UNIVERSAL
LIBRARY

INTRODUCTION TO

EXPERIMENTAL PHYSICS

PRENTICE-HALL PHYSICS SERIES

Dr. Donald H. Menzel, Editor

INTRODUCTION TO EXPERIMENTAL PHYSICS

By

WILLIAM B. FRETTER, Ph.D.

Associate Professor of Physics
University of California

New York
PRENTICE-HALL, INC.
1954

Library of Congress Catalog Card Number: 54-6004

PREFACE

The techniques of modern experimental physics, which increased in number and complexity under the impact of the wartime emergency, are again being applied to research in physics. In an attempt to introduce our students to some of these techniques, a course has been instituted at the University of California in which lectures are given on the various aspects of experimental procedures. The lectures are presented by experts who are thoroughly familiar with the experimental details of the subject. This book has been compiled from notes taken on the lectures given in this course during the first years of its existence. The short, condensed presentation is intentional, so that graduate students might find time to read the book, even if they cannot take the course. The reader should acquire a *vocabulary* in experimental physics, and the book is directed toward that end.

It was clear from the beginning that any such introduction to experimental physics could not be complete. There is not time enough during the one-year course, or space enough in this book, to include the vast amount of existing experimental knowledge of techniques. A "Handbook of Experimental Physics" would require many large volumes and, although some excellent books on experimental physics have appeared recently, they are limited to particular fields of research and most of them require advanced knowledge to understand them.

A large part of this book is devoted to atomic and nuclear physics, and it is natural that the emphasis is on the research problems studied at Berkeley. Since a very considerable proportion of young physicists currently go into these fields on graduation, the author believes that this selection of topics is justified. The techniques basic to all research in physics, i.e. measurements, electronics, vacuum techniques, and magnet design are also discussed. There is a chapter on laboratory hazards and another on the design of experiments. Throughout we have tried to include as many "tricks of the trade" as possible, which rarely appear in scientific journals. The illustrations are mostly adapted from blackboard sketches made to illustrate a point and, in general, the lecture style of presentation is preserved.

The author appreciates the cooperation and help of his colleagues in preparing the manuscript. He is also much indebted to Professors Menzel, Schmitt, and Havens, who read the manuscript and made many helpful suggestions. A list of contributors follows: L. W. Alvarez, V. Ashby, W. Blocker, A. L. Bloom, W. Boggs, H. Bradner, A. Brate-

nahl, W. Birnbaum, W. Cartwright, O. Chamberlain, R. L. Chasson, O. El-Mofty, R. Evans, F. L. Fillmore, R. H. Fox, T. Geballe, W. E. Glenn, J. D. Gow, J. W. Hadley, R. W. Hales, A. C. Helmholz, F. Hurlbut, A. J. Hudgins, F. A. Jenkins, W. D. Knight, A. F. Kip, E. Lilge, E. M. McMillan, R. G. Mills, B. J. Moyer, R. Mozely, W. A. Nierenberg, M. Packard, W. K. H. Panofsky, W. M. Powell, R. V. Pyle, J. H. Reynolds, A. Rosen, L. Schecter, A. Schulz, E. Segrè, J. S. Steller, O. Stern, M. Summerfield, D. H. Templeton, R. L. Thornton, J. F. Tracy, R. Wallace, H. E. White, R. Wright, H. F. York.

WILLIAM B. FRETTER

CONTENTS

*

INTRODUCTION TO

EXPERIMENTAL PHYSICS

1

ELECTRICAL MEASUREMENTS

NOISE

THERMAL NOISE. Consider a generator together with its internal impedance, as a source, transmitting power to a load. Reasoning based on the second law of thermodynamics shows that the power generated by thermal noise across the terminals II is $P \approx kT\Delta\nu$, where k is the Boltzmann gas constant, T is the absolute temperature, and $\Delta\nu$ is the bandwidth of the frequencies employed.

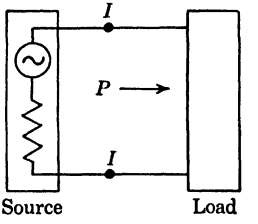


Fig. 1. Equivalent noise source, transmission line, and load.

