x, and $d\mathcal{R}_G/dx$ has the constant value $1/A$, where A is the effective pole face area. This is applicable to any case of plane mating surfaces having an appreciable separation, including the configuration usually employed with a residual screw as separator.

Pull for a Domed Pole Face

In the case of a domed pole face there is a concentration of the field near the point of contact, which varies with the effective air gap at this point. An expression for the reluctance can be developed for the idealized configuration shown in Fig. 9. In this, R is the radius of a spherical surface mating with a plane over the projected area A bounded by the radius αR. The separation x is that measured at the center of A. As indicated in the figure, an expression can be obtained for the gap reluctance $\mathcal{R}_G$ in terms of its reciprocal, or permeance. The latter is given by the integration of the permeances of the differential rings within the projected area. $\mathcal{R}_G$ is conveniently expressed in terms of the ratio $\mathcal{R}_\infty/\mathcal{R}_G$

given by the equation:

$$\frac{\mathcal{R}_\infty}{\mathcal{R}_G} = 2\pi R\mathcal{R}_\infty \ln\left(1 + \frac{1}{2\pi R\mathcal{R}_\infty}\right), \tag{14}$$

in which $\mathcal{R}_\infty = x/A$, the value of $\mathcal{R}_G$ which would apply if the spherical radius were infinite, and both mating areas plane. On substituting this expression for $\mathcal{R}_G$ in (13), there is obtained the following equation for the pull at a domed pole face:

$$F = \frac{\varphi^2}{8\pi A}\left(\frac{\mathcal{R}_G}{\mathcal{R}_\infty}\right)^2 \frac{2\pi R\mathcal{R}_\infty}{1 + 2\pi R\mathcal{R}_\infty}. \tag{15}$$

The right hand side of (14), and the expression given by (15) for the ratio of F to $\varphi^2/(8\pi A)$ are both functions of a single dimensionless parameter: $R\mathcal{R}_\infty$. These two ratios are plotted against this parameter in Fig. 10. These curves show how the reluctance and pull values compare

For $\theta < \alpha$ small,

$$r = R \sin\theta = R\theta$$

$$h = R(1 - \cos\theta) = \frac{R\theta^2}{2}$$

Gap Reluctance

$$\frac{1}{\mathcal{R}_G} = \int_0^{\alpha R} \frac{2\pi r \times dr}{x + h}$$

$$= \int_0^{\alpha} \frac{2\pi R^2\theta d\theta}{x + \frac{R\theta^2}{2}}$$

$$= 2\pi R \ln\left(1 + \frac{R\alpha^2}{2x}\right)$$

Let: $\mathcal{R}_\infty = \frac{x}{A} = \frac{x}{\pi(\alpha R)^2}$

$$\frac{\mathcal{R}_\infty}{\mathcal{R}_G} = 2\pi R\mathcal{R}_\infty ln\left(1 + \frac{1}{2\pi R\mathcal{R}_\infty}\right)$$

Pull for Flux φ

$$F = \frac{\varphi^2}{8\pi} \times \frac{d\mathcal{R}_G}{dx} = \frac{\varphi^2}{8\pi A} \times \frac{d\mathcal{R}_G}{d\mathcal{R}_\infty}$$

$$= \frac{\varphi^2}{8\pi A} \times \left(\frac{\mathcal{R}_G}{\mathcal{R}_\infty}\right)^2 \times \frac{2\pi R\mathcal{R}_\infty}{1 + 2\pi R\mathcal{R}_\infty}$$

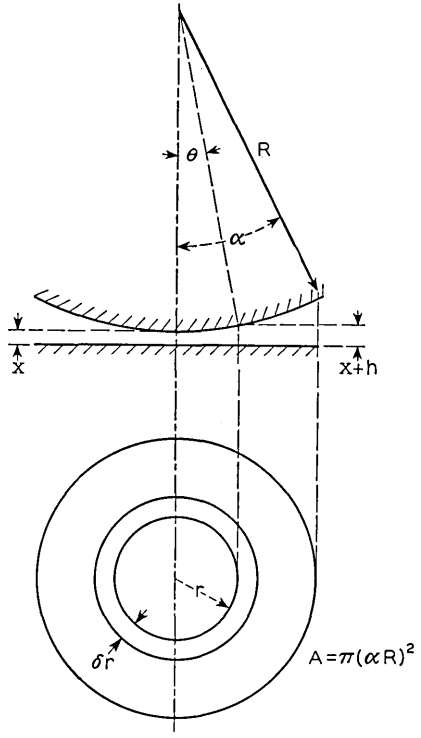


Fig. 9 — Reluctance and pull of a domed pole face.

with those for a gap of the same separation and projected area, but with plane pole faces (a dome of infinite radius).

Ampere Turn Sensitivity

The effect of the relation between F and φ on the release time cannot be conveniently deduced from (6), which involves both the term in z, which varies with φ, and the reluctance $\mathcal{R}_i$, which also enters the pull relation. If the pull is expressed in terms of $\mathfrak{F}$, however, (9) shows that the time varies inversely with $\mathfrak{F}$, and is substantially independent of the reluctance, provided z is in the usual range from 2 to 6. It follows that maximum release time can be obtained for a given load if the value of $\mathfrak{F}$

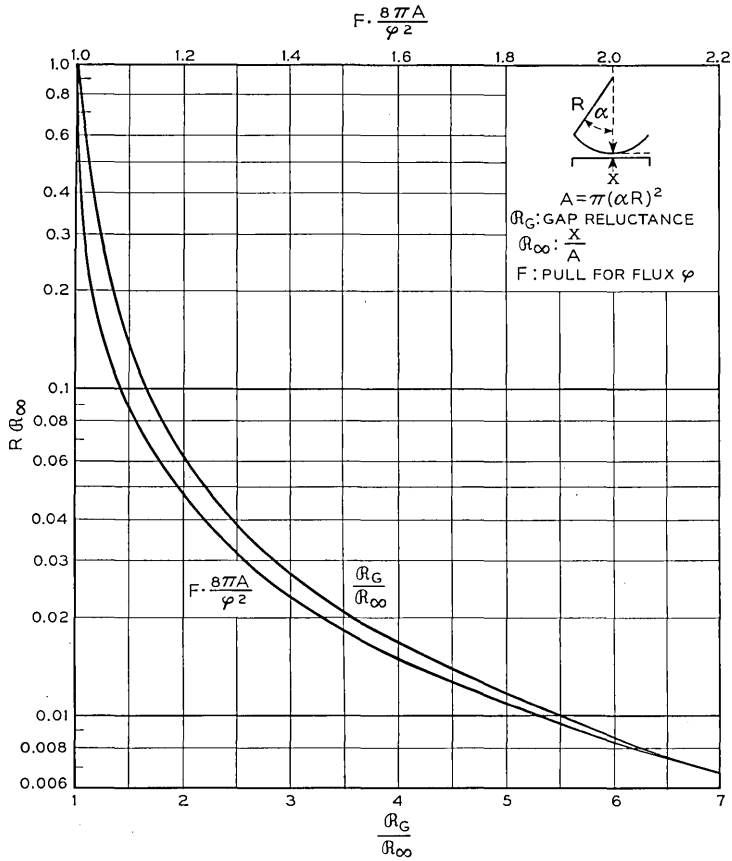


Fig. 10 — Reluctance and pull relations of a domed pole face electromagnet.

required to operate this load is minimized. Thus maximum release time is attained by providing maximum ampere turn sensitivity.

A general expression for the relation between F and $\mathfrak{F}$ may be obtained by substituting in (13) the expression for $(\mathfrak{F}_C + \mathfrak{F})/\varphi$ given by (10). A much simpler relation is given by the linear approximation to the decreasing magnetization curve in which $(\mathfrak{F}_C + \mathfrak{F})/\varphi$ is taken as equal to $\mathfrak{R}_i$. This approximation has the same slope at φ_0 as the curve given by (3) or (10). The release pull usually corresponds to values of φ much nearer φ_0 than φ'', and the approximation is satisfactory in this range if φ_0/φ'' is small. Writing $(\mathfrak{F}_C + \mathfrak{F})/\mathfrak{R}_i$ for φ in (13), there is obtained the equation:

$$F = \frac{2\pi(NI + (NI)_C)^2}{\mathfrak{R}_i^2} \frac{d\mathfrak{R}_G}{dx}, \tag{16}$$

in which $(NI)_C$ is written for $\mathfrak{F}_C/(4\pi)$. This approximation is used for convenience and simplicity. The general expression, which is required for higher values of $\mathfrak{F}$, is obtained by substituting the right hand side of (10) for $\mathfrak{R}_i$ in (16) and in the expressions derived from it.

By comparison with (12), $\mathfrak{R}_i$ may be taken as equal to the gap and joint reluctance $\mathfrak{R}_E$ plus a modified and minor term for the iron reluctance. For the present purpose, it is convenient to take $\mathfrak{R}_i$ as the sum of the main gap reluctance $\mathfrak{R}_G$, which varies with x, and $\mathfrak{R}_F$, which includes the iron reluctance and the constant reluctances of the heel gap and of any joints in the magnetic structure. For a plane pole face, $\mathfrak{R}_G$ is given by x/A, where A is the effective pole face area, and (16) reduces to the equation:

$$F = \frac{2\pi(NI + (NI)_C)^2}{A\left(\mathfrak{R}_F + \frac{x}{A}\right)^2}.$$

As shown in one of the companion articles,[6] the pull F at travel x for a given value of NI is a maximum when the gap reluctance x/A is equal to the reluctance $\mathfrak{R}_F$ external to the gap. A similar relation applies for a domed pole face gap, as can be shown from the expressions given above. Taking the linear approximation to apply, $(\mathfrak{F}_C + \mathfrak{F})/(\mathfrak{R}_F + \mathfrak{R}_G)$ may be substituted for φ in (15). Writing $x/\mathfrak{R}_\infty$ for A in the resulting equation, this may be written in the form:

$$F = \frac{2\pi(NI + (NI)_C)^2}{x\mathfrak{R}_F} \Omega,$$

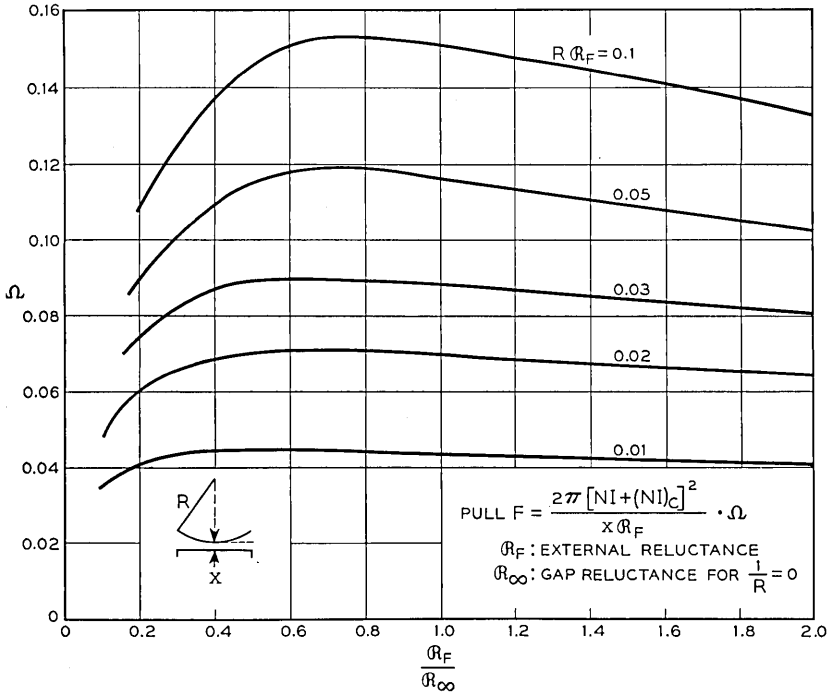


Fig. 11 — Ampere turn sensitivity of a domed pole face electromagnet.

where: (17)

$$\Omega = \frac{\Re_G^2 \Re_F}{\Re_\infty (\Re_G + \Re_F)^2} \frac{2\pi R \Re_\infty}{1 + 2\pi R \Re_\infty}.$$

The ratio Ω is a function of $\Re_F/\Re_\infty$ and of $R\Re_\infty$, which determines $\Re_G/\Re_\infty$, as shown by (14). Alternatively, Ω may be taken as a function of the ratios $\Re_F/\Re_\infty$ and $R\Re_F$, and Ω may be represented, as in Fig. 11, by a family of curves giving Ω versus $\Re_F/\Re_\infty$ for various values of $R\Re_F$.

From (17), maximum ampere turn sensitivity is attained by minimizing the separation x and the reluctance $\Re_F$ external to the gap. Assuming these to be made as small as engineering considerations permit, the pull for a given value of NI varies as Ω. With x and $\Re_F$ fixed, Ω now depends only on the dome radius R, to which $R\Re_F$ is now proportional, and on the projected area A, to which $\Re_F/\Re_\infty$ is now proportional. From the curves of Fig. 11 it is apparent that maximum ampere turn sensitivity is attained by using as large a value of dome radius as possible. For a given dome radius, there is an optimum value of $\Re_F/\Re_\infty$,

and hence an optimum value of A, corresponding to the maximum shown by each curve in Fig. 11.

AG Relay Pull

The AG relay[3] is a slow release relay with a domed pole face, to which the relations given above apply approximately. M. A. Logan and O. C. Worley developed a more exact expression for the pull in this case, in which an increment to the pull is given by the secondary pole faces on the side legs of the armature, which mate with the side legs of the E shaped core member. A further increment to the pull is given by the area at the main gap which lies outside the dome proper. Their analysis may be summarized in the notation used here by writing:

$$\mathcal{R}_i - \mathcal{R}_I = \mathcal{R}_S + \frac{\mathcal{R}_G \mathcal{R}_P}{\mathcal{R}_G + \mathcal{R}_P},$$

where R_I is the iron reluctance, $\mathcal{R}_S$ is the side gap reluctance, and the main gap reluctance is that of the domed gap $\mathcal{R}_G$ in parallel with that of the remaining area, $\mathcal{R}_P$. Substituting $R_i - \mathcal{R}_I$ for $\mathcal{R}_G$ in (16), there is obtained the following expression for the pull:

$$F = \frac{2\pi(NI + (NI)_C)^2}{\mathcal{R}_i^2} \left(\frac{d\mathcal{R}_S}{dx} + \left(\frac{\mathcal{R}_G}{\mathcal{R}_G + \mathcal{R}_P} \right)^2 \frac{d\mathcal{R}_P}{dx} + \left(\frac{\mathcal{R}_P}{\mathcal{R}_G + \mathcal{R}_P} \right)^2 \frac{d\mathcal{R}_G}{dx} \right). \tag{18}$$

An expression for $d\mathcal{R}_G/dx$ is given in Fig. 9, while $d\mathcal{R}_S/dx$ and $d\mathcal{R}_P/dx$ are given by $1/A_S$ and $1/A_P$, where A_S and A_P are the effective pole face areas for these two gaps. For the dimensions applying to the AG relay, these additional terms introduce minor but significant corrections to the values of F computed from (17).

Engineering Pull Data

In determining the requirements for slow release relays, the estimation of the range of variation in pull is a major problem. A procedure for such estimation developed by M. A. Logan makes effective use of the relation between F and $(NI + (NI)_C)$ in (16). This relation gives a linear plot of slope two when F is plotted against $(NI + (NI)_C)$ on logarithmic paper. A basic plot of this nature may be experimentally determined for a model having nominal values of coercive force H_C, to which $(NI)_C$ is proportional, and of finish thickness and side gap separa-

tion. The effect of known changes in these two latter factors may be determined by the observed vertical shift in this logarithmic plot, corresponding to changes in $\mathfrak{R}_i$ and $d\mathfrak{R}_G/dx$. The effect of a change in H_C is a change in the value of $(NI)_C$ to be subtracted from $(NI + (NI)_C)$ in determining NI. From these results there could be prepared curves giving the relation between F and NI for different combinations of the several variables for which allowance should be made.

5 EVALUATION OF CONDUCTANCE

The preceding sections have described procedures for estimating and measuring all the terms entering the expressions for the release time except the conductance G. This quantity is the sum of the sleeve conductance G_S and the eddy current conductance G_E. In some cases a short circuited winding may be used instead of a sleeve: in such cases G is the sum of G_E and the coil constant of the winding, G_C or N^2/R.

Sleeve Conductance

The conductance of a cylindrical sleeve is the sum of the conductances of the differential shells of length ℓ, radius r and thickness δr, as indi-

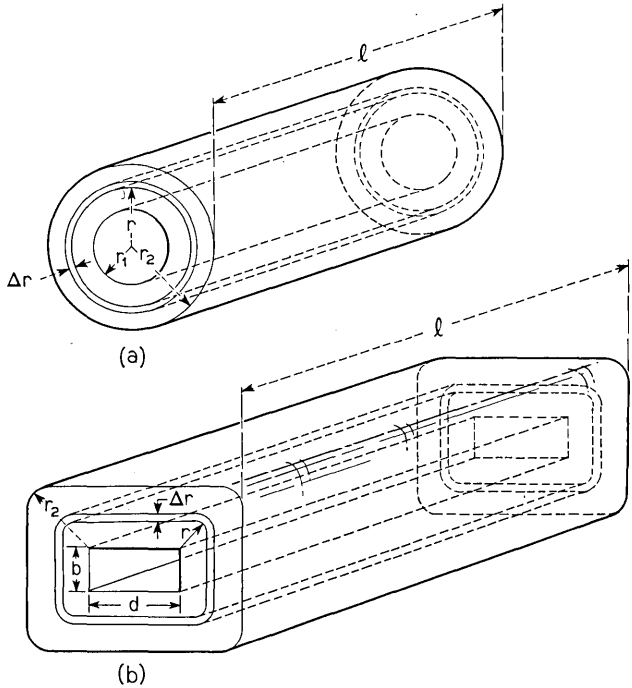


Fig. 12 — Sleeve conductance relations.

cated in Fig. 12(a). Thus the total conductance G_S is given by:

$$G_S = \int_{r_1}^{r_2} \frac{\ell dr}{2\pi\rho r},$$

where ρ is the resistivity of the material. On integration, there is obtained:

$$G_S = \frac{\ell}{2\pi\rho} \ln \frac{r_2}{r_1}. \tag{19}$$

The value of ρ for copper is 1.73×10^{-6} ohm-cm. If r_2 is twice r_1, for example, and $\ell = 5$ cm, the value of G_S given by (19) for a copper sleeve is 320,000 mhos.

In the case of a sleeve of rectangular section, as shown in Fig. 12(b), an approximation may be obtained by taking the sleeve as made up of shells with straight sides parallel to the center hole, connected by quarter circles. Then the perimeter of the shell at a distance r from the center hole is $2(b + d + \pi r)$. The total sleeve conductance is therefore given by:

$$G_S = \int_0^{r_2} \frac{\ell\, dr}{2(b + d + \pi r)\rho}.$$

in which r_2 is the wall thickness. On integration, there is obtained the equation:

$$G_S = \frac{\ell}{2\pi\rho} \ln \left(\frac{b + d + \pi r_2}{b + d} \right). \tag{20}$$

This expression is identical with that for the cylindrical sleeve, as given by (19), when the ratio r_2/r_1 of the radii is equal to the ratio of $(b + d + \pi r_2)$ to $(b + d)$.

Coil Conductance

For a cylindrical coil, the number of turns N is determined by the area of the coil section cut by a plane through the axis, or $\ell(r_2 - r_1)$, where ℓ is the length of the coil and r_1 and r_2 are its inner and outer radii. If a is the cross sectional area of the wire, and e the fraction of the coil space occupied by the conductor,

$$Na = e\ell(r_2 - r_1).$$

The mean length of turn is $\pi(r_2 + r_1)$, and the total length of conductor is N times this. Hence the resistance R is given by:

$$R = \frac{\rho N \pi (r_2 + r_1)}{a}.$$

The expression for N/R given by these equations gives the following expression for the coil constant N^2/R, or G_C :

$$G_C = \frac{e\ell}{\pi\rho} \frac{r_2 - r_1}{r_2 + r_1}. \tag{21}$$

For unit length of coil, values of G_C are shown in Fig. 13, plotted against r_2/r_1 for various values of e, together with the corresponding relation for G_S, as given by (19). This shows that the sleeve provides the maximum value of conductance for the space occupied. It also shows that the space used for a given value of sleeve conductance reduces the value of G_C that can be provided. When the full depth of the winding space is used by both coil and sleeve, called a slug in this case, each occupying part of the length, the value of G_C attainable is reduced in proportion to the length used by the slug. When the coil is outside the sleeve, the outer radius of the sleeve is the inner radius of the coil, and the value of G_C is reduced in proportion to the value of G_S. For a given value of e (given wire size and insulation) the value of G_C is fixed by the winding space and the value of G_S.

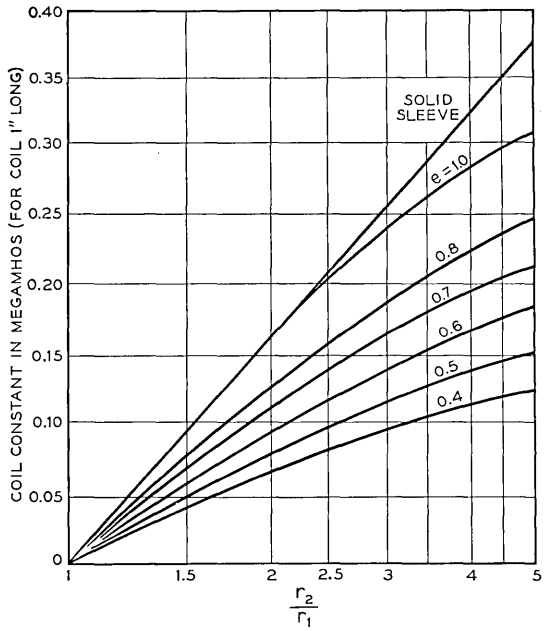


Fig. 13 — Relation between coil constant and coil dimensions.

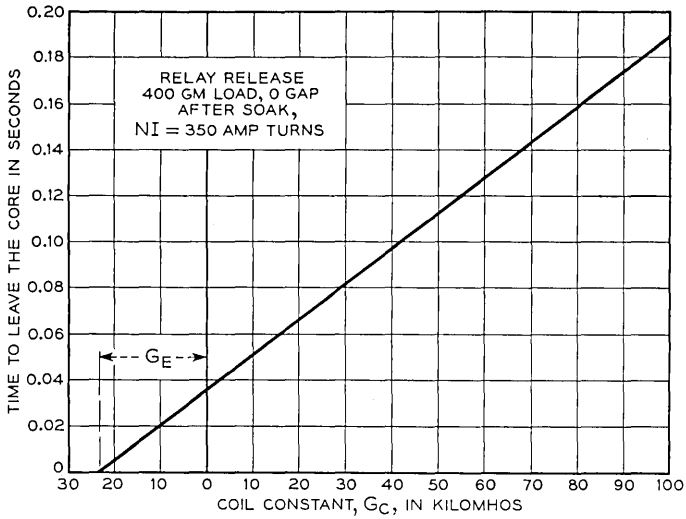


Fig. 14 — Experimental evaluation of eddy current conductance for release of relay.

Eddy Current Conductance

In one of the companion articles[4] it is shown that when the eddy current conductance G_E is a minor term in G, as with slow release relays, it is given by the equation:

$$G_E = \frac{\ell}{8\pi\rho}, \tag{22}$$

where ρ is the resistivity of the material, and ℓ is the length of the magnetic path. For iron, $\rho = 11 \times 10^{-6}$ ohm-cm, so for $\ell = 5$ cm, the value of G_E given by (21) is 17×10^3 mhos. The equation applies to a path of uniform circular cross section, so that the effective value of ℓ for most relay structures is intermediate between that of the core and that for the complete path.

The eddy current conductance of a specific model may be experimentally determined by measuring the release time with a winding shorted through an external resistance. A series of measurements are made in which this resistance is varied, while the initial current, which determines the "soak NI" value, is kept constant. The different values of the resistance correspond to different values of the coil constant G_C or N^2/R. From (6), the time t in this series of measurements varies directly as G, or $G_C + G_E$. Then a plot of t versus G_C, as illustrated in Fig. 14, is linear, and has a negative intercept giving the value of G_E.

Values of G_E thus determined agree with those estimated from (22) within the level of uncertainty as to the applicable value of ℓ in this equation.

6 DESIGN OF SLOW RELEASE RELAYS

The following discussion is confined to the bearing on design decisions of the performance relations developed in the preceding sections, and does not cover the manufacturing considerations involved. The development of a specific design depends on the initial choice between certain alternative features which require description.

Design Alternatives

The features considered here are: (1) the adjustment means, (2) the form of sleeve, (3) the criterion of adjustment.

The two methods of adjustment that have been used are (a) residual screw adjustment, (b) spring load adjustment. The former affects the release time by changing the reluctance and the residual flux, the latter by changing the flux or ampere turn value at which release occurs. The differences in the two methods relate more to the stability of the adjustment than to its ease or initial accuracy. Spring adjustment permits the use of the domed pole face, which is inherently stable except as the finish thickness may be affected by wear.

The two forms of sleeve are the interior sleeve, which uses the full length of the winding space, but only part of the depth, and the slug, which uses the full depth, but only part of the length. As shown in Section 5, the values of G_S and G_C, or coil constant N^2/R, attainable with a given winding space are independent of which arrangement is used. The slug provides some flexibility as to operate time, on which its retarding effect is a maximum when it is near the gap, and a minimum when it is away from it. It is subject to smaller temperature changes, with consequent changes in conductance, than the sleeve. The other differences between the two arrangements relate to manufacture, and to the costs of both coil and sleeve. The interior sleeve is used with the flat type relays of the Bell System.

The two different criteria of adjustment used are the release current, and the release time. The former is an indirect control, using a measurement of release ampere turns to determine the time that will result for the sleeve conductance used: the other is a direct measurement of the quantity to be controlled. In principle, the latter method would appear preferable, but its use is attended with several disadvantages.

The most important of these is the uncertainty as to the sleeve temperature that applies to any measurement made in central office maintenance, unless the relay to be tested is cut out of service for an hour or more before measurement. The conductivity of copper varies approximately as its absolute temperature, or, for engineering estimates, as $390 + T_F$, where T_F is the temperature in degrees Fahrenheit. Coil temperatures of 225°F are permitted in normal relay operation. As the time varies as the sleeve conductance, a relay with its sleeve at this temperature would have a release time in the ratio 470/615 or 76/100 to the rated release time for 80°F. Allowance for variation in this range is made in circuit design, but a corresponding uncertainty as to the condition applying in adjustment would effectively double this variation. Current flow adjustment is free of this difficulty, and the variations in the correlation of release time with release ampere turns are less than those resulting from the temperature uncertainty in any convenient procedure for timing measurements. Current flow adjustment has the further advantage of using equipment that is employed for other relays in central office maintenance. It is the more commonly used criterion of adjustment for Bell System relays.

Operate Considerations

A slow release relay must not only provide the desired release performance: it must also operate its load. The operate pull characteristics are similar to those of other relays of the same general type, as the domed pole face, in particular, gives nearly the same pull at an open gap as a plane pole face of the same total area. Thus the pull characteristics of the *AG* relay are similar to those of the *AJ* relay[3] for the same travel. The sleeve retards the flux development, and makes operation slower than that for the same coil input without the sleeve. In most applications of slow release relays, this has little or no effect on circuit operation. Faster operation can be obtained by increasing the steady state power applied, but this is limited by heating considerations. The large part of the winding space used for the sleeve limits the operate sensitivity of slow release relays, and increases the power required for a given load. The load for a given relay design determines a minimum ampere turn value for operation, and this is related to the steady state power by the identity: $(NI)^2 = I^2R\ N^2/R$. As the coil constant N^2/R, or G_C, is determined by the available winding space available for the coil, the power requirements of slow release relays are higher than those of similar relays having the full winding space available for the coil.

Optimum Design Conditions

In considering the design features that are favorable to slow release operation it is convenient to refer again to the expression for the release time given by equation (9):

$$t = \frac{4\pi G(\varphi'' - \varphi_0)}{4\pi NI} f(z), \tag{9}$$

where $f(z)$ is the function shown in Fig. 5 to be nearly a constant in the range of interest. It is also convenient to refer to the expression for the pull given by the equation (16):

$$F = \frac{2\pi(NI + (NI)_C)^2}{\mathfrak{R}_i^2} \frac{d\mathfrak{R}_G}{dx}.$$

Together with reference to the operate requirements, these two equations indicate the characteristics that are important for slow release operation.

The winding space determines the value of G attainable, as shown by the relations of Section 5. It limits the combined values of G_S and G_C, of which the latter controls the operate power sensitivity, while the former, from (9), determines the release time. Thus both the operate sensitivity and the release time attainable vary directly with the winding space and hence with the over-all size of the relay.

From (9), the attainable release time is nearly proportional to φ'', and hence to the cross section of the core, assuming φ_0/φ'' to be small. The external core dimension is the internal dimension of the sleeve, so that increases in φ'' are offset by decreases in G if this dimension alone is varied. In any case, sufficient section must be provided for φ'' to have a margin over the field required to operate the maximum load.

In telephone use, the slow release relays are a minority group in a relay population which must, for maximum economy in manufacture and use, have common overall dimensions and as few differences as are consistent with the requirements of specific uses. Thus the core section, winding space, and over-all dimensions reflect an optimum choice for the whole relay population, and not for the slow release relays alone. The latter are distinguished by as few special features as are essential to their special function. In the AG relay these are: the armature, the heat treatment of the magnetic parts, the sleeve and coil, and a buffer spring for load adjustment.

The material and its heat treatment determine the iron reluctance and the coercive mmf, $4\pi(NI)_C$, of which the latter is the more important quantity. It enters the timing relation indirectly in the minor term φ_0,

which varies directly with $(NI)_c$, and enters the pull relation (16) directly as a major term. If high coercive material were used to increase $(NI)_c$ and thus reduce the release ampere turns NI, the effect of the latter on the time would be offset by the increase in φ_0. The latter varies also with the reluctance $\mathcal{R}_i$, and if φ_0/φ'' is not small, the time for a given value of NI is no longer independent of the reluctance variation, and the release ampere turn value is unsatisfactory as a criterion of adjustment. Thus a small stable value of $(NI)_c$, as given by low coercive material, is preferable when the release current is used as the criterion of adjustment.

The pole face dimensions are determined by the need for a low stable value of gap reluctance. As shown in Section 4, maximum release sensitivity is obtained by using as large a dome radius as will assure a consistent configuration, together with a projected area optimum with respect to the reluctance in series with the gap. Fig. 11 shows this optimum to be broad, allowing considerable latitude in the choice of the projected area with reference to other design considerations.

The finish applied to the magnetic parts at the pole face is critical with respect to wear and stability, as well as to the danger of mechanical adhesion which would affect the release load. A nickel chrome finish is used in the Y and AG relays. The effective air gap thus introduced is the dimension x of Figs. 9 and 10. As these figures show, this has a relatively large effect on the pull relation, and any change in this dimension as a result of wear tends to increase the pull and the release time. The analysis of Section 4 shows that maximum sensitivity is attained with as small a finish separation as can be provided, which is fortunately in agreement with the manufacturing considerations for the finish.

The adjustment means used with a fixed pole face, e.g., the dome type, is some form of spring adjustment for controlling the operated load. The normal operated load of a relay is the product of the contact force and the number of contacts, plus the back tension and the force required for spring flexure. Adjustment of the back tension is used, but is limited by the requirements for operation. Hence a buffer spring is employed in AG relays, which is picked up in the last few mils of armature travel and affords control of the operated load. This spring is adjusted by bending, and a tolerance range of 50 gm wt is allowed in this adjustment.

Performance Variations

Performance variations fall into two categories: the initial differences among individual relays that are nominally identical, and the further differences that develop in service as a result of wear and other changes.

In the latter category, the changes may be magnetic or mechanical. The mechanical changes appear in the value of the operated load, and are practically negligible in relays of the AG type, because of the lift-off actuation used.[3] In this they compare favorably with Y-type relays, which are liable to load changes through stud and contact wear. The magnetic changes include "ageing" of the magnetic material, a gradual change in permeability and coercive force which occurs with magnetic iron. Hydrogen annealed silicon steel, used in the AG relay, is free from this defect. The other magnetic changes are those appearing in the pole face configuration and finish thickness as a result of wear. No significant changes of this type have been observed in the AG relay, which is therefore highly stable in maintaining the performance of its initial adjustment.

In use, the release time varies inversely with the sleeve or slug temperature. As shown above, the 140°F temperature rise allowed for coils in normal operation will, if experienced by the sleeve, reduce the release time by about 25 per cent. Allowance for variation in this range must be made in planning the applications of slow release relays. Relays fitted with slugs are less affected by coil temperatures than are those using interior sleeves. The slug arrangement, with an insulating air gap between slug and coil, would be essential to a slow release relay designed for closely controlled release time.

The initial differences from nominal performance may be sub-divided into those arising from the adjustment tolerance, and those resulting from using release current as the adjustment criterion. Only the former would apply if the direct timing criterion were used, provided a procedure was employed in which the sleeve temperature was fixed in measurement. A tolerance range of 50 gm wt is a 25 per cent variation for a representative load of 200 gm wt, and corresponds, from (16) to about a 12 per cent variation in $NI + (NI)_C$. If $(NI)_C$ is about equal to NI, the variation in the latter is about 25 per cent, which results, from (9) in a similar variation in the release time. A smaller tolerance range could be attained by more expensive adjustment means than the buffer spring of the AG relay, with a corresponding reduction in the range of time variation.

With release current as the criterion, the value of NI in (9) is subject to the adjustment variation, of the order of 25 per cent for the case cited above. In addition, the release time is subject to variations in G, $\varphi'' - \varphi_0$, and $f(z)$. In the optimum range of operation where $f(z)$ is nearly constant, variations in this term are minor. φ_0 is subject to the relatively large percentage variations in $(NI)_C$ and $\Re_i$, but the net effect

of these is small if φ_0/φ'' is as small as 10 per cent, as in the AG relay. Hence the principle variations in the release time are those associated with the variations in the sleeve dimensions and the core cross section, which determine the values of G and φ'' respectively. They cause variations in the AG relay of from 20 to 30 per cent. Combined with the variation in NI resulting from the adjustment tolerance, they result in release time variations of 40 or 50 per cent.

For the great majority of slow release relay applications, much larger variations have no effect on circuit performance, and for only a small minority is it desirable to employ requirements assuring this measure of control. It is apparent from the preceding discussion that changes in construction and adjustment procedure might be made which would result in closer time control: as they would add to the costs of both manufacture and use, they are not economically justified.

Engineering Data

A specific relay design is termed a code: all relays of the same code are nominally identical. A code is distinguished from other codes of the same relay type by the contacts and coil provided, and by the sleeve or other special features used. To design slow release relay codes, and to specify their adjustment requirements and the performance provided, requires means for determining from the general pull and timing characteristics the values applying to specific cases. For this purpose the pull and timing data are prepared in the form illustrated in Fig. 15.

The pull information appears in this chart in the form of the two curves marked "hold pull" and "release pull." These show the relation between operated load and release ampere turns for the two limiting cases of minimum and maximum pull, and reflect the variations in the terms of (16). No individual relay releases on ampere turns above the hold values, all relays release on ampere turns below the release values.

The timing information appears in the form of a pair of curves for each sleeve size that is used: one giving the minimum time, and one the maximum. Only one such pair is included in Fig. 15. The difference between the two curves represents the 20 to 30 per cent variation associated with the variations in G and $\varphi'' - \varphi_0$.

The information on a chart of this type is used in conjunction with engineering data for the operated loads of the various spring combinations, and the changes in these loads that can be provided by back tension and buffer spring adjustment.

For each spring combination, there is a lower limit to the adjusted

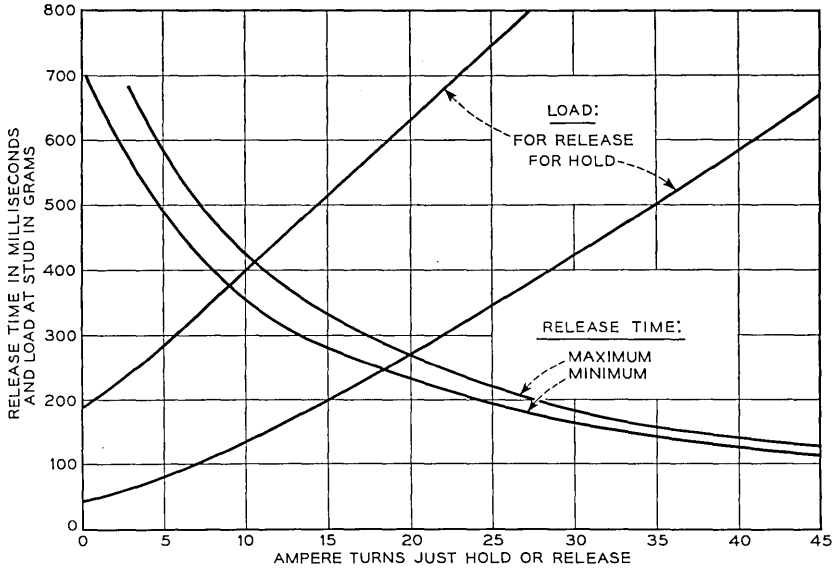


Fig. 15 — Engineering data for release time estimation.

load attainable in all cases. This corresponds to the case in which the contact force, which is not subject to adjustment, has its maximum value. The NI value read from the hold curve for this load is the lowest that can be specified as a "hold" requirement. Each individual relay has a release F versus NI characteristic intermediate between the release and hold capability curves. To meet the hold requirement its load must be adjusted, and this adjustment is subject to the tolerance cited above. The minimum load is set by the hold requirement, and the maximum load is set by a release requirement of a lower NI value, the difference between the two NI requirements corresponding to the tolerance range in load adjustment.