To give an approximate derivation of this important relationship, let us imagine that we replace the source of power by a resonant transmission line of length L. The resonant modes of oscillation of such a transmission line have wavelengths λ, where $n\lambda = 2L$, or frequencies ν, where $nc = 2L\nu$. The total number of modes is $n = (2L/c)\nu$ and the number of resonant modes in a given wavelength range $\Delta\nu$ is $\Delta n = (2L/c)\Delta\nu$. Since each oscillating mode represents kT of energy, we can obtain the desired result by multiplying both sides of the equation by kT and rearranging the terms. Then

$$\frac{kT\Delta n}{2L/c} = kT\Delta\nu$$

where the quantity on the left side is equivalent to the power developed at the terminals TT, since $2L/c$ is the mean propagation time along the transmission line and $kT\Delta n$ represents the energy per mode times the number of modes propagated. Thus the energy per unit time, or power, $P \approx kT\Delta\nu$.

Since low thermal noise power is concomitant with low absolute temperature and narrow bandwidth, it is important to design the bandwidth as narrow as possible, compatible with the proposed measurement, if the noise is to be reduced to a minimum.

Another important fundamental limitation involving the bandwidth is that the product of Δt, the time available for measurement, and $\Delta\nu$ is of the order of unity, i.e., $\Delta t\Delta\nu \approx 1$. For example, if we

are measuring pulses of time width $\Delta t = 1$ μsec (microsecond), then according to the relation $\Delta t \Delta \nu \approx 1$, a bandwidth $\Delta \nu = 1$ megacycle is required.

Although the approximate theoretical thermal noise limit cannot actually be improved on, one method of apparently doing so is useful when the pulse or signal is periodic. If the sweep frequency of an oscilloscope is the same as the periodic frequency of the signal, the oscilloscope is tuned to the signal, but not to the noise. Since several successive sweeps must be observed, however, the bandwidth is thereby reduced and the effective signal-to-noise ratio increased. In such a case the oscilloscope acts as a narrow-band easily tuned filter, and the observed noise can never be reduced to less than the theoretical minimum when referred to the bandwidth of the oscilloscope sweep.

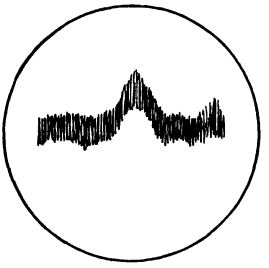


Fig. 2. Signal and noise on oscilloscope.

In applications to vacuum tube circuits a minimum value for the equivalent noise impedance appearing in series with the grid of the tube can be taken to be 200 ohms.[1] Accordingly, the noise voltage developed across this resistance as a function of bandwidth is $V^2 \approx 200kT\Delta\nu$. In the previous example of a 1 μsec pulse being passed by an amplifier having the required bandwidth of 1 megacycle the noise voltage would amount to approximately 1 μv (microvolt). In practice, the actual noise found in a single stage amplifier falls within a factor of 2 of that predicted by the above considerations, and for a multiple stage amplifier it is usually within a factor of 5 or 10.

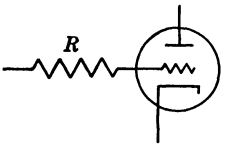


Fig. 3. Equivalent noise impedance with a vacuum tube.

FLUCTUATION NOISE. When the measurement of a small current occurs within a sufficiently short time interval, we must take into consideration the fact that the finite number of charges constituting the current pass a given point in a statistically arbitrary way. The fluctuation in the number of charges counted in a short time interval Δt, i.e., $n = i\Delta t/e$, is $\delta n = \sqrt{n}$, and the corresponding current fluctuation is $\Delta i = \delta_n e/\Delta t$. In terms of Fig. 1, the power generated across the terminals *II* due to fluctuation noise is $P = (\Delta i)^2 R = ieR/\Delta t =$

[1] This figure may vary widely with different tube types and with the circuit used. A table of equivalent noise resistances of various receiving tubes is given by Twiss and Beers in Valley, G. E., and Wallman, A., *Vacuum Tube Amplifiers*. New York: McGraw-Hill Book Co., Inc., 1948, p. 636.

$Ve\Delta\nu$, so that Ve now appears in place of the kT of the thermal noise expression.