Thus adjustment values are determined in the form of limits to the ampere turn value at which release occurs: the lower limit is the release value, the upper the hold. The release time limits can then be read from the timing curves. The maximum time is read from the maximum curve at the release ampere turn value. The minimum time is read from the minimum curve at the hold ampere turn value. This minimum time, when determined for the largest sleeve, is the longest time that can be guaranteed for the load in question, and is subject to reduction in service by the temperature variation previously discussed. When a shorter release time is desired, a smaller sleeve may be used, or a higher hold

value specified, subject to the capacity of the adjustment springs to supply the necessary increase in load.

In the common case for which only the minimum time is of circuit importance, the release value is chosen without reference to the hold value, solely for the purpose of assuring that the relay will not lock up indefinitely. This procedure takes advantage of the simpler requirements of this case by widening the adjustment tolerance and reducing adjustment effort.

7 CONCLUSIONS

The relation between the release time of slow release relays and the design parameters can be more accurately expressed in analytical form than the other time characteristics of relays. These analytical relations, as presented in this article, can be used for the estimation of release time, and particularly for the determination of the effect on this time of variations in the design parameters. The need for a low reluctance magnetic circuit makes the performance of slow release relays highly sensitive to dimensional and material variations, and adjustment is required to assure the timing limits required in their use. Such use usually permits a wide spread in release time, provided a minimum value is assured. Advantage is taken of this in providing slow release relays which perform their function at a minimum cost in manufacture and use, materially lower than that for the construction and adjustment practices which would be required for closer timing control.

ACKNOWLEDGEMENTS

The specific references made to individuals and prior studies do not include all the work which has been drawn on in the preparation of this article. In particular, much of the discussion of the applications of the analysis is based on work carried out by M. A. Logan, Mrs. K. R. Randall, O. C. Worley and others in the development of the *AG* relay.

REFERENCES

1. F. A. Zupa, The Y-type Relay, Bell Lab. Record, **16,** p. 310, May, 1938.
2. H. N. Wagar, The U-type Relay, Bell Lab. Record, **16,** p. 300, May, 1938.
3. A. C. Keller, A New General Purpose Relay for Telephone Switching Systems, B.S.T.J. **31,** p. 1023, Nov., 1952.
4. R. L. Peek, Jr., and M. A. Logan, Estimation and Control of Operate Time of Relays, pages 109 and 144 of this issue.
5. H. N. Wagar, Slow Acting Relays, Bell Lab. Record, **26,** p. 161, April, 1948.
6. R. L. Peek, Jr., and H. N. Wagar, Magnetic Design of Relays, page 23 of this issue.

Economics of Telephone Relay Applications

By H. N. WAGAR

(Manuscript received July 6, 1953)

Today's telephone central office is largely built around the telephone relay. As each office may use some 60,000 of them, their performance characteristics, as well as their first cost, have a very important influence on the cost of the office. By properly balancing such factors economically, the lowest cost office can be planned.

This paper shows how performance features and design factors may all be expressed as "equivalent first cost," and so be related to manufacturing cost itself. The influence of lot-size on manufacturing cost is considered including a determination of optimum lot-sizes, to aid the relay designer in deciding on standardized models.

With a basic performance language formulated — equivalent first cost — optimum relations among the variables may be established. Such results are given for (a) optimum coil-plus-power costs, (b) optimum power-plus-speed costs, (c) optimum number of relay codes, and other related problems. Methods are also given to evaluate how serious the effect may be when optimum conditions are not satisfied.

Were it not for the application of these methods, central office costs would be much higher.

INTRODUCTION

The telephone central office of today is in part an enormous computing machine which, upon instructions from the customer, given by his dial, calculates how to find the called party; then the remaining mechanism completes the call. This machine is composed largely of interconnected relays, used in enormous quantities. Every dial call in a large city involves around 1000 relays. A typical large office contains more than 60,000 of them; in fact, they are used so extensively that the Western Electric Company manufactures about ten of them for every new subscriber, and their output is figured in tens of millions per year. It is not

surprising that relay use has an enormous influence on the cost of the central office — not just because of relay purchase cost but also the relay's influence on other office factors. For example, the size of the power plant and the total number of equipments for common control, which depend on the functioning time of the relays, are decided by the relay characteristics. In the application of relays, then, it may well prove that the largest economies can be realized by spending a little more for each relay in the beginning in order to save still more in the cost of the power plant, common control equipment, size of the building, and so forth. It has been found that attention to ways of optimizing the application costs for relays can lead to an appreciably lower cost central office, and this paper will illustrate a few cases of how such a problem is approached. Because the telephone system is so large, the influence of each relay, taken in total, is also large. Fortunately, at Bell Telephone Laboratories, it has been possible to take the over-all view of the subject through familiarity with all phases of the relay application problem; and large economies which will eventually benefit the customer are resulting.

The basic problem to be discussed is how best to realize maximum economy of the central office so far as relays and their uses are concerned. As in most engineering problems, it is necessary to evaluate and compromise between oppositely varying cost effects of such things as efficiency, manufacturing cost, amount of equipment, ease of mounting and wiring, maintenance in all its aspects, and the like. The end result of each of these variables is its economic effect on the office as a whole, and they can only be compared if their values can be stated on a comparable basis. It has been found that each such effect may be considered as an incremental cost over and above a reference cost for its particular ideal condition. If properly chosen so as to be independent, then all such incremental costs may be added. They then represent the net "cost penalty," compared to the design with all ideal conditions taken together, and describe the merit of the design. They can also be used to find the optimum design. The methods apply generally to many other similar problems.

BREAKDOWN OF THE PROBLEM

Consider first how relays are applied in the telephone switching system. To the greatest extent possible a basic relay structure is chosen, carefully planned for low maintenance effort, all of whose basic parts can be made by mass production methods. On this basic framework one

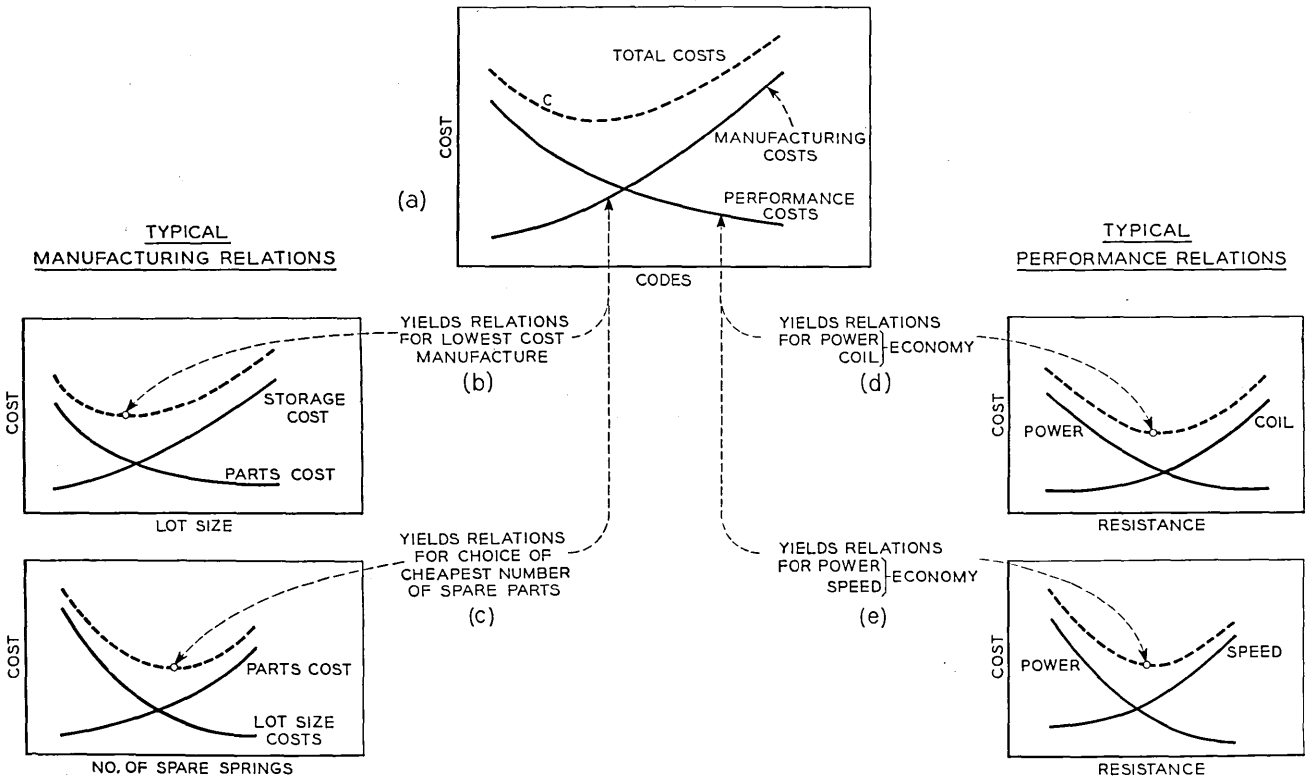


Fig. 1 — Simplified view of procedure for gaining maximum system economy.

may apply (a) suitable windings to satisfy the various conditions imposed by a need for sensitivity, speed, lowest possible first cost, or some other of the many specialized requirements to be encountered; and (b) particular arrangements of contact springs which will provide for the desired functions in the circuit ranging from a single "make" up to complicated sequences such as 12 sets of "transfers." Every different combination of the basic parts is termed a "code," meaning another *kind* of relay built up from the parts common to the basic *type* in question. If an increasingly large number of such codes is encouraged, each tailored to some special circuit function, then the advantages of mass production begin to be lost as the total demand is split into more and more subdivisions, each with comparatively low demand. Each additional *kind* of relay is a problem in "coding economics," and can lead to excessive telephone office costs unless it is properly considered.

Satisfactory solutions to this coding problem in turn depend on a detailed knowledge of all the factors governing either the first cost of the relays, or the first cost of other features of the system, as controlled by the relays. When all such effects are properly stated, they can be compared and brought into economic balance. Though such studies soon branch out into many fields, the enormous dollar savings are strong incentives to do whatever work is needed.

Best economic balance in the system is formed through a series of optimizing steps, generally illustrated by Fig. 1. Graph (a) shows how the total cost of an office may vary with number of codes. The performance part of the cost diminishes with increasing codes because each code will more nearly satisfy the precise circuit needs. But the manufacturing costs of the relays will rise with increasing codes because the lot-sizes grow ever smaller and hence more costly. The lowest point on the summation curve represents a desirable goal. The two base curves of (a) in turn are built from detailed facts about parts costs, coil costs, power plant and equipment costs, and many other similar data. The curve of manufacturing cost, for instance, is built from graphs like (b) and (c) which tell how factory costs will vary. The curve of performance cost is built from graphs like (d) and (e) which tell how office costs will vary for things like power and speed. In each such case, important individual economies result by following similar optimizing steps. Procedures for the practical application of these ideas have been of particular help in recent redesigns of the No. 5 crossbar system to use the newly developed wire spring relay family.[1]

[1] A. C. Keller, A New General Purpose Relay for Telephone Switching Systems, B. S. T. J., **31**, pp. 1023–1067, Nov., 1952.

In carrying out an actual study, a point of reference is needed first. This involves (a) fixing a standard against which all costs will be compared, and (b) expressing the value of all features in a common language. Once this has been done, every factor may be evaluated as an incremental "cost penalty," over and above the reference standard. Finally, in making comparisons, the reference standard will always subtract out, so that one needs only to add all incremental cost penalties to get the total cost of the design changes in mind.

The evaluation of all variables on a common basis is covered in Part I, which considers various manufacturing costs, power costs, and the cost of functioning time.

The remaining parts of this paper then develop relations for maximum economy in the switching system, for some important cases that arise in practice:

Coil design for maximum power economy.

Economical number of occasionally unused extra parts.

Economical adjustment for speed relay.

Coil design for maximum combined power and speed economy.

So far as possible, results are given in general form, permitting one to work from charts when considering specific application problems. Of further benefit are charts which permit the designer to decide the economic disadvantage of a design which may depart from optimum.

Design for systems economy through relay design includes many other important topics not considered here, as for example: how to make best use of molding, welding and other manufacturing processes; or how to design for long life, reliability, and other maintenance concerns. This article covers only some typical problems for which optimizing methods are readily applied.

Part I — Expressing Systems Performance and Manufacturing Variables in a Common Language

In the over-all switching system, one must evaluate various effects whose costs seem quite different, and a decision must be made as to what design course to follow, no matter how varied the conditions may be. Many of the most important such cases can be handled quite easily by a process of converting actual cost in the telephone plant back to an equivalent cost in terms of the factory production cost of each individual relay. Once all such costs are so stated, it is possible to examine the incremental effect of each change, and confidently draw conclusions. For example, if it can be stated that a change in a relay coil will save an amount of power that is worth 50 cents per relay equivalent first cost,

then there is no doubt about designing such a new coil if its manufacturing cost will increase by only 20 cents. So the relay designer needs to understand how various factors, from original manufacture to final application, may be stated on a cost basis that is comparable throughout. Among the most important special cases are the following:

1. Manufacturing cost of winding a coil,
2. Manufacturing cost as a function of annual demand,
3. Equivalent manufacturing cost of power consumed by relays,
4. Equivalent manufacturing cost of functioning time of a relay.

The method of evaluating each on the same basis will now be briefly outlined. In each case, it should be noted that since only comparisons between variable portions of the system are considered, only incremental values need be considered.

1.1 THE COST OF A WINDING

In comparisons between coils, certain common operations, such as soldering the leads, using and cementing spoolheads, etc., will always subtract out, leaving only the difference due to varying amount of copper, which is paid for by the pound, (or per ohm), and the number of turns, which depends on the speed of the winding machine and number of coils wound at one time. Thus, winding costs, which depend only on the electrical design, may be considered as varying only with the cost per ohm and cost per turn, as given in equation (1):

$$C_W = c_R R + c_N N, \tag{1}$$

where C_W = total cost of the variable part of the winding,
c_R = cost per ohm, which may be tabulated by wire size,
c_N = cost per turn of winding, also given in tables by wire size,
R = resistance of coil, ohms,
N = number of turns in winding.

Typical values of c_R and c_N for a particular kind of wire and coil are shown in Fig. 2.

1.2 COST AS A FUNCTION OF ANNUAL DEMAND

Various components of relays, as well as their assembly routines, occur in several variants of a general basic pattern. In the course of their manufacture it is necessary to stop the manufacture of one unit, reset the machine, and proceed with manufacture of another. Whenever this happens, production stops; and work must be done to reset the machine, all of which may be evaluated as a set-up cost (designated S). As more

and more variants are required in the over-all general structure, the number of set-ups will increase to the point where the economical manufacture of the part is seriously penalized. Thus a point will be reached where it is advantageous to make more parts than are immediately needed, and store the remainder, so as to avoid excessive cost of shut-downs. When the variable effects of (a) set-up costs and (b) inventory costs are fully considered, the ideal quantity to manufacture in one lot may then be found by comparing the results of (a) and (b), as will now be shown. The costs which correspond to these ideal quantities will immediately follow.

1.21 *Lot-Size Costs*

In manufacturing practice there are two outstanding expenses which may affect the cost of any particular item which is classed as a code, or kind, of the basic family of articles. These are the administration costs, and the set-up costs.

1.211 *Administration Costs*

For certain kinds of operations there is a considerable amount of paper-work, drafting, checking, etc. to maintain in normal up-to-date condition. Occasionally, this cost is enough to influence the cost of the

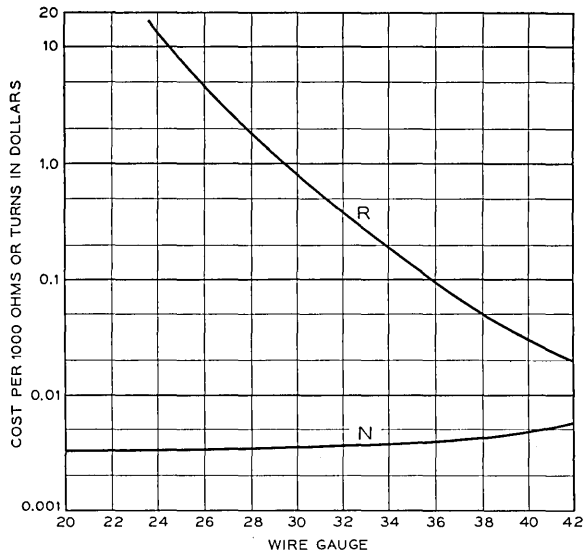


Fig. 2 — Typical winding costs.

item. The resulting individual effect may then be stated as

$$C_A = \frac{A}{n}, \tag{2}$$

where C_A = cost penalty per unit due to administration,
A = annual administration cost for maintaining one code in good standing,
n = annual demand for a given code.

1.212 *Set-up Costs*

If only one kind of part or assembly were needed, it could be continuously built in the same way, with no time lost for changeover to other parts, no particular bookkeeping necessary to control the proper flow of differing parts, and with more mechanized action. Usually this condition is far from realized in practice; nevertheless it represents the peak of manufacturing economy, and may be taken as a standard of reference for comparing all other less favorable conditions. The cost incurred, per item, due to its lot size, as against its cost if there were always but one lot, will be called the "lot-size cost penalty." The manufacturing lot-size cost penalty per part for any particular process, then, may be stated as follows:

$$C_L = \frac{S}{L}, \tag{3}$$

$$= \frac{S\ell}{n}, \tag{4}$$

where C_L = lot size cost per item,
S = cost of one "set-up",
L = size of lot for which one set-up is made,
$= n/\ell$,
n = annual demand for a given code,
ℓ = number of lots per year, or lot-frequency.

Equation (4) gives the cost over and above single-kind manufacture, so far as lot-size is concerned. Hence, when values for each of these variables can be established, the lot-size cost penalty for any particular process can be determined.

1.213 *Over-all Lot-Size Costs, Related to Annual Demand*

Combining the results of equations (2) and (4), the total cost penalty per part due to lot-size is

$$C_L = \frac{A}{n} + \frac{S\ell}{n}. \tag{5}$$

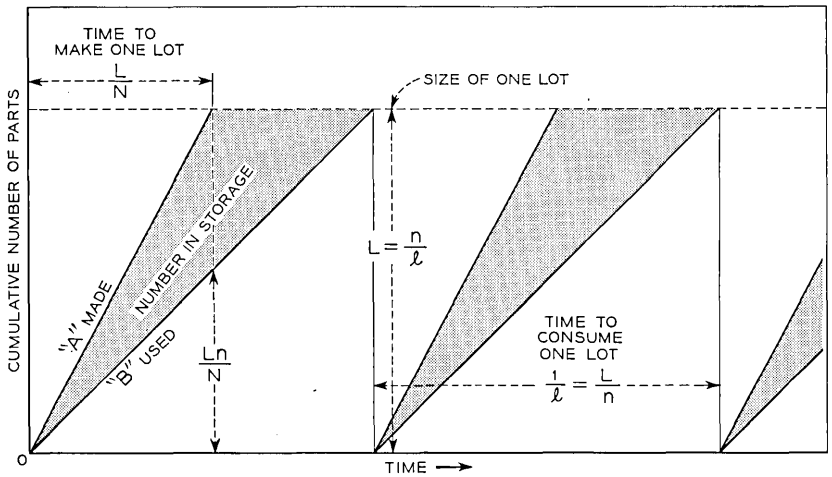


Fig. 3 — Lot-size cost penalty relationships.

1.22 *Inventory Costs*

When more parts are made than are immediately needed, it becomes necessary to store them until they are all used, when the process will be repeated. The resulting cost penalty, compared to no storage at all, may be stated analytically with the help of Fig. 3, which is a plot of production as a function of time. Line A represents the number of parts *made* in a lot as a function of time, while line B represents the parts *used* in the next manufacturing stage, as a function of time. The difference between the two lines at any time represents the number in storage, from which their value and hence the interest charges may be found.

The cost penalty per part due to storage is:

$$C = \frac{\text{Annual interest factor} \times \text{value of one part} \times \text{av. no. of parts in storage}}{\text{Annual production of this part}}. \quad (6)$$

The following steps can be taken to supply values for this equation. From the figure, the maximum quantity of parts stored equals

$$L - \frac{Ln}{N},$$

or

$$\frac{n}{\ell}\left(1 - \frac{n}{N}\right).$$

The average number of parts stored is one-half this, or

$$\frac{n}{2\ell}\left(1 - \frac{n}{N}\right), \tag{7}$$

where N = rate of output of machine per year.
The penalty per part due to lot-size and administration was given in (5), and when added to its base cost C_1 gives the total value of one part as

$$C_1 + \frac{S\ell}{n} + \frac{A}{n}. \tag{8}$$

The penalty per part, due to storage, equation (6), is then the product of k/n times (7) times (8), where k = interest charges on stored parts expressed as a ratio. Then the storage cost penalty, C_S, is

$$C_S = \frac{k}{n}\left(C_1 + \frac{S\ell}{n} + \frac{A}{n}\right)\frac{n}{2\ell}\left(1 - \frac{n}{N}\right). \tag{9}$$

1.23 *Total Cost Penalty*

The entire penalty resulting from lot size charges and inventory charges may now be stated as a function of annual demand by summing equations (5) and (9). Writing q for $(1 - n/N)$, and rearranging terms, the total cost penalty in terms of annual demand is

$$C_P = \frac{kq}{2\ell}\left(C_1 + \frac{A}{n}\right) + \frac{kqS}{2n} + \frac{1}{n}(S\ell + A). \tag{10}$$

Equation (10) is the basic relation between costs and annual demand, and is of the general shape shown in graph (b) of Fig. 1.

It is of greatest interest to determine conditions where this curve has its lowest value, which occurs when $dC/d\ell = 0$. The result is

$$\ell_0 = \sqrt{\frac{kqn}{2S}\left(C_1 + \frac{A}{n}\right)}. \tag{11}$$

which gives the lot-frequency, ℓ_0, for which total cost penalty is a minimum. When ℓ_0 is substituted for ℓ in equation (10), the resulting optimum cost, as a function of annual demand, C_0, is given:

$$C_0 = \sqrt{\frac{2kqS(C_1 + A/n)}{n}} + \frac{kqS}{2n} + \frac{A}{n}. \tag{12}$$

Equations (11) and (12) furnish the means for deciding how to plan manufacture under differing conditions of annual demand, and how much the product is penalized even after planning is carried out as

efficiently as possible. Two illustrations will be given: for a comparatively low-cost part involving expensive set-up charges and no administrative charges, and for a more costly operation with comparatively inexpensive set-up costs but high administrative costs. For Case I assume a part for which the following constants apply:

$$k = 0.15 \qquad C_1 = \$0.10$$
$$N = 2{,}000{,}000 \qquad A = 0$$
$$S = \$10.00$$

In Cast II assume that

$$k = 0.15 \qquad C_1 = \$1.00$$
$$N = 2{,}000{,}000 \qquad A = \$100$$
$$S = \$2.00$$

The results for these two cases are given in Fig. 4.

Results such as these are of the utmost value both in manufacturing planning for lowest cost, and in applications planning to show when the lot-size penalties can be tolerated. However, no case is to be expected in practice where optimum conditions can be met more than a part of the time. What with unexpected rush orders, fluctuation of annual demands, rescheduling, possible moves, or breakdown of machinery, it is usually impossible to plan production to remain at all times at the optimum value. The variation from optimum cost that results from a departure from optimum lot size is thus of interest; it is readily stated

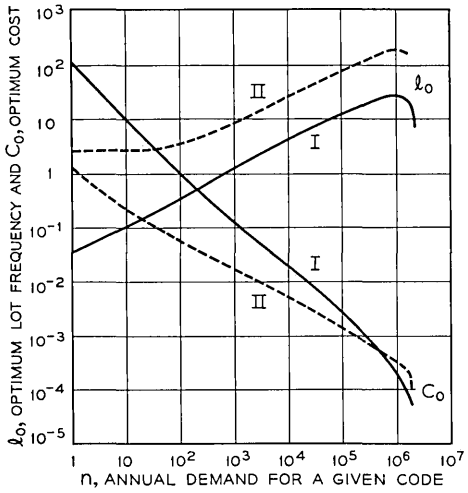


Fig. 4 — Optimum lot sizes and costs.

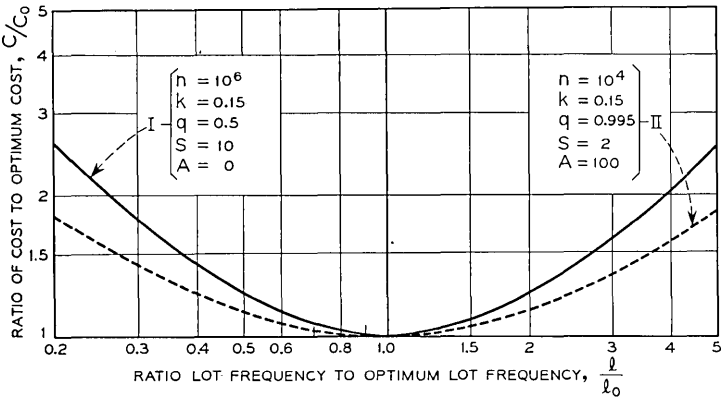


Fig. 5 — Cost variations with variations from optimum.

by rewriting equation (10) in terms of the values of ℓ_0 and C_0 just found in equations (11) and (12). The result is

$$C/C_0 = \frac{\ell_0/\ell + \ell/\ell_0 + D/\ell_0}{2 + D/\ell_0}, \tag{13}$$

where

$$D = \frac{kq}{2} + \frac{A}{S}.$$

This relationship is seen to depend on annual demands, and also administrative and set-up costs, which will be peculiar to each case. Therefore, it will be necessary to use the equation with appropriate values of these constants for each case that arises. Typical values for the two previous illustrations are shown in Fig. 5. Inspection of these curves shows that the cost penalty of departing from optimum lot-frequency is somewhat variable depending on the ideal lot-frequency for the particular case. However, the higher values of lot-frequency, which produce the highest penalties, are those most easily under control of the factory, and variations averaging in excess of 2 or 3 to one from ideal are not at all likely. When the ideal lot-frequency is low, manufacturing control is not so easy, but the cost penalty ratio of departing from the ideal is far less. From inspection of these curves, then, the value of C/C_0 must be judged by the conditions of each problem; a value of cost penalty amounting to about 1.25 C_0 is often found to be a quite satisfactory value to assume.

Thus, (a) a method is available to the manufacturing engineer to help plan optimum lot size, and (b) a method is available to the relay

and circuit designer for deciding how much his designs may be penalized by varying the quantities used. It consists of using the costs determined from equation (12), and modified by an amount indicated by (13), or in ordinary cases by arbitrarily increasing (12) by the factor 1.25.

1.3 THE EQUIVALENT MANUFACTURING COST OF POWER CONSUMED BY A RELAY

The largest part of the power plant used in telephone central offices is the 50-volt equipment needed for the switching apparatus. Because of its special voltage range and the requirements on its stability and reliability, it must be provided as a part of every central office to convert the normal power company voltages into the desired telephone values. This equipment involves generators, banks of storage batteries, and switchboard, bus-bar and control equipment; it is an expensive portion of every central office which, of course, it is desirable to minimize as far as possible. For every relay required in the central office, one must associate a small "chunk" of this power plant; so that each relay needed implies an associated investment in the power plant. In a later section the means for minimizing the power and relay costs are described, and since they will involve comparisons on the same basis, it is first necessary to state the value of the power plant in terms related to the amount of power consumed by any individual unit. The method of evaluating the power plant is given in this section.

The problem may be broken into two parts: (a) the equivalent first costs assignable to the power plant equipment, and (b) the equivalent first costs of the charges per kilowatt-hour paid to the power company. Then the figures are combined.

1.31 *The Power Plant*

Power plants furnished for central offices vary over a wide range of size and cost, depending on the size of the office and its estimated activity. This plant must of course be planned so as to carry the load during the period of peak activity, even though this occurs for only a small part of the time, so that in the long run the cost of the power plant is decided by the share of busy-hour power taken by each relay. To get such costs one must first find the cost of the plant as a function of total power requirements.

Present practice in power plant planning results in the purchase of basic power units which may be combined to cover certain ranges of power supplied over roughly a two-to-one range. Within this range, by

adding generators and batteries in parallel the needed capacity can be attained. Beyond the range, basically heavier machines are needed. The resulting installed price of such power plants has the character shown in Fig. 6. Since prices and installation charges change from year to year, the values shown here are given in qualitative form as illustrations — they must of course be evaluated as concisely as possible when considering new central office designs. As seen in the figure, the power plant price varies in two ways: in fairly big jumps as basic plant size is changed, and in smaller but rather uniform manner as the basic size varies within its lower and upper limits.

From Fig. 6 it is seen that for each basic size of plant the rate of increase in price with small increases in capacity is about the same for all plants. The average rate for this incremental power capacity, corresponding to the expense if relatively small power changes are made in any one basic type of power plant will be called a. As the power requirements of the office change beyond a certain point, it is necessary to install basically larger or smaller equipment with correspondingly different base prices. At the points of maximum capacity, there is an abrupt change in the basic price of the plant, amounting often to many thousands of dollars. These abrupt increments, designated A, represent a lump sum of money that can be saved whenever the power plant size can be re-

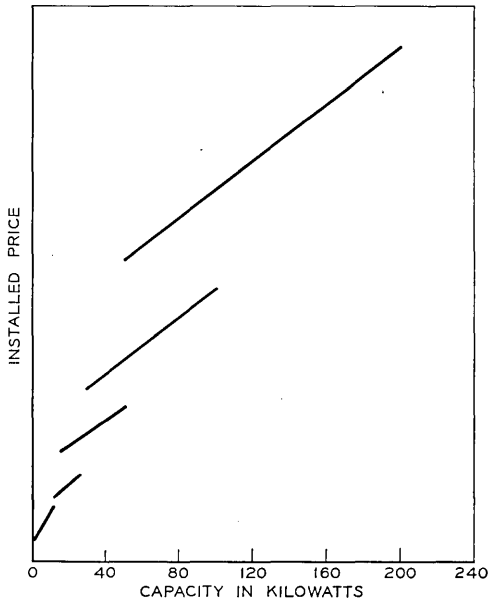


Fig. 6 — Installed value of power plants.

duced to the next lower-range unit. Such a saving can be realized only part of the time, but it is desirable to attempt to weight its value so that over the telephone system as a whole there is adequate encouragement to strive for power savings.

The method of weighting is complicated by such factors as the fraction of power that may be saved in an office, by the likelihood of certain size offices being representative, and by the extent to which offices are designed with safety factors for growth and overload. Nevertheless a reasonable estimate can be made by assuming differing amounts of the office power can be saved, and by assuming that all sizes of office are equally likely. This gives the result that the over-all incremental value of power is

$$k = \frac{C}{W} = a + \frac{2AW_0}{(1-p)(W_2^2 - W_0^2)}, \qquad \text{(dollars per kw)}, \quad (14)$$

the average value of power plant per kilowatt saved over the range between W_0 and W_2, taking into account the proportionate number of plants that can be converted to the next cheaper size, where

C = total price saving,
W = total power saving,
p = fraction of power saved, not to exceed about .5, (above which point two steps of base price saving might be realized),
W_0 = lower limit of range,
W_2 = upper limit of range,
a = incremental price of one size power plant, assumed to be about constant for all plants,
A = difference in base price to next lower plant, as may be tabulated for any particular case.

By these methods, incremental values of installed power plant may be found; it is even possible in many cases to assume an average slope for the family of curves in Fig. 6, without greatly changing the results in practice.

It is now desired to convert these figures of installed expense to figures comparable to manufacturing cost, in the factory, of the relays which will consume this power. This may be done through standard economic-study practice by which first cost is converted to annual charges through a knowledge of charges for interest, taxes, depreciation, etc. As these factors are not necessarily alike for power equipment and relays, the factors used will differ to suit the problem. For the present study, with recognition of the accounting practices then in effect, it was found that

on a basis comparable to relay manufacturing cost, savable power was worth between $500 and $2,500 per kilowatt, depending on size, savable power, etc.

Now these figures must be associated with the power needed by any particular relay. The total switching power during the busy part of the day, which is the period determining the size of the power plant, is the sum of all the individual relay power drains at any particular moment. This in turn may be taken as the sum of the power required by each relay times the probability that it will be energized at any particular time, which is taken as the fraction of the busy hour that the relay is expected to be energized. For every relay, this ratio can be determined with some certainty, by consultation with the circuit designers.

Summarizing, the switching power capacity required by the office may be stated as the sum, for all relays, of power w for each relay, times m, the fraction of the busy hour that it is energized (i.e., hrs. per busy hr.). The cost of the power plant capacity required for each relay is then

$$C = kmw \text{ dollars}, \tag{15}$$

where k is the equivalent value of power.

Very often it is desirable to state this cost on the basis of the annual power drain. In this way, the plant investment can be easily correlated with the costs paid to the power companies. In telephone offices, the annual power has been found to correspond to 3,000 times the power consumed in a busy hour, and as a matter of convenience power drain in the central office is often stated in terms of the energy per year based on a 3,000-busy-hour year. Then the ratio of use during the busy hour may also be stated as

$$m = \frac{h}{3000},$$

where h = hours per year energized.

Writing $E^2/1000R$ for the power in kilowatts drawn by any relay, the equivalent first cost of the power plant assignable to the relay is then

$$\begin{aligned} C_{\text{power plant}} &= \frac{E^2}{1000R} \times \frac{kh}{3000}, \\ &= c_{P_1} \frac{E^2 h}{R}. \end{aligned} \tag{16}$$

For this case, in other words, the equivalent first cost of power plant has a value c_{P_1} dollars for each watt-hour-per-year of switching power required.

1.32 *Cost of Power Purchased from Power Companies*

The power used by the switching apparatus corresponds to a larger amount of power furnished by the power company, since the power machines are not 100 per cent efficient. Its cost may be stated as

$$c_P = \frac{c_0}{e},$$

where c_0 = average marginal cost of power furnished,
e = conversion efficiency.

The annual power bill will then be

$$\frac{c_0}{e} \frac{E^2 h}{1000R} \text{ dollars.}$$

Converting from annual charges to first cost, through the same steps as above, gives

$$C_{\text{power}} = c_{P_2} \frac{E^2 h}{R}, \tag{17}$$

or c_{P_2} dollars per watt-hour-per-year.

1.33 *Total Power Cost*

With both power plant costs and purchased power costs now stated on the same basis, they may be combined to give the total power cost, C_P:

$$C_P = (c_{P_1} + c_{P_2}) \frac{E^2 h}{1000R} = \frac{c_P E^2 h}{R}, \tag{18}$$

where c_P represents the dollars per watt-hour-per-year for power consumed by a particular relay, and is the equivalent first-cost value of power.

Thus a method is available to find the value of the power that magnet coils may need — to be balanced against the first cost of the coils (Section 1.1). Procedures for best results are given later.

1.4 THE EQUIVALENT MANUFACTURING COST OF THE FUNCTIONING TIME OF A RELAY

Another large part of the investment in every crossbar-type central office is the common-control equipment, most important of which is the "marker." This equipment has the sole duty of selecting the proper path for each call, and sending it on its way. After each such function,

the equipment restores to normal, ready to serve the next call. It takes a marker about a half-second to do its job; nevertheless many markers may be needed to handle the traffic during each day's heaviest calling period. Since the cost of each marker is measured in the tens of thousands of dollars, it becomes a matter of great importance to shorten the marker work time — in other words, to shorten the time of each relay in the chain of marker events. Since faster relays are usually more expensive it is important to find the dollar value of speed so as to guide an intelligent design of each speed relay.

As in the case of power, it has been found possible to state the value of speed in terms equivalent to the manufacturing costs of relays. This may be done through a knowledge of the number of markers and associated equipment needed in relation to their work time, and their cost. Such relationships are shown in Fig. 7, which gives a typical curve for the markers per line needed to handle the traffic. Then, when the value of the marker is known, the value per line per millisecond, c_t, follows. Corresponding to this, a value per millisecond per marker, c_M, can then be found, which varies inversely as the number of markers needed:

$$c_M = \frac{c_t}{p},$$

where p = number of markers per line that are needed. This figure is the value of a millisecond for any complete event or series of events

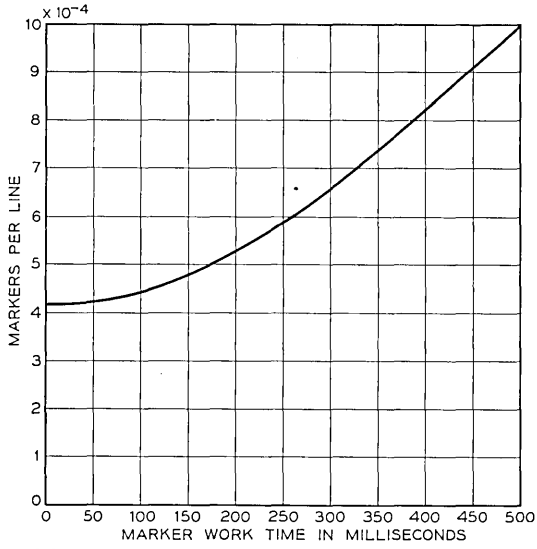


Fig. 7 — Markers as a function of work time.

within the marker which directly affects its working time, referred to as a "major event."