Both thermal noise and fluctuation noise are found in any amplifier, and in general the two types are indistinguishable. Under ordinary conditions the amount of thermal noise is a fixed quantity determined by the temperature and properties of the circuit. When the noise measured is appreciably greater than this, the predominant source is often fluctuation noise. The actual rms voltage of the thermal noise depends not only on the temperature, but also on the input resistor and the bandwidth. The actual thermal noise is

$$\overline{e_{\text{th}}^2} = 4kTR\Delta\nu$$

and at $T = 290°\text{K}$,

$$e_{\text{rms}} = 1.26\sqrt{R\Delta\nu} \times 10^{-10} \quad \text{v}$$

TIME-BANDWIDTH PRODUCT

The relation $\Delta t\Delta\nu \approx 1$ used several times in the above discussion has its theoretical basis in the fact that a Fourier integral of a Gaussian function of the time gives a Gaussian function of the frequency. If the wave shape under consideration is known and reproducible, we can effectively gain a factor of 10 over the uncertainty product above, i.e., $\Delta t\Delta\nu \approx 0.1$.

For example, it was recently desired to measure the coincidence between two events A and B, the latter being delayed relative to the former by approximately 10^{-8} sec. The relation above states that a 100 megacycle bandwidth should be required, yet the measurement was carried out with only a 20 to 30 megacycle bandwidth. The experimenters plotted the coincidence rate vs. the delay time, and found the resulting curve askew because of the finite probability that pulse A would occur later than pulse B. Under the most favorable circumstances then, $\Delta t\Delta\nu \approx 0.1$.

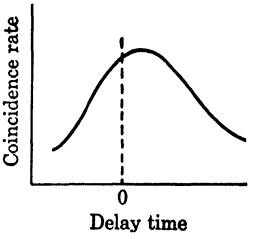


Fig. 4. Coincidence curve.

Reversing the question of what bandwidth is required to pass a given pulse width, we can inquire how long it takes to measure a given frequency to a given accuracy, and the above relation states, for example, that to measure a 1 megacycle frequency to 1 part per million requires a time of measurement $\Delta t \approx 1$ sec.

GAIN-BANDWIDTH PRODUCT

In the design of a broad-band amplifier (which would be necessary in making a measurement involving pulses of short duration), we en-

counter the limitation $G2\pi\Delta\nu = g_m/C$, where G is the amplifier gain, $\Delta\nu$ its bandwidth, g_m and C the tube transconductance and plate-cathode capacitance, respectively. The maximum gain is $G = g_mZ$, where Z is the plate load impedance. The plate load impedance cannot, however, be increased without limit because the capacitance of the tube and stray circuit capacitances become the limiting impedances, rather than Z, as Z is increased. The maximum load impedance that can be used is given approximately by $Z_{max} = 2/\omega C(\omega = 2\pi\nu)$, and therefore the maximum gain becomes $G = 2g_m/\omega C$. Thus the broader the bandpass of an amplifier the lower is its gain. The largest bandpass for the best small sized tubes available is approximately 50 megacycles (i.e., for unity gain).

We conclude that measurement of events of short duration requires large bandwidths with accompanying sacrifices in gain and increased noise. Paralleling of ordinary tubes in the usual way increases both g_m and C and provides no remedy. However, the recent development of a distributed amplifier (Chapter 3) allows tubes to be paralleled and the gain-bandwidth product to be increased. Without presenting a detailed explanation, we may say that each tube "sees" only its own capacitance at the time it passes the signal and that, although the gain per tube is less than unity, the over-all gain is greater than unity, and amplifiers with bandwidths of 250 megacycles can be constructed.

One further remark must be made in regard to large bandwidths. The required output current is sometimes quite high, as will be seen from the following example. Suppose that a 100 v output is ultimately desired so that a large oscilloscope response may be observed, and that the signal is such as to require the 250 megacycle bandwidth of the traveling-wave amplifier. Then for a reasonable oscilloscope deflecting plate capacitance of 30 $\mu\mu$f the required driving current amounts to $i = q/\Delta t = CV/\Delta t = CV\Delta\nu = 30 \times 10^{-12} \times 100 \times 250 \times 10^6 = 0.75$ amp, and a power amplifier tube must be used in the output stage.