1.41 *The Value of Speed per Event*

Within each major event there may be a whole cycle of operations happening partly in parallel and partly in sequence. Each separate action is called an "event." In the case of events happening in parallel, the circuit designer will usually know whenever one path is in outstanding control of the work time as a whole. Whenever such is the case, it is a "major event," and the other paths may be ignored. Often, however, the times for each of several paths will be about the same, and special consideration is needed to determine their value. In breaking down these operations so as to arrive at the cost per relay, one must evaluate the events within any major event as in the region marked in Fig. 8. Supposing that the number of subevents within a major event are labelled a, b, c, etc. for each of the several parallel paths, then the total number of subevents is $a + b + c + d + \cdots$. The total time for the major event is t, which has the dollar value, c_t/p, per millisecond, as just previously given.

Along any path, A, of events, the total time for the events is the time taken by the group as a whole. If this total time is t, then the effective time per event along any path is

$$t_A = t/a \text{ for path } A \text{ (seconds)},$$
$$t_B = t/b \text{ for path } B \text{ (seconds)},$$
$$= \text{etc.}$$

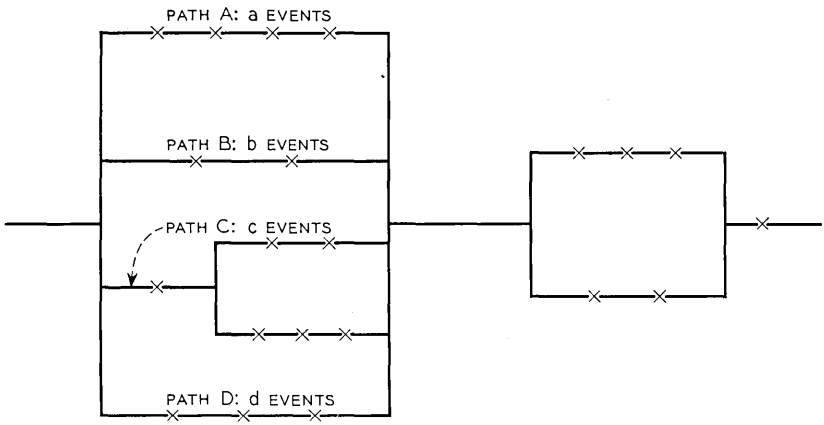


Fig. 8 — Typical chain of events in controller operation.

The worth of the time taken for the group as a whole is then

$$C_M = c_t t/p = c_t a t_A/p = c_t b t_B/p = \text{etc., dollars.}$$

The value per event is usually of principal interest. Since there are $(a + b + c + \cdots)$ events in any one "major event," then the total value per event C_E is

$$C_E = \frac{C_M}{(a + b + c + \cdots)} \text{ dollars per event,}$$

$$= \frac{a}{(a + b + c + \cdots)p} c_t t_A, \qquad \text{for path } A,$$

$$= \frac{b}{(a + b + c + \cdots)p} c_t t_B, \qquad \text{for path } B,$$

$$= \text{etc.}$$

Then the value of an event, per millisecond, is

$$c_E = q_A c_t/p = q_B c_t/p = \text{etc., dollars per event per millisecond.}$$

where $q_A = \dfrac{a}{a + b + c + \cdots}$, $\quad q_B = \dfrac{b}{a + b + c + \cdots}$, etc.

In general, then, the value of an event is

$$C_E = c_E t_E \text{ dollars,}$$

where t_E = the duration of the event,

$$c_E = \left(\frac{q}{p}\right) c_t \text{ dollars per millisecond.} \tag{19}$$

The quantity q will be called the "path factor." It is seen to simplify to a value of unity when a major event is involved, $(q = a/a)$, and in all other cases to have a value less than this.

1.42 *The Value of Speed, per Relay Operation*

Each event may consist of the action of one relay out of several possible alternates. This means that several relays must be purchased for each such event, though only one determines the time. Calling w the number of relays provided per marker to perform one function, each function being designated by subscript corresponding to path and se-

quence, then the value per relay of one millisecond is

$$\begin{aligned} c_r &= \frac{q_A c_t}{w_{A_1} p}, \qquad \frac{q_A c_t}{w_{A_2} p} \quad \text{etc., for path } A, \\ &= \frac{q_B c_t}{w_{B_1} p}, \qquad \frac{q_B c_t}{w_{B_2} p} \quad \text{etc., for path } B, \\ &= \frac{q c_t}{w p}, \qquad \text{in general, (dollars per millisecond).} \end{aligned} \tag{20}$$

These equations give the general expression for the value of functioning time in terms of cost. They permit the direct comparison with cost of a relay design, including its coding cost, manufacturing cost, or power drain cost, in deciding on how much investment is warranted in switching apparatus in order to shorten the marker work time.

1.43 *Simple Examples of Relay Design for Speed*

The above relations make it possible to rapidly evaluate design changes for speeding up relays, or to state the value of seeking a design change that will make speed possible. Several examples will be given.

(*a*) *Relay Involved in a Major Event*

Suppose the following typical conditions are given:

$$\begin{aligned} \text{No. of markers per line} &= p = 0.0006. \\ \text{Path factor} &= q = 1. \\ \text{Relays per marker function} &= w = 1. \\ \text{Value per millisecond per line} &= c_t = \$0.04. \end{aligned}$$

Then value per millisecond per relay is

$$c_r = \frac{0.04 \times 1}{0.0006 \times 1} = \$67.$$

In this case a very large investment per relay could be justified in order to gain small savings in time.

(*b*) *Relay Involved in a Subevent*

Suppose the following typical conditions are given:

$$\begin{aligned} \text{No. of markers per line} &= p = 0.0006. \\ \text{Path factor} &= q = \tfrac{2}{8} = 0.25. \end{aligned}$$

$$\text{Relays per marker function} = w = 2,$$

$$c_t = 0.04.$$

Then value per millisecond per relay is

$$c_r = \frac{0.04 \times 0.25}{0.0006 \times 2} = \$8.33.$$

In this case, too, it is worth a considerable manufacturing cost penalty to shorten the functioning time of the relay.

(c) *Relay Involved in a Subevent — Release*

A particularly common form of subevent is the release of numerous relays at the end of one phase of marker action. In this case, assume values of the constants are:

$$\text{No. of markers per line} = p = 0.0006.$$

$$\text{Path factor} = q = 0.02.$$

$$\text{Relays per marker function} = w = 5.$$

$$c_t = \$0.04.$$

Then value per millisecond per relay is

$$c_r = \frac{0.04 \times 0.02}{0.0006 \times 5} = \$0.266$$

Thus even in this case, an appreciable first cost value is associated with the releasing speed of each relay.

1.5 REVIEW OF EQUIVALENT FIRST COST EXPRESSIONS

The previous sections have shown how some of the variable portions of the relay design that affect its performance may be stated in terms of first cost. Likewise methods are given for evaluating sensitivity and speed performance in the same terms. These may be summarized as

Cost of a coil:

$$C_W = c_R R + c_N N. \tag{1}$$

Cost of an item as a function of demand:

$$C_o = \sqrt{\frac{2kqS(C_1 + A/n)}{n}} + \frac{kqS}{2n} + \frac{A}{n}. \tag{12}$$

Cost of power:

$$C_P = \frac{c_P E^2 h}{R} . \tag{18}$$

Cost of functioning time:

$$C_r = \frac{q c_t}{wp} . \tag{20}$$

Examination of these equations will show that, in nearly every case, a cheapening of the cost of the structure leads to more expensive (i.e., poorer) performance, and vice versa. However, the availability of these relations, together with accurate values for the constants, gives the information needed to vary the relay design to the point where maximum over-all economy (i.e., the cheapest central office) can be realized.

In the following sections, some examples of methods for finding this optimum point, and how much it is worth to find it, will be given.

Part II — Design for Maximum Power-Plus-Coil Economy

Among the more common problems facing the designer of switching circuits is the selection of the relay magnet which gives best economy in use. On the one hand, equation (1) has shown that the relay coil cost will be less if the resistance is less (but power consumed is more). On the other hand, equation (18) shows that the equivalent first cost of that part of the power plant furnished to operate the relay increases as the resistance is reduced (but coil is more expensive). These opposite effects are indicated in Fig. 9 where the curve labelled C_W shows how wind-

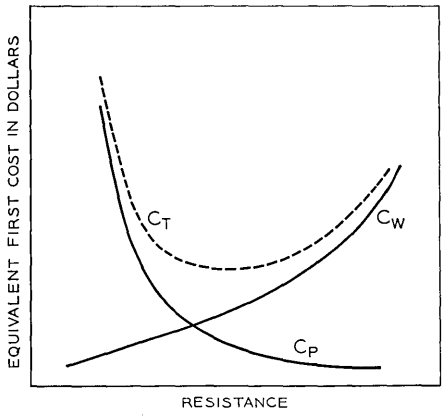


Fig. 9 — Power and winding costs.

ing cost for a certain wire gauge varies while the curve labelled C_P shows how power plant cost is changing. The desirable end result is a coil with resistance value to permit the sum of these to be as low as possible, as in the curve marked C_T. The following methods will give such a design.

The value of C_T, made up of power cost, C_P, and winding cost, C_W, represents the cost of all the parts of the central office which are influenced by a change in resistance of the coil. Thus we wish to minimize this equation:

$$C_T = C_P + C_W,$$

where, from the previous relations,

$$C_P = \frac{K_1}{R},$$

$$C_W = K_2 R,$$

$$K_1 = c_P h E^2,$$

$$K_2 = c_R \left(1 + \frac{c_N}{c_R} \frac{N}{R}\right).$$

The resulting expression for this total cost is

$$C_T = \frac{K_1}{R} + K_2 R. \tag{21}$$

In this equation K_1 is a constant, but K_2 varies both with resistance and wire size, as the latter affects the values of c_N, c_R, and the quantity N/R, which is determined by the coil design equation

$$R = AN\,(h + d),$$

in which, using the special notation common to coil design problems:

N = turns in winding = $h\ell/K$,

R = coil circuit resistance,

A = π times the resistivity of wire in ohms per inch length, given in tables for each size of wire,

h = depth of coil space occupied by wire,

ℓ = length of winding space,

d = inside diameter of coil,

K = effective cross-sectional area required by one turn of wire, given in tables for each size.

For a given relay, the dimensions ℓ and d are fixed, and only the wire size and the coil depth h can be varied. For a given wire size, therefore, the values of h and N are fixed by the resistance. As C_N and C_R are known quantities, the values of K_2 can be determined and plotted against R for each wire size, as illustrated in Fig. 10. Such a chart applies to a particular relay, as the curves depend on the fixed values of ℓ and d applying. Here it can be seen that for a given resistance the finest gauge wire is the cheapest. Because of this, and since power costs are minimized with the highest resistance, the finest gauge wire would always give the cheapest coil. However, this becomes a theoretical case. As finer and finer wires are used, $A = \pi\sigma$ becomes larger, and thus N/R or NI/E becomes smaller. In all relay design, an ampere-turns requirement must be met. Thus a practical solution to this problem of optimum costs is to choose the finest gauge wire which will meet the ampere-turns requirement.

If the same relay is assumed as for Fig. 10, the variations of ampere turns with resistance are as shown in Fig. 11. Starting with the NI required, the wire size can be determined from a plot like Figure 11. Then K_2 is determined from a plot like Fig. 10.

Now, with the gauge of wire chosen, K_2 is almost a constant and can be considered independent of R. Thus, upon differentiating (21), C_T will be found to be a minimum when R has a value designated by

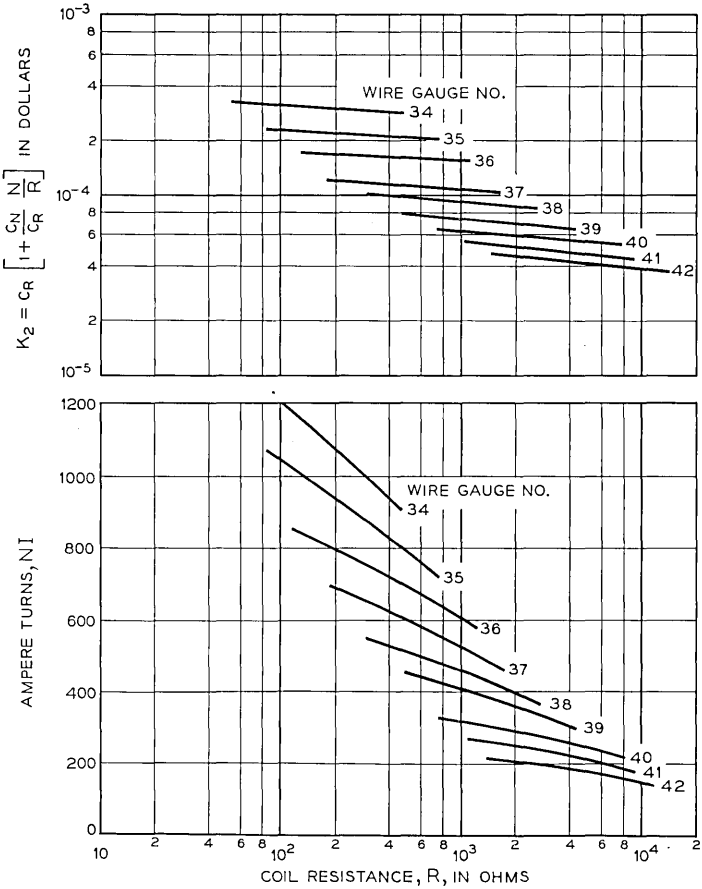
$$R_0 = \sqrt{K_1/K_2}\,. \tag{22}$$

The most economical coil, resulting from use of a resistance R_0, will have a system cost of

$$C_0 = 2\sqrt{K_1K_2}\,. \tag{23}$$

As can be seen by the slopes of the curves on Fig. 11, if the optimum resistance, R_0, is larger than the resistance value at the desired ampere turns, the NI requirement will not be met. Then either the selection of R_0 must be made for the next coarser wire, or the resistance value chosen at just the required NI, whichever is cheaper. Conversely, if there is a very large NI margin, a finer gauge wire could be tried. In all, two trials should give the optimum values.

A procedure has now been outlined for choosing the coil design for optimum power-plus-winding cost. Curves similar to Figs. 10 and 11 should be drawn for the applicable relay parameters. Then the K_2 value is obtained for the finest gauge meeting the ampere-turns requirement, and R_0 and C_0 found. Then, the NI value applying for R_0 must be checked and adjustments made as above, if necessary.



Figs. 10 and 11 — K_2 and NI values versus resistance for various wire gauges for a typical relay.

COST WHEN OPTIMUM RESISTANCE CONDITION IS NOT SATISFIED

There will, of course, be cases in practice where the optimum resistance is not used, for such reasons as standardizing of certain coil sizes, need for speed, or insufficient winding space. The penalty in cost from departing from the ideal winding size may be found by comparing the cost for any value of resistance, (equation 21), with the cost if resistance were optimum, (equation 23):

$$\frac{C}{C_0} = \frac{K_1/R + K_2R}{2\sqrt{K_1K_2}},$$

$$= \frac{1}{2}\left(\frac{R_0}{R} + \frac{R}{R_0}\right). \qquad (24)$$

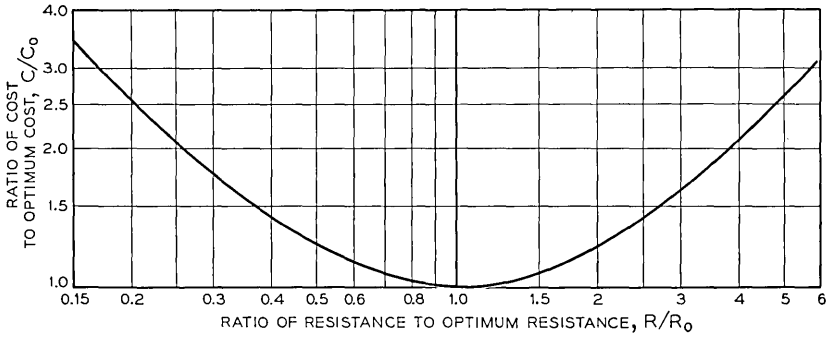


Fig. 12 — Cost penalty when resistance is not optimum.

Thus, the fractional change in cost caused by any departure from the best resistance is readily estimated, expressed in ratio form as R/R_0. This relation is shown in Fig. 12.

SUMMARY

In summary of Part II, methods have been given for deciding on the magnet coil design which will provide the cheapest net cost in the switching system. When the conditions of use are known, a series of charts may be prepared such as the one shown, from which the applications engineer may decide the proper coil design, and how much it may be worth to the system. As experience will show, the cost of departing from an optimum design may run well over $1.00 per individual relay, with the result that aggregate savings in the whole central office may run to thousands of dollars compared to the cost when this problem is ignored.

New relay development may be guided by the use of these same relations, since through a process of considering various hypothetical designs of various sizes, work outputs and magnetic qualities, the potential optimum costs may be found. Design of the recently adopted wire spring relay was very materially guided by considerations of this sort.

Part III—Choice of Contact Sets for Maximum Economy

One of the many problems involving lot-size costs of relays is the number of standardized sets of contacts that should be provided: i.e., how many kinds of contact arrangements give maximum economy for the product as a whole? By providing only a few choices of kinds of contacts, for example, the annual output is less divided and there are fewer setup charges, with the result that the initial manufacturing economy is large. However, because of the small selection of kinds of contact sets, there will result many cases where extra contacts are provided at greater than

needed cost. The problem is how best to combine these opposing factors.

In illustration, consider a relay which is to be equipped with molded spring arrangements, capable of providing any number of "transfers" from 1 to 12, each of which is used in approximately equal numbers in the system. At the one extreme, only one molding might be provided, always equipped with 12 transfers. At the other extreme 12 different moldings might be provided, covering each individually needed quantity of "transfers." In the case of the single molding there is no lot-size penalty at all, but on the average a large penalty in surplus contacts; while for the twelve kinds of moldings, lot-size penalties corresponding to one-twelfth of the possible annual output are incurred, but no spare parts at all will be needed. Such a problem may be treated as given below.

Let the number of kinds of spring sets chosen be designated by v. Then the annual output of each kind is the total output N divided by v, or

$$n = \frac{N}{v}. \tag{25}$$

Now it was shown in Section 1.2 that the cost penalty for making things in lots is given by equation (12). For the present problem, this may be expressed in terms of v, by substituting equation (25) into equation (12). The resulting lot-size cost penalty is then

$$C_0 = \frac{v}{N}\left[\sqrt{\frac{2kSN}{v}\left(1 - \frac{1}{v}\right)\left(C_1 + \frac{Av}{N}\right)} + \frac{kS\left(1 - \frac{1}{v}\right)}{2} + A\right]. \tag{26}$$

The penalty due to providing unneeded springs is the dollar value of the average number of spare springs. The average number of spares is approximately

$$E = \frac{12 - v}{2v}.$$

Thus for only one kind furnished, always 12 transfers, the number of spares can vary between 0 and 11, an average of 5.5, as given by the equation. If there were eight kinds, consisting possibly of moldings of 1, 3, 5, 6, 7, 8, 10 and 12 sets of transfers, the average number of extras would be 0.25. The dollar penalty compared to no extras ever needed is

$$C_E = C_1\left(\frac{12 - v}{2v}\right). \tag{27}$$

and the sum of this equation and that for lot size represents the system cost for any one condition.

Fig. 13 shows this variation for an assumed case for which

$$N = 500{,}000 \text{ parts per year.}$$

$$k = 0.15.$$

$$S = \$100.$$

$$C_1 = \begin{cases} 0.20 \text{ for spring set cost.} \\ 0.005 \text{ for spare spring cost.} \end{cases}$$

$$A = 0.$$

Here it is found that the greatest benefit comes from using in the neighborhood of six kinds of spring sets. It is further clearly seen that the system as a whole will not be particularly penalized if the number of sets chosen is a few different from this quantity — but a penalty of almost two cents per relay, or about two thousand dolllars per year would be incurred by erring too far toward the extreme.

Part IV — Choice of Special Adjustment for Maximum Speed Economy

Speed can be an exceedingly valuable property of a relay, as shown in Section 1.4. There are cases for example where every saving of 1

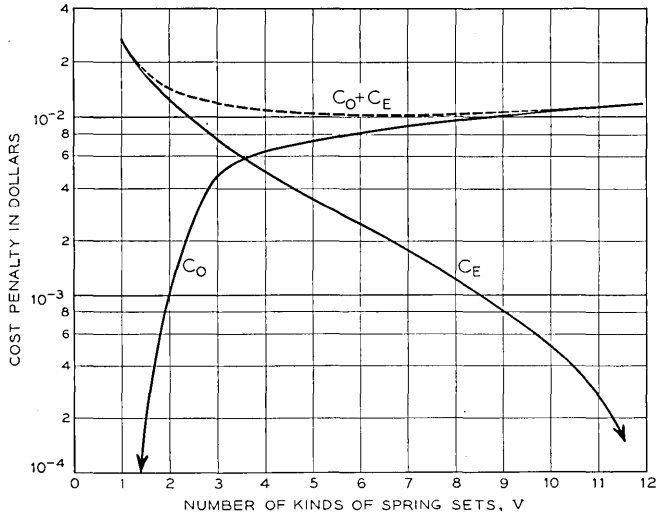


Fig. 13 — Optimum number of spring sets.

millisecond in working time will reduce the average quantity of equipment needed by an amount equivalent to $35, or even more. For such cases a considerable sum of money can be spent on each relay in order to make it faster. The following approach helps guide the circuit and relay designer in the choice of relay they should make in order to gain the greatest over-all advantage.

It is well known to relay designers that one of the most important design parameters controlling operating time of a relay is the spacing between the contacts when they are at rest. This space cannot be less than a certain small distance of about 0.005″ to avoid contact failures due to vibration, buildups, sparkover, etc. However, because it is costly to hold adjustments accurately to a few thousandths of an inch, the maximum spacing is often quite large, sometimes as much as 0.050″. The spread in this distance is primarily a matter of economics; if it were clear that more money could be spent on the relay to narrow down this spacing while gaining material reductions in operating time as a result, then some closer adjustment would be specified. Such a problem is readily treated by summing (a) the cost penalties for differing values of this spacing and (b) the cost penalties of the corresponding functioning time, and finding the optimum condition that results. One such case is shown below.

The cost penalty due to changing spacing may be estimated by considering various actual values, and determining what procedures would be used in the factory in each case. Such a cost study can be made in the factory. In illustration, assume a particular relay type deemed to suffer no cost penalty if its armature stroke is allowed to take on any value up to 0.050″, but which cannot be less than 0.010″ without impairing performance. The actual cost per unit of various kinds of manufacturing procedures might be found according to the hypothetical conditions given in Table I. Thus a curve can be plotted of the cost penalties for each setting of gap.

The design may also be studied for the influence of the gap settings on speed. The operating time of the magnet as a function of spacing is readily determined by experiment, and may also be checked against known operating principles, to be sure the figures are reasonable. For a typical case, operating times corresponding to various spacings might be as given in Table II. At the same time, the value of this functioning time is given by equation (20) to be

$$C_r = \frac{qc_t}{wp}.$$

TABLE I — COST PER UNIT OF VARIOUS KINDS OF MANUFACTURING PROCEDURES

Assumed Max. Stroke	Manufacturing Procedure	Cost
0.050″	Acceptance of wide dimensional tolerances on all parts, and no adjustment permitted	$0
0.045″	As above, but crude adjustments added	0.05
0.040″	Closer dimensional tolerances, no adjustment	0.07
0.034″	As just above, with rough adjustment	0.10
0.025″	Very close tolerances on all parts, no adjustment	0.15
0.020″	As just above, with touch-up permitted	0.20
0.015″	Screw adjustment added, (more expensive parts, but simpler adjustment)	0.35
0.010″	Precision setting of all parts	0.65

Assuming for this hypothetical case that the values applying are

$$c_t = \$0.05,$$

$$w = 100,$$

$$q = 0.2, \text{ and}$$

$$p = 0.0006,$$

then the cost of the resulting time is found in the last column of Table II.

A summation of these two sets of cost penalties is given in Fig. 14, which clearly shows that the ideal value of stroke is about 0.020″, and that that it may be permitted to vary between the limits 0.015″ and 0.025″ without a serious economic penalty.

In summary, a method has been indicated for deciding how important it may be to build in, or omit, expensive design features which have an effect on functioning time. An illustration has been given for an assumed common control cirucit and an assumed change in a design feature of the relay that importantly affects the time. The method, however, is equally

TABLE II — OPERATING TIMES CORRESPONDING TO VARIOUS SPACINGS

Assumed Max. Stroke	Functioning Time (Milliseconds)	Equivalent First Cost Value of Time
0.050	6.1	$1.02
0.045	5.6	0.933
0.040	5.2	0.867
0.035	4.75	0.792
0.025	3.75	0.625
0.020	3.2	0.533
0.015	2.55	0.425
0.010	1.9	0.317

applicable to other speed problems, once the value of speed is known, and once the influence of the design change and its cost are known.

Part V — Choice of Winding for Maximum Combined Economy of Speed and Power

In the selection of magnet designs for use in circuits where speed is important, much can be gained by modifying the mass, the stroke, the force against the stop, etc., but when all these devices have been exhausted there remains the possibility of supplying more power to the magnet coil. If this is done in enough cases, an over-all penalty in central office cost is incurred through the added over-all power plant capacity that may be needed. As already seen in Section 1.3, this cost penalty is proportional to the base value of one watt-hour-per-year, the voltage squared, and the hours per year energized, but is inversely proportional to the resistance. In telephone usage, the voltage is usually fixed at 50 volts, and equation (18) may be rewritten as

$$C_P = \frac{k_P h}{R}, \tag{28}$$

where

$$k_P = c_P E^2.$$

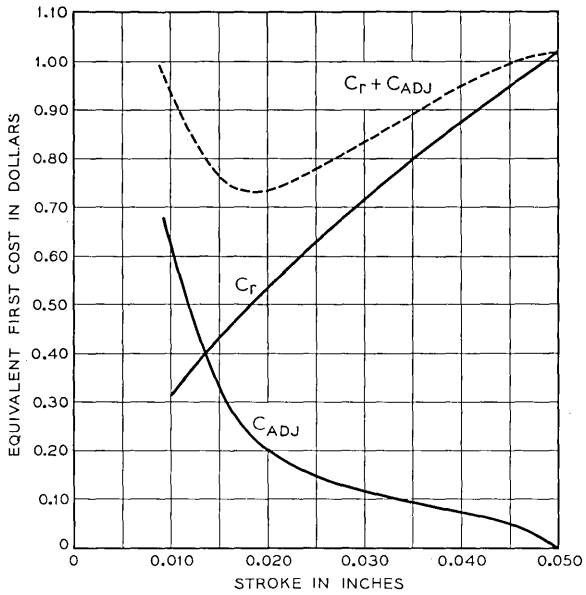


Fig. 14 — Optimizing a speed relay design.

The above expression gives the equivalent first cost of the power taken by one relay. For each relay operation, or event, however, there may be several relays energized. Then the total cost penalty per event, due to power, is

$$C_P = \frac{k_P h g}{R} \text{ dollars,} \tag{29}$$

where g = number of relays energized, per marker, for the particular function in question even though only one appears in the time sequence. For the power cost to be compared with the cost due to speed, this cost must be spread over the number of relays, w, involved in the function. Then

$$C_P = \frac{k_P h g}{w R} \tag{30}$$

This is the equivalent first cost penalty incurred against each relay, due to its power consumption. As R increases, it is seen to decrease.

However, as R increases, the operate time increases. For speed relays, the time t has been found to be a complicated function of the resistance, varying approximately as $R^{1/3}$. Thus when R increases, time increases, and the equivalent first cost of this action time increases, as already seen in Section 1.4.

As power is changed, then, there results two oppositely varying costs — one due to power, and one due to speed. The problem is graphically shown in Fig. 15, where Curve A shows the typical cost variation due to power drain and Curve B shows the variation in cost due to speed. The point where the sum of the two is a minimum is the ideal point to operate, assuming that the coil costs are about equal in each case. Curve C shows how the total cost will vary. Actually, it may be necessary to add the effects of a third variable: the cost of the coil itself as it is redesigned for different power consumption. This can be easily done by the methods shown above, but is ignored in the present treatment for the sake of simplicity. In many actual cases, coil cost has been found to be a much smaller factor than the other variables.

CHOICE OF OPTIMUM COIL

In practice the above result may be obtained by a knowledge of how a specific magnet design will vary in operate time as its coil resistance is changed. The values may be substituted in the following equation for total cost:

$$C = \frac{qc_t t}{wp} + \frac{k_P hg}{wR},$$
$$= \frac{g}{w}\left(\frac{qc_t t}{pg} + \frac{k_P h}{R}\right). \tag{31}$$

The expression may be evaluated when a relation between t and circuit resistance R is established, usually a matter for experiment in any given case.

Illustration

For a particular case let us assume that $p = 0.0006$, so that equation (31) is

$$C = \frac{g}{w}\left(1670\,\frac{q}{g}\,c_t t + \frac{k_P h}{R}\right). \tag{32}$$

Typical values of operate time, t, for a speed relay in 50-volt telephone applications are given in Table III. From this information, one may readily evaluate equation (32) to cover the practical range of conditions that are met in service. This has been done, in Fig. 16, to give the total

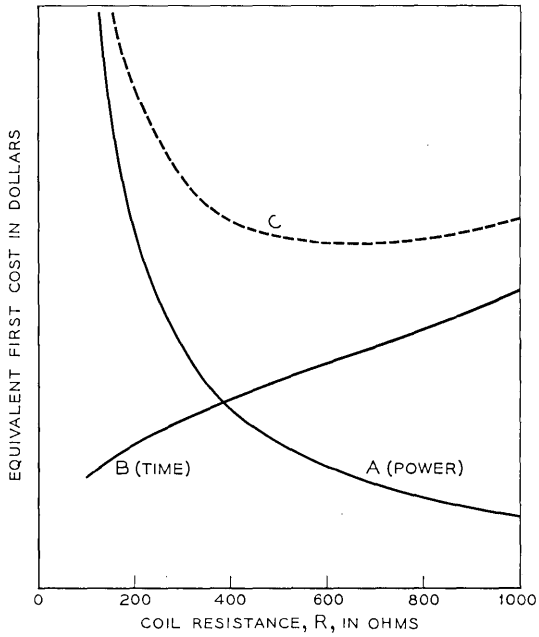


Fig. 15 — Optimizing costs of speed plus power by relay coil design.

TABLE III — TYPICAL VALUES OF OPERATE TIME

Coil Resistance (Ohms)	Optimum Time (Milliseconds)
100	2.45
200	3.1
400	4.15
600	5.0
800	5.8
1000	6.7

cost penalty for the case of very short holding time (h = 100 hr. per year) and for long holding time (h = 2000 hr. per year). The optimum coil resistances for these cases are seen to lie between 200 and 400 ohms for the long holding time condition, and to be as low as possible for the case of short holding time. As a result we see an appreciable economic incentive to plan relay designs which are capable of operating on low resistances. Further study of heat dissipation, operation with holding windings, contacts to withstand heavy currents, and other devices to gain fast action, therefore take on added importance to the relay designer.

PART VI — CHOICE OF THE ACTUAL CODE

With the background of the previous sections, the systems designer can now decide what codes he needs to use in his circuits so as to be

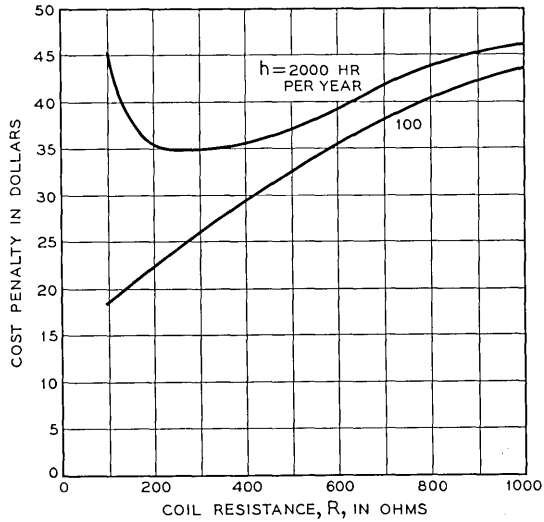


Fig. 16 — Best coil design for both speed and power economy.

most economical, over-all. The problem will first be discussed in a generalized form, and the means for getting the economy will then be reviewed.

6.1 GENERAL DISCUSSION OF CODING

If enough information were available to the circuit designer, he should be able to choose a relay for any particular application by considering

(a) The penalties in performance (expressed in terms of first cost) resulting from having only a certain limited number of available combinations of relays, as against complete flexibility, i.e., the penalties due to standardizing, and

(b) The penalties in first cost resulting from many variations of a basic type, as compared with one single standard type, i.e., the penalties due to *not* standardizing.

By weighing both the penalties and the advantages of standardization in each case, they may be maintained in approximate balance, and give the economical number of codes.

In the development of an entire new switching system using relays, there will be approximate ideas on the number of relays needed in the system, and on the number of kinds of relays required to fairly well satisfy the circuit functions. Based on past experience, the statistical distribution of kinds of relays for the various uses in relation to annual demand may also be approximated. The number of codes provided will determine the number of cells into which the statistical distribution is divided, and also the relative annual demand for each. In the extreme, only one code might be provided. It would have the advantage of very large lot-size, but also it would have to do the hardest as well as the easiest function. Such a requirement could lead to such absurdities as requiring that all relays have three windings, twelve transfers, best grade of contact metal, and the like; or to a greatly increased number of simpler relays. In addition, with but little flexibility in the design, there would be the added performance disadvantages due to imperfectly matching certain needs for power economy, speed, and the like. Hence, there is a large performance disadvantage, if but one kind of relay were provided.

At the other extreme would be cell divisions of kinds of relays to give the perfect match for every circuit use. Then, with no spare contacts, the best possible power economy, and every other condition at its ideal value, no performance penalty at all would be entailed. But every code increase would further subdivide the manufacture, till the lot-size penalties grew excessive.

These variations are shown by the curves in Fig. 1, and it is their sum,

as seen in the curve marked *c*, that should decide where best to work. In planning a whole new central office which may use from 40,000 to 80,000 relays, and a new design of relay whose manufacturing costs are not yet well established, the data to give numerical values for Fig. 1 can only be quite tentative, and the results can only serve as guides. Yet such work has been most useful in recent effort on a redesigned No. 5 crossbar system to use the wire spring relay. Calculations could be made with enough assurance to determine the general shape of the total penalty curve (Curve c), and showed an optimum value for kinds of relays of about 200. Even more important was the indication that the curve was extremely flat in its optimum region, giving but a small change in the total penalty if an error, even as great as 2 to 1, were made in the number of codes to be used.

6.2 THE DETAILED PROCEDURES OF CODE SELECTION

Background information as gained above helps in guiding the early steps in picking codes of relays for the job, but there is a better method of picking codes when the circuit designer is actually down to cases. By following the steps below, the optimum number of codes will automatically result.

The first step in the coding program is to pick a list of basic codes. This can be done quite arbitrarily, initially, and should be based on general knowledge of types of circuit functions, together with specific knowledge of certain uses with very large demands. For example, certain coils will be needed to cover the range from very fast operation with no concern for current drain, to slow release and emphasis on economy of power; there will be single and multiple windings as dictated by circuit functions. Also, certain arrangements of contacts may be assumed to span the range of needs from a single make up to the capacity of the design — twelve transfers or twenty-four makes in the case of the new wire spring relay. Further, certain combinations of springs and coils will be evident at once as ideally suited to particular large-demand uses, and these should be on the original list. Soon the likely minimum demand for each will be quite evident. Such a list of tentative codes and demands is the first step in the coding scheme.

Now, as each circuit application arises, this basic list may be consulted. For some cases, an available code will be ideal, but for many others a new arrangement may be desired. In each such case a simple calculation will show the economical choice between a new code or the old code with more features than are really needed. Steps in the calculations are given

below, where the extension of the old code is identified as Plan A, and the use of a new code as Plan B.

For Plan A, there will be a new demand, $(X + Y)$, which is greater than its previous demand by Y, the quantity of the new application. The demand cost penalty can then be read from charts such as Fig. 4. Performance penalties for pertinent features such as speed, power, extra contacts, etc. may be read from charts of the kind described in the earlier sections above. The sum of these values, multiplied by Y, is the cost penalty for the new application of an existing code.

For Plan B, there will be a lot-size cost penalty due to its demand. But there will also be a penalty imposed on the original code, (whose demand was X), because its demand was not enlarged to value $X + Y$. Thus the total lot-size cost penalty is made up of two parts:

(a) Penalty due to demand Y, multiplied by Y.