CHOICE OF MEASURING INSTRUMENT FOR SMALL VOLTAGE AND CURRENT

IMPEDANCE REQUIREMENTS FOR GALVANOMETERS. In elementary physics students learn to use a high-impedance device for measuring voltage and a low-impedance device for measuring current. However, for measurement of very small voltages or currents, just the reverse is true, since power is required to operate the instrument. For example, if the number of turns per unit cross section is changed in a galvanometer coil, the quantity that is held constant for a given deflection is the

power (i.e., I^2R or E^2/R). Hence high current sensitivity requires a high-resistance coil, and high voltage sensitivity requires a low-resistance coil.

For the most sensitive galvanometer made the power per unit deflection is approximately 10^{-19} w. Thus the best voltage sensitivity is about 10^{-9} v (for $R \approx 10$ ohms) and the best current sensitivity is about 10^{-11} amp (for $R \approx 1000$ ohms).

ELECTRONIC TYPE INSTRUMENTS. Although it is frequently more convenient to use an electronic device than an electromechanical one, the basic limitations of the electronic instrument should be considered in making a choice between the two types. The stability, for example, of a d-c amplifier used to measure small voltages is approximately 50 μv over short periods of time, and varies from 1 to 10 mv over periods of a day or so. In the a-c amplifier, as mentioned before, the primary limitation upon sensitivity is the noise, which at room temperature cannot be reduced much below the order of a microvolt. The a-c amplifier in conjunction with a vibrating reed or diaphragm[2] for the conversion of direct current to alternating current can be employed to measure direct voltages in the microvolt range. The galvanometer, however, remains preferable to an electronic device for extremely low voltage measurements.

For measurement of currents smaller than about 10^{-10} or 10^{-11} amp, a galvanometer becomes impractical, and an electrometer, which is essentially a voltage measuring instrument with a sensitivity of the general order of 1 mv per scale deflection, can be used in conjunction with a shunt resistor of 10^{12} ohms to obtain a sensitivity of the order of 10^{-15} amp. Such resistors are readily available commercially and are best calibrated by measurement of the time constant of the resistor and a known capacitance (as determined by a capacitance bridge, say). Since these very high resistors are usually nonlinear, it is not advisable to perform the calibration by the simple method of using a 100 v source and a galvanometer when the resistor is ultimately to be used in the millivolt region. The accuracy of the small current measurement depends critically upon a careful calibration of the resistor being employed, both in the electrometer method described above and the one following.

Use of an electronic device such as a d-c feedback amplifier is practical down to currents of the order of 10^{-12} amp. The instrument stability is here dependent upon so-called passive components, i.e., the external resistances, regardless of the general instability of the amplifier components.

[2] Pavlevsky, H., Swank, R. K., Grenchik, R., *Rev. Sci. Instr. 18*, 298 (1947).

A simple nonfeedback type electronic amplifier consists of an electrometer type tube, the grid characteristics of which are shown in Fig. 6. The grid current i_o for large bias voltages is made very small by proper tube design, and may be about 10^{-16} amp for the best tubes made and about 10^{-14} amp for some of the smaller tubes. By such methods as reducing the filament temperature, etc., the i_o of an ordinary tube can be made as small as 10^{-8} or 10^{-9} amp. It should be mentioned that when an electrometer circuit is used in the 10^{-16} amp

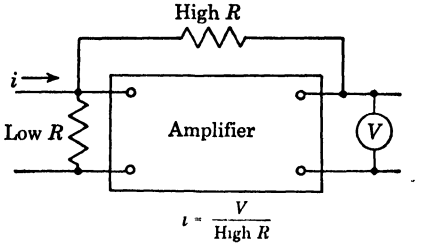


Fig. 5. Feedback amplifier.

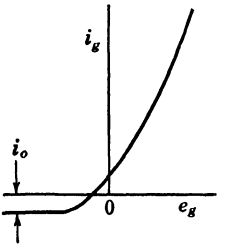


Fig. 6. Grid characteristic.

range, proper precautions must be taken to isolate the equipment from mechanical vibration, to which it usually is sensitive.