(b) [Penalty due to demand X − Penalty due to demand $(X + Y)$] multiplied by X.

The unit lot-size cost penalty for Plan B is the sum of these two factors, divided by the quantity Y. The design of relay for Plan B may be chosen by the methods previously outlined to be as nearly optimum as is feasible, and then performance penalties itemized as in Plan A above. The sum of these penalties, added to the lot-size penalties just mentioned, give a number for comparison with Plan A.

After the penalties due to either plan are compared, the least costly should be chosen. In cases where the costs are approximately equal, preference should probably be given to Plan A, as encouraging standardization.

In summary, the ideal number of codes in a newly designed system

TABLE IV — CHECK LIST FOR RELAY CODE SELECTION
AMOUNT OF EQUIVALENT COST PENALTY

Kind of Penalty	Plan A		Plan B
Lot-size (a)	Note (1)	Note (2)	Note (3)
(b)		$[(2) - (1)] \times \frac{X}{Y} =$	
Power			
Speed			
Extra parts			
Other			
Total			

Note (1): Fill in for demand $X + Y$.
Note (2): Fill in for demand X.
Note (3): Fill in for demand Y.

may be approximated by choosing the more economical alternative resulting from filling in the check list given in Table IV.

Part VII — Summary

In the telephone central office some 60,000 relays are needed to carry out the automatic switching functions. Since the entire office is built around them, they determine the cost of the office, not only by their first cost, and maintenance cost, but at least equally importantly by their influence on the power plant, and the amount of common control equipment that is needed. There is opportunity for great economy, overall, when manufacturing cost is put in proper balance with the cost of power, cost of functioning time, and similar performance variables.

The preceding pages have shown how each factor in the total system cost may be stated in a common language — the equivalent first cost value. When each variable, expressed as an incremental figure denoting the cost compared to an ideal standard, is so stated, all relay factors in the switching system may be cross-compared. This permits a large number of optimizing procedures to be carried out.

Optimizing methods for several important cases have been illustrated, and relations or procedures of value to the relay designer are presented. Particularly outstanding cases are the means for realizing (a) power-plus-coil economy and (b) power-plus-speed economy.

On the side of manufacturing cost, relay designers need some guidance as to how their applications problems will be influenced by the lot-size, or annual demand. Optimizing procedures are also given for this problem, to yield approximate values for optimum lot-sizes under various manufacturing conditions, and the corresponding cost penalty as a function of annual demand.

The steps to be taken in order to strike the best balance between standardizing for maximum manufacturing economy and diversifying for maximum performance economy are also given.

The problem as a whole is a striking example of how design for service can be applied on a very large scale. The economies realized by the approach used here have avoided the expenditure of many additional thousands of dollars yearly to the telephone companies, with corresponding economy to the customer.

ACKNOWLEDGMENT

Appreciation is here expressed to R. L. Peek, Jr., and Mrs. K. R. Randall for helpful technical suggestions, and also to Mrs. Randall for help with the preparation of figures and examples.

Symbols

(Manuscript received October 20, 1953)

The following list gives the symbols used through the several articles appearing in this issue of the JOURNAL. The list does not include all the variant forms of the different symbols distinguished by subscripts, as these distinctions are indicated in the context where they are used.

MECHANICAL:

a Cross sectional area
ℓ Length
m Effective armature mass
T Kinetic energy
V Work done to overcome static load
x Armature displacement

ELECTRICAL:

e Copper efficiency or fraction of the copper volume occupied by conductor
E Applied battery voltage
G Total equivalent single turn conductance
G_C Equivalent single turn conductance of coil; N^2/R
G_E Equivalent single turn eddy current conductance
G'_E Effective single turn eddy current conductance; $G_E e^{-G_E/(G_C+G_S)}$
G_S Equivalent single turn conductance of sleeve
i Instantaneous current
I Steady state current
m Mean length of turn in winding
N Number of turns in winding
L_1 Single turn inductance; $\dfrac{4\pi}{\mathcal{R}'(x)}$
NI Ampere turns
NI_0 Just operate or just release ampere turn value
q Ratio of just operate to final ampere turns; $\dfrac{NI_0}{NI}$

ρ Resistivity of material
R Resistance
S Coil volume
v Ratio of flux attained at time t to steady state flux
W Power; I^2R

MAGNETIC:

A Equivalent pole face area
A_2 Effective pole face area (design value)
B Induction (flux density)
B' Value of B for maximum permeability
B'' Saturation density
B_M Density at $\mu = 1000$
B_R Remanence (density at $H = 0$)
C_L $\dfrac{\mathfrak{R}_0 + \mathfrak{R}_L}{\mathfrak{R}_0}$
F Magnetic pull
$\mathfrak{F}$ Magnetomotive force; $4\pi NI$
$\mathfrak{F}_C$ Coercive magnetomotive force
H Field intensity
H_C Coercive force
L Inductance
μ Permeability
μ_A Permeability of air
μ' Maximum value of permeability
$\mathfrak{R}$ Reluctance
$\mathfrak{R}_C$ Reluctance of the core
$\mathfrak{R}_C'$ Minimum value of core reluctance
$\mathfrak{R}_1$ $\mathfrak{R}(x)$ for $x = x_1$
$\mathfrak{R}_i$ Initial incremental reluctance
$\mathfrak{R}_0$ Equivalent closed gap reluctance
$\mathfrak{R}_L$ Equivalent leakage reluctance
$\mathfrak{R}_{L2}$ Effective leakage reluctance (design value)
$\mathfrak{R}_{02}$ Effective closed gap reluctance (design value)
$\mathfrak{R}(x)$ Equivalent relay reluctance $\dfrac{\mathfrak{R}_L\left(\mathfrak{R}_0 + \dfrac{x}{A}\right)}{\mathfrak{R}_0 + \mathfrak{R}_L + \dfrac{x}{A}}$
φ flux
Φ Steady state flux

φ_1 Initial equilibrium flux
φ' Flux for maximum permeability or minimum reluctance
φ'' Flux at saturation
φ_0 Residual flux
φ_G Gap flux
U Field energy
u $x/A\mathfrak{R}_0$ or x/x_0
W Mechanical work done by magnet
x_0 $A\mathfrak{R}_0$

TIME:

t Time
t_0 Operate or release time
t_1 Waiting time
t_2 Motion time
t_3 Stagger time
t_E Eddy current time constant; $L_1G'_E$
t_C Winding time constant; L_1G_C
t_S Sleeve time constant; L_1G_S

Abstracts of Bell System Technical Papers* Not Published in this Journal

BENNETT, W.[1]

Telephone-System Applications of Recorded Machine Announcements, A.I.E.E., Trans., Commun. & Electronics **8,** pp. 478–483, Sept., 1953.

BIELING, D., see D. EDELSON.

BIGGS, B. S.[1] and W. L. HAWKINS[1]

Oxidative Aging of Polyethylene, Modern Plastics, **31,** pp. 121–122, 124+, Sept., 1953 (Monograph 2155).

Thermal oxidation of polyethylene follows the pattern set by lower homologues such as paraffinic waxes and oils. It is an autocatalytic free radical chain reaction and is subject to inhibition by typical antioxidants. The rate of degradation in the dark at room temperatures is found to be extremely low. Photo-oxidation of polyethylene is rapid in contrast with that of saturated low molecular weight aliphatic hydrocarbons. Furthermore, antioxidants are of little benefit in protecting against exposure to light. Opaque pigments are of great value in reducing the effects of light, finely divided carbon black being particularly effective. By proper compounding polyethylene can be made to last many years outdoors.

BOZORTH, R. M., see H. J. WILLIAMS.

BROWN, W. L.[1]

n-Type Surface Conductivity on p-Type Germanium, Phys. Rev., **91,** pp. 518-527, Aug. 1, 1953 (Monograph 2173).

* Certain of these papers are available as Bell System Monographs and may be obtained on request to the publication Department, Bell Telephone Laboratories, Inc., 463 West Street, New York 14, N. Y. For papers available in this form, the monograph number is given in parentheses following the date of publication, and this number should be given in all requests.

[1] Bell Telephone Laboratories.

A positive charge on the surface of a *p*-type germanium crystal induces a net negative space charge within the crystal adjacent to the surface. This space charge is composed of ionized acceptor atoms and also of electrons under certain conditions. When electrons occur they provide a layer of *n*-type conductivity immediately under the *p*-type germanium surface. Such a layer has been found on the *p*-type region of some *n-p-n*-transistors. In the *n-p-n* structure the layer of electrons appears as an extra conducting path—"a channel"—across the *p*-type material between the two *n*-type ends. The conductance of a channel and the capacity between the channel and the *p*-type material have been measured and compared with the theoretical predictions based on a simple model.

CORNELL, L. P., see E. P. SMITH.

CRANE, G. R.[4], F. HAUSER[4] AND H. A. MANLEY[4]

Westrex Film Editor, J.S.M.P.T.E., **61,** Part 1, pp. 316–323, Sept., 1953.

This paper describes a film-editing machine which employs continuous projection resulting in quiet operation. It accommodates standard-picture and photographic or magnetic sound film as well as composite sound-picture film. Differential synchronizing of sound and picture while running, automatic fast stop and simplified threading features in the film gates with finger-tip release materially increase operating efficiency.

CRISSMAN, L.[3]

See Both Sides of the Game, Tele-Tech., **12,** p. 84, Sept., 1953.

DACEY, G. C.[1] AND I. M. ROSS[1]

Unipolar "Field-Effect" Transistor, I.R.E., Proc., **41,** pp. 970–979, Aug., 1953.

Unipolar "field-effect" transistors of a type suggested by W. Shockley have been constructed and tested. The idealized theory of Shockley has been extended to cover the actual geometries involved, and design nomographs are presented. It is found that these structures can be designed in such a way as to yield a negative resistance at the input terminals. The characteristics of several units are presented and analyzed. It is shown that these characteristics are in substantial agreement with the extended theory. Finally a speculative evaluation of the possible future applications of field effect transistors is made.

[1] Bell Telephone Laboratories.
[3] Western Electric Company.
[4] Westrex Corporation.

EDELSON, D.[1], C. A. BIELING[1] AND G. T. KOHMAN[1]

Electrical Decomposition of Sulfur Hexafluoride, Ind. and Eng. Chem., **45,** pp. 2094–2096, Sept., 1953.

EHRBAR, R. D., see C. H. ELMENDORF.

ELMENDORF, C. H.[1], R. D. EHRBAR[1], R. H. KLIE[7], AND A. J. GROSSMAN[1]

The L3 Coaxial System, A.I.E.E., Trans. Commun. & Electronics, **8,** pp. 395–413, Sept., 1953.

The L3 coaxial system is a new broad-band facility for use with existing and new coaxial cables. It makes possible the transmission of 1,860 telephone channels or 600 telephone channels and a television channel in each direction on a pair of coaxial tubes. The principal system design problems and the methods used in their solution are described in terms of its components and their location in the system.

FINE, M. E.[1]

C_p-C_v in Silicon and Germanium, Letter to the Editor, J. Chem. Phys., **21,** P. 1427, Aug., 1953.

FRAYNE, J. G.[4] AND E. W. TEMPLIN[4]

Stereophonic Recording and Reproducing Equipment, J.S.M.P.T.E., **61,** Part 2, pp. 395–407, Sept., 1953.

This paper describes new stereophonic recording channel equipment including a six-position mixer and portable three-channel recorder. For re-recording, the previously described triple-track recorder-reproducer is available. For review-room and theater reproduction, a theater-type dummy equipped for three-channel stereophonic reproduction is described.

GOERTZ, M., see H. J. WILLIAMS.

GRAMELS, J.[1]

Problems to Consider in Applying Selenium Rectifiers, A.I.E.E., Trans., Commun. & Electronics, **8,** pp. 488–492, Sept., 1953.

GRAY, A. N.[3]

Room-Temperature Compound Process, Mech. Eng., **75,** pp. 625–628, Aug., 1953.

[1] Bell Telephone Laboratories.
[3] Western Electric Company.
[4] Westrex Corporation.
[7] Sandia Corporation.

GROSSMAN, A. J., see C. H. ELMENDORF.

HAGSTRUM, H. D.[1]

Electron Ejection From Ta by He^+, He^{++}, and He_2^+., Phys. Rev., **91,** pp. 543–551, Aug. 1, 1953 (Monograph 2148).

Measurements of total yield (γ_i) and kinetic energy distribution are reported for electrons ejected from tantalum by the ions, He^+, He^{++} and He_2^+ in the kinetic energy range 10 to 1000 ev. The evidence presented indicates that the electrons are released by a collision of the second kind of the ion with the metal surface (potential ejection). One internal secondary electron is produced per incident ion. The probability of this electron escaping is reduced by the possibility of internal reflection at the image barrier at the metal surface. γ_i for the slowest ions is observed to be 0.14, 0.52, and 0.10 for He^+, He^{++}, and He_2^+, respectively. The data presented must be considered representative of gas-covered tantalum, since no gas was observed to desorb from the target on heating to 175°K. γ_i was found not to vary with time after cooling the target indicating rapid re-establishment of the equilibrium gas layer on the surface from within the metal. The work function of the covered Ta surface is found to be ca 4.9 ev, some 0.8 ev higher than that of atomically clean Ta. Considerations based on a theory which includes variation of energy levels near the metal surface show resonance neutralization of He^+ at the covered Ta surface not to be possible. Thus only the so-called direct process of potential ejection occurs, with which conclusion the measured energy limits of ejected electrons are in agreement.

HAUSER, F., see G. R. CRANE

HAWKINS, W. L., see B. S. BIGGS.

HEFFNER, H.[1]

Backward-Wave Tube, Electronics, **26,** pp. 135–137, Oct., 1953.

Electron-stream amplifier utilizing backward-wave mode forms microwave oscillator continuously tunable over a three-to-one bandwidth by a single voltage control. Tubes have been built for frequency centering about 6,000, 10,000 and 50,000 mc.

HENDERSON, O.[5]

Magnetic Amplifier Controls for Rectifier Protecting Underground Metallic Structures Cathodically, Corrosion, **9,** pp. 216–220, July, 1953.

This paper covers the initial attempt on the part of the Ohio Bell Telephone

[1] Bell Telephone Laboratories.
[5] Ohio Bell Telephone.

Company to use a magnetic amplifier to give continuous control of the output of a copper-oxide rectifier used for cathodic protection of underground lead covered cables. The controlled rectifier was designed for use at a location where stray current "end effects" were damaging underground telephone cables and municipal light cables and were threatening high-pressure water mains. When properly adjusted, the output of the rectifier will increase or decrease automatically so that at each instant the amount of forced drainage will be adequate to protect the underground telephone cable sheath but will not be in excess of the value at which neighboring underground metallic structures would become anodic and would thus become subject to corrosion. For satisfactory operation the amplifier had to be designed to give a large gain so that a change in the control voltage of only 0.2 volt (+0.1 to −0.1 volt) would be sufficient to vary the output of the rectifier from practically zero current to full rating. In the installation described in this paper the gain of the magnetic amplifier is in the order of 12,000. The magnetic amplifier type of control is well suited to outdoor installations subject to wide changes in temperature and at remote locations where frequent maintenance inspections are not feasible. The magnetic amplifier has advantages over other available control devices which accomplish this same purpose in that it has no moving parts, no vacuum tubes, batteries, motors, relays or contactors. The magnetic amplifier gives a continuous output control which is superior to the stepped increment changes that result from relay or contactor operation.

JERONE, M. G., see E. P. SMITH.

JOHNSON, J. B.[1] AND K. G. MCKAY[1]

Secondary Electron Emission of Crystalline MgO, Phys. Rev., **91,** pp. 582–587, Aug. 1, 1953 (Monograph 2163).

Secondary emission is measured from single crystals of MgO cleaved along the (100) plane. The maximum ratio of secondary to primary current, δ_{max}, is about 7 at about 1,000 volts and room temperatures. The cross-overs are at 33 volts and far above 5,000 volts. Most probable energy of emission is 1 ev or less. A definite effect of temperature is established, decreasing with increasing temperature, in accord with expectations for an insulator.

JONES, T. A.[1] AND W. A. PHELPS[1]

Level Compensator for Telephotograph Systems, Elec. Eng., **72,** pp. 787–791, Sept., 1953.

To eliminate interference in telephotograph transmission through broadband carrier equipment, it was decided to cancel it from the signal delivered by the carrier facility instead of modifying the carrier equipment. Conse-

[1] Bell Telephone Laboratories.

quently, a recently developed telephotograph level compensator, consisting of a pilot channel arrangement designed for insertion in the telephotograph connecting circuits, is utilized.

KISTLER, R. E.[6]

Radio Links U.S. and Canada, Telephony, **145,** pp. 16–17, 33, Sept. 5, 1953.

KLIE, R. H., see C. H. ELMENDORF.

KOHMAN, G. T., see D. EDELSON.

KRUSE, P. F., Jr.[7] and W. B. WALLACE[7]

Identification of Polymeric Materials, Anal. Chem., **25,** p. 1156, Aug., 1953.

KUH, E. S.[1]

Potential Analog-Network Synthesis for Arbitrary Loss Functions, J. App. Phys., **24,** pp. 897–902, July 1953.

A general method is developed for designing networks with arbitrary loss functions based on the potential analogy. An appropriate potential problem is formed on the basis of the given loss function by introducing continuous charge distribution on the complex frequency plane. After the potential problem is solved, the technique of quantization of charge is used to find the natural modes of the network function.

LANGE, R. W.[1]

40- to 4,000-Microwatt Power Meter, A.I.E.E. Trans., Commun. & Electronics, **8,** pp. 492–494, Sept., 1953.

LEWIS W. D.[1]

Electronic Computers and Telephone Switching, I.R.E., Proc., **41,** pp. 1242–1244, Oct., 1953.

Automatic telephone switching and digital computation have much in common. Both rely upon discrete rather than continuous devices. Development of recent switching systems with a close functional resemblance to large digital computers has increased this overlap. The next big step in

[1] Bell Telephone Laboratories.
[6] Pacific Telephone and Telegraph Company.
[7] Sandia Corporation.

telephone switching should be towards electronics. In making this step, switching scientists and engineers can be much helped by modern electronic computer technology. To be successful, they must also contribute to this technology.

LINDHOM, P.[3]

"Feedback" — Heart of Automatic Process Control, Factory, **111,** pp. 106–109, Oct., 1953.

LOGAN, R. A.[1]

Thermally Induced Acceptors in Single Crystal Germanium, Letter to the Editor, Phys. Rev., **91,** pp. 757–758, Aug. 1, 1953.

MAITA, J. P., see M. TANENBAUM.

MALLERY, P.[1]

Transistors and Their Circuits in the 4A Toll Crossbar Switching System, A.I.E.E. Trans., Commun. & Electronics, **8,** pp. 388–392, Sept., 1953.

MALTHANER, W. A.[1] AND H. E. VAUGHAN[1]

Automatic Telephone System Employing Magnetic Drum Memory, I.R.E., Proc., **41,** pp. 1341–1347, Oct., 1953 (Monograph 2151).

The use of magnetic drum memory in an automatic telephone switching office is described. A capacitive scanner acts as a time-division connector through which information generated by subscribers' telephone sets is conveyed to storage on magnetic drums. Information thus accumulated is combined with "permanent" information on the magnetic drums, processed in accordance with built-in programs and dispatched to control call switching circuits. Technical feasibility of this system has been demonstrated by the construction and successful operation of a large-scale laboratory model.

MANLEY, H. A., see G. R. CRANE.

MATTHIAS, B. T.[1]

Superconducting Compounds, Letter to the Editor, Phys. Rev., **91,** p. 413, July 15, 1953.

MCAFEE, K. B., see K. G. MCKAY.

[1] Bell Telephone Laboratories.
[3] Western Electric Company.

McCarthy, R. H.[3]

Organization for Production Engineering, Mech. Eng., **75,** pp. 785–788, 793, Oct., 1953.

McGuigan, J. H.[1]

Combined Reading and Writing on a Magnetic Drum, I.R.E., Proc., **41,** pp. 1438–1444, Oct., 1953 (Monograph 2152).

This paper points out that the characteristics of magnetic recording make it possible to combine reading and writing in the same cell as it passes just once under the head. Amplifier requirements for this method of operation are discussed and a suitable design presented. A single head is used for both reading and writing. The process can be repeated in every successive cell at a cell rate of 60 kc. The techniques described, which are applicable to either parallel or serial systems, extend the utility of magnetic drums by allowing date processing as well as data storage.

McKay, K, G., see J. B. Johnson.

McKay, K. G.[1] and K. B. McAfee[1]

Electron Multiplication in Silicon and Germanium, Phys. Rev., **91,** pp. 1079–1084, Sept. 1, 1953 (Monograph 2162).

Electron multiplication in silicon and germanium has been studied in the high fields of wide p-n junctions for voltages in the pre-breakdown region. Multiplication factors as high as eighteen have been observed at room temperature. Carriers injected by light, alpha particles, or thermal-generation are multiplied in the same manner. The time required for the multiplication process is less than 2×10^{-8} second. Approximately equal multiplication factors are obtained for injected electrons and injected holes. The multiplication increases rapidly as "breakdown voltage" is approached. The data are well represented by ionization rates computed by conventional avalanche theory. In very narrow junctions, no observable multiplication occurs before Zener emission sets in, as previously reported. It is incidentally determined that the efficiency of ionization by alpha particles bombarding silicon is 3.6 ± 0.3 electron volts per electron-hole pair produced.

McSkimin, H. J.[1]

Measurement of Elastic Constants at Low Temperatures by Means of Ultrasonic Waves — Data for Silicon and Germanium Single Crystals and for Fused Silica, J. App. Phys., **24,** pp. 988–997, Aug., 1953 (Monograph 2171).

[1] Bell Telephone Laboratories.
[3] Western Electric Company.

Ultrasonic waves (shear on longitudinal) in the 10-30 mc range are transmitted down a fused silica rod, through a polystyrene or silicone one-quarter wavelength seal, and into the solid specimen. Measurement of reflections within the specimen yields values for velocities of propagation and elastic constants. Data obtained over a temperature range of 78° to 300°K for silicon and germanium single crystals, and 1.6° to 300°K for fused silica are listed. For the latter, a high loss is noted, with an indicated maximum near 30°K.

MERZ, W. J.[1]

Double Hysteresis Loop of $BaTiO_3$ at the Curie Point, Phys. Rev., **91,** pp. 513–517, Aug. 1, 1953 (Monograph 2166).

It is known that the Curie point of the ferroelectric $BaTiO_3$ shifts to higher temperatures when a dc bias field is applied. If the crystal shows a sharp transition, we expect by applying an ac field at the Curie temperature that the crystal would become alternately ferroelectric and nonferroelectric in the cycle of the ac field. This can be seen in the shape of the hysteresis loop at temperatures slightly above θ. In the center of the polarization P versus field E plot, we observe a linear behavior corresponding to the paraelectric state of $BaTiO_3$ above θ. At both high voltage ends, however, we observe a hysteresis loop corresponding to the ferroelectric state. A change in temperature causes a change in size and shape of the double hysteresis loops, ranging from a line with curves at the ends (higher temperature) to two overlapping loops (lower temperature). The results obtained allow us to calculate the different constants in the free-energy expression of Devonshire and Slater. One of the results shows that the transition is of the first order since the P^4 term turns out to be negative. The properties of the hysteresis loops are discussed, especially the large spontaneous electrical polarization and the low coercive field strength.

MOORE, E. F., see C. E. SHANNON.

MORRISON, J.[1]

Leak Control Tube., Rev. Sci. Instr, **24,** pp. 564–547, July, 1953.

PHELPS, W. A., see T. A. JONES.

PRINCE, M. B.[1]

Experimental Confirmation of Relation Between Pulse Drift Mobility and Charge Carrier Drift Mobility in Germanium, Phys. Rev., **91,** pp. 271–272, July 15, 1953 (Monograph 2168).

Experimental data of drift mobilities of minority carriers in germanium are

[1] Bell Telephone Laboratories.

brought into agreement with theoretical predictions by distinguishing between group velocity and particle velocity of a pulse of minority carriers. Corrected high temperature measurements of electron drift mobility are consistent with the theoretical prediction $\mu = AT^{-3/2}$. The experimentally determined value of A is 2.0×10^7 cm^2 deg$^{3/2}$/volt-sec.

REISS, H.[1]

Chemical Effects Due to the Ionization of Impurities in Semiconductors, J. Chem. Phys., **21,** p. 1209, July, 1953 (Monograph 2172).

This paper contains a theoretical account of the possible effects which the ionization of donor and acceptor impurities can induce in the thermodynamic phase relations involving their solutions with semiconductors. These effects reflect the energy band structure of the semiconductor. Furthermore, the phenomenon has an interest of its own, for within a certain range of experimental conditions the effects can be attributed to a chemical-like, mass action behavior of the electrons which play the roles of negative ions. Section V is a brief discussion of a fine point concerning the Fermi level. It is shown that although the Fermi level is certainly the electronic electrochemical potential, it is not the Gibbs free energy per electron unless the density of electron energy levels is linear in the volume of the system.

ROBBINS, R. L.[1]

Measurement of Path Loss Between Miami and Key West at 3675 mc., I.R.E., Trans., P.G.A.P. **1,** pp. 5–8, July, 1953.

Radio transmission measurements have been made at 3,675 megacycles on the 130-mile path between Miami and Key West, Florida, which is largely over watery wastes of the Everglades and shallow sea waters of the Florida Keys. Path loss and fading characteristics for this terrain were not found to differ materially from the characteristics of hilly or mountainous paths in the northeastern section of the country.

ROSS, I. M., see G. C. DACEY.

SHANNON, C. E.

Turing's Formulation of Computing Machines and Von Neumann's Models of Self-reproducing Machines., I.R.E., Proc., **41,** pp. 1235–1241, Oct., 1953 (Monograph 2150).

This paper reviews briefly some of the recent developments in the field of automata and non-numerical computation. A number of typical machines are described, including logic machines, game-playing machines and learning machines. Some theoretical questions and developments are discussed, such as a comparison of computers and the brain.

[1] Bell Telephone Laboratories.

SHANNON, C. E.[1] AND E. F. MOORE[1]

Machine Aid for Switching Circuit Design, I.R.E., Proc., **41,** pp. 1348–1351, Oct., 1953 (Monograph 2153).

Design of circuits composed of logical elements may be facilitated by auxiliary machines. This paper describes one such machine, made of relays, selector switches, gas diodes, and germanium diodes. This machine (called the relay circuit analyzer) has as inputs both a relay contact circuit and the specifications the circuit is expected to satisfy. The analyzer (1) verifies whether the circuit satisfies the specifications, (2) makes systematic attempts to simplify the circuit by removing redundant contacts, and also (3) obtains mathematically rigorous lower bounds for the numbers and types of contacts needed to satisfy the specifications. A special feature of the analyzer is its ability to take advantage of circuit specifications which are incompletely stated. The auxiliary machine method of doing these and similar operations is compared with the method of coding them on a general-purpose digital computer.

SMITH, E. P.[6], L. P. CORNELL[6] AND M. G. JEROME[6]

Co-ordinating M1 and N1 Telephone Carrier Systems, Elec. Eng., **72,** p. 780, Sept., 1953.

TANENBAUM, M.[1] AND J. P. MAITA[1]

Hall Effect and Conductivity of InSb Single Crystals, Letter to the Editor, Phys. Rev., **91,** pp. 1009–1010, Aug. 15, 1953.

TEMPLIN, E. W., see J. G. FRAYNE.

TOWSLEY, L. M., see E. A. WOOD.

TUCKER, C. J., JR.[8]

Emergency Reporting System Installed by Southern Bell, Telephony, **145,** pp. 26, 38, Aug. 29, 1953.

New fire and emergency reporting system utilizing telephone facilities for the general public to make verbal reports of fires and other emergencies — the first of its kind in the nation — was put into service in Miami, Fla., on Aug. 1, by Southern Bell Telephone & Telegraph Co.

VAN ROOSBROECK, W.[1]

[1] Bell Telephone Laboratories.
[6] Pacific Telephone and Telegraph Company.
[8] Southern Bell Telephone Company.

Transport of Added Current Carriers in a Homogeneous Semiconductor, Phys. Rev., **91,** pp. 282–289, July 15, 1953 (Monograph 2165).

Taking into account the thermal equilibrium minority carrier concentration and employing the formulation which includes, as one of two fundamental equations, the continuity equation for added carrier concentration Δp, this equation is derived in a form which exhibits the ambipolar nature of the diffusion, drift, and recombination mechanisms under electrical neutrality. The general concentration-dependent diffusivity is given. The local drift velocity of Δp has the direction of total current density in an n-type semiconductor and the reverse in a p-type semiconductor, differing in general in both magnitude and direction from the minority-carrier drift velocity. Specifying a model for recombination fixes the dependence of a lifetime function for Δp on Δp and the electron and hole mean lifetimes. Negative Δp, or carrier depletion with electrical neutrality, may occur. For known total current density, the continuity equation alone suffices, as for the case of $|\Delta p|$ small, for which the equation is linear. A condition for this comparatively important case is derived, and theoretical relationships are given with the aid of a parameter specifying the Fermi level, which determine for germanium the minority carrier-Δp drift velocity ratio as well as the ambipolar diffusivity and group mobility in terms of resistivity and temperature.

VAUGHAN, H. E., see W. A. MALTHANER.

VOGELSONG, J. H.[1]

Transistor Pulse Amplifier Using External Regeneration, I.R.E., Proc., **41,** pp. 1444–1450, Oct., 1953.

Pulse-regenerative amplifier using a point-contact transistor has been operated at a basic frequency of 3 megacycles. To produce regenerated pulses with waveshapes which are practically independent of the waveshapes of the input pulses, a germanium diode circuit has been used in conjunction with an external feed-back path. This arrangement also provides for the synchronization of the output pulses with a master clock. Transformer coupling has been incorporated into the circuit to provide dc restoration.

WALKER, A. C.[1]

Hydrothermal Synthesis of Quartz Crystals, Am. Ceramic Soc., J., **36,** pp. 250–256, Aug., 1953 (Monograph 2146).

Research at the Bell Telephone Laboratories on the problem of growing large single crystals of quartz has now progressed to a point where it is possible to grow crystals weighing more than 1 lb. each in a period of 60 days or less. Equipment now in use includes autoclaves 4 inches in inside diameter and 4 ft. long, weighing about 1,150 lb. each. In developing the

[1] Bell Telephone Laboratories.

hydrothermal process used to grow these quartz crystals it has been necessary to solve many problems in the little-known field of high pressure. The results point to the possibility of growing other types of crystals, and the field of usefulness of this process now appears to be much more extensive than was the case at the beginning of the investigation in 1946. Of prime importance is the fact that crystals grown from solution are likely to be better formed and of more perfect quality than those grown from the melt or by other methods. Many of the difficulties inherent in this work have been due to corrosion of steel in alkaline solution. This was a rather unexpected problem, since in high pressure steam boilers alkali is added in small amounts to prevent corrosion of the boiler tubes. Such corrosion has been shown to be responsible for the appearance of electrical twinning on the growing faces of the quartz crystals. Other causes of such twinning have also been found in the course of this work.

WALKER, L. R.[1]

Starting Currents in the Backward-Wave Oscillator, J. App. Phys., **24,** pp. 854–859, July, 1953.

The starting current of a simple model of the backward-wave oscillator described by Kompfner and Williams has been calculated. The effect of space charge is included. The starting current I_0 may be written in the form

$$\frac{4V_0}{Z_0}\left[\frac{a_0(4QC)}{2N}\right]^3,$$

where V_0 is the beam voltage, Z_0 is the impedance of the circuit, N is the length of the oscillator in wavelengths measured on the circuit and $a_0(4QC)$ is a dimensionless quantity which has been evaluated as a function of the space-charge parameter $4QC$.

WALLACE, W. B., see P. F. KRUSE, JR.

WASHBURN, S. H.[1]

Application of Boolean Algebra to the Design of Electronic Switching Circuits, A.I.E.E. Trans., Commun. & Electronics, **8,** pp. 380–388, Sept., 1953.

WERNER, D. R.[2]

Effects of Polarization on Telephone Cable Buried Through a Salt Bed, Corrosion, **9,** pp. 232–236, July, 1953.

Cathodic protection has been applied to a copper jacketed cable in a salt lake bed about one mile wide in which the earth resistivity was apparently uni-

[1] Bell Telephone Laboratories.
[2] American Telephone and Telegraph Company.

form at about 20 ohm-centimeters. Current of about one ampere flowed on the copper jacket into the low resistivity areas on either side and was found to be sustained by polarization effects when the cathodic protection current was removed. The current loss on the copper jacket was found to be concentrated in an area about 720 feet wide, 360 feet each side of the point where the cathodic protection current had been drained from the copper jacket. The copper jacket to soil potential tested most negative to a copper sulfate half cell in the 720-foot area where the current loss was concentrated and was of the order of -1.0 to -1.1 volts. Permanent remedial measures will consist of installations of magnesium anodes distributed throughout the low earth resistivity area and insulating joints in the copper jacket at locations where large changes in the earth resistivity occur.

WILLIAMS, H. J.[1], R. M. BOZORTH[1] AND M. GOERTZ[1]

Mechanism of Transition in Magnetite at Low Temperatures, Phys. Rev., **91,** pp. 1107–1115, Sept. 1, 1953 (Monograph 2149).

When magnetite is cooled through $-160°C$ it is known to undergo a transition (cubic to orthorhombic) that is influenced by the presence of a magnetic field. Our experiments are in agreement with the following mechanism of the transition: The orthorhombic c axis is parallel to one of the original cubic axes and is the axis of easiest magnetization. Generally, different regions of the original crystal will transform with their c axes lying along different cubic axes, and when no field is applied there are 6 different orientations which different regions assume. When a field is applied during a cooling a c axis tends to lie along the original cubic axis that is nearest to the applied field, the a and b axis having less but different tendencies to lie parallel to the field. Six magnetic crystal anisotropy constants are derived from torque curves measured in the (100) and (110) planes. From them magnetization curves are calculated for the 100 and 110 directions, and these are in agreement with experiment.

WOOD, E. A.[1] AND L. M. TOWSLEY[1]

Manganese Film-Shield for FeK X-Rays, Rev. Sci. Instr., **24,** p. 547, July, 1953.

[1] Bell Telephone Laboratories.

Contributors to this Issue

ARTHUR C. KELLER. B.S., Cooper Union, 1923; M.S., Yale University, 1925; E.E., Cooper Union, 1926; Columbia University, 1926–30; Western Electric Company, 1917–25; Bell Telephone Laboratories, 1925–. Special Apparatus Development Engineer, 1943; Switching Apparatus Development Engineer, 1946; Assistant Director of Switching Apparatus Development, 1949; Director of Switching Apparatus Development, 1949–. Mr. Keller's experience in the Bell System includes development and design of telephone instruments; development of systems and apparatus for recording and reproducing sound; and, during World War II, the development, design, and preparation for manufacture of sonar systems and apparatus. His department, in addition to being responsible for a number of military projects, is responsible for the fundamental studies of switching apparatus and the development, design, and preparation for manufacture of electromagnetic and electromechanical switching apparatus for telephone systems. Member of the American Physical Society, A.I.E.E., Acoustical Society of America, I.R.E., S.M.P.T.E., and the Yale Engineering Association. Representative for Bell Telephone Laboratories in the Society for Experimental Stress Analysis. For his contributions to the Navy during World War II, he received awards from the Bureau of Ships and the Bureau of Ordanance.

MASON A. LOGAN, B.S. in Physics and Engineering, California Institute of Technology, 1927; M.A. in Physics, Columbia University, 1933; Bell Telephone Laboratories, 1927–. His early Laboratories' projects were concerned with wire transmission problems particularly those of circuits, noise and cross induction in local, manual and dial telephone circuits. This was followed by circuit research on alternating current methods of signaling including the use of non-linear elements and electronic terminal equipment. From 1941 to 1948 he worked on military projects, including a mine fire control system, anti-aircraft gun director, magnetic proximity fuses, and guided missiles. For the past five years he has been a member of the Switching Apparatus Development Department in which he is supervising a group concerned with static and dynamic behavior of

new electromagnets and relays. He is also engaged in investigations of the performance of electrical contacts on telephone relays.

ROBERT L. PEEK, JR., A.B. and Met.E., Columbia University, 1921 and 1923; Bell Telephone Laboratories, 1924–. In the Chemical Research Department and later the Apparatus Development Department, Mr. Peek's work related to the analytical and testing aspects of materials development. Since 1936 he has been engaged in apparatus development projects, including the wire spring relay and, during World War II, underwater ordnance and magnetostriction sonar.

H. N. WAGAR, B.S. in Physics, Harvard University, 1926; M.A. in Physics, Columbia University, 1931; Bell Telephone Laboratories, 1926–. Mr. Wagar has worked on the design of nearly all types of telephone relays as well as magnets, insulating methods and allied apparatus and practices. He was also associated with the preparation of text and presentation of training courses in this field. During World War II his projects for the military included work on an antiaircraft director and a proximity fuse for magnetic mines. He is currently Switching Apparatus Engineer, in charge of fundamental studies on electromagnetic switching apparatus, including contacts and magnetics.