MEASUREMENT OF HIGH VOLTAGE

DIRECT CURRENT MEASUREMENTS. The problem is to construct a suitable voltage divider having small power dissipation and corona losses, and with the divider ratio calibrated to the desired accuracy. For example, we might wish to measure a voltage of 1 kv while having available only the means to measure 1 v accurately.

To keep the power in the divider small (say 1 w), resistance of 1 megohm would be required, and this is about the highest wirewound resistor available commercially. If still higher voltages are to be measured, if the power dissipation is required to be smaller, or if the voltage gradient is greater than the rated value for the resistance unit employed, a number of units must be connected in series, thus requiring a long string of resistors (frequently several feet long).

If corona currents are allowed to flow through the parts of the divider that are at higher potential and sprayed into the surrounding air, the effective dividing ratio will be increased by an unknown and variable amount, and the voltages indicated will be too low. The corona can be reduced by electrostatic shielding of the resistor units and still further by immersion of the divider in oil if the voltage

gradient is not too great. Another method is to surround the divider column by a cylindrical insulating can coated with aquadag so that each resistor of the divider sees only a small field and thus no corona current exists.

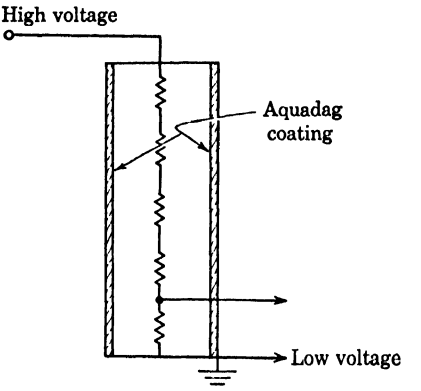


Fig. 7. High voltage divider.

A method for obtaining a divider ratio of 10^4 to 1 with the extremely high accuracy of one part in 10^7 has been described in the literature[3] and will serve as an example for precision divider calibration. Two especially constructed resistance boxes, one containing 10 sections of 1000 ohm manganin coils and the other containing 10 sections of 100,000 ohm coils were employed in series parallel combinations as follows:

Let: R_s = 10^5 ohm coils in series = 10^6 ohm
R_p = 10^5 ohm coils in parallel = 10^4 ohm
r_s = 10^3 ohm coils in series = 10^4 ohm
r_p = 10^3 ohm coils in parallel = 10^2 ohm

The ratio of the almost equal resistance combinations R_p and r_s can be determined with a Wheatstone bridge by an interchange method to an accuracy of 1 part in 10^6, with care being taken to eliminate the effects of contact resistance and cross leakage, to determine d, where $R_p/r_s = 1 + d$. Analysis then leads to the result

$$\frac{R_s}{r_p} = 10^4(1 + d)$$

With this procedure, the errors in individual resistor values cancel out to the first order, and enter only in the second order. For example, if each resistor is accurate to 1%, the value of R_s/r_p is accurate to 1 part in 10^4. The whole method can be applied to individual resistor groups to obtain even higher divider ratios if required.

ALTERNATING CURRENT MEASUREMENTS. A further problem arises in measuring pulsed voltages in that the same division ratio must appear to all frequency components of the pulse if distortion of the waveform is to be avoided. Since capacitance leakage shunts the resistors, the division ratio may be smaller for the higher frequency

[3] Wenner, *J. Research Nat. Bur. Standards, 25*, 229 (1940), RP 1323.

components, i.e., for the fast rise time at the beginning of a pulse, with the resultant distortion as shown.

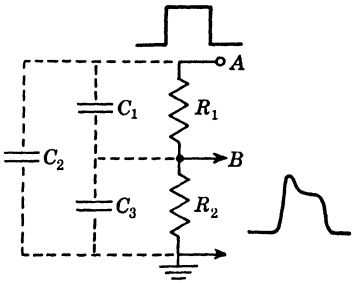


Fig. 8. High voltage divider equivalent circuit.

A capacitance compensated divider can be made to have no frequency dependence at moderate frequencies by adjustment of the shunt capacitances (including that of the output cable at B) by definite amounts such that $R_1C_1 = R_2C_2$. If the compensation is made in C_2 only, the whole divider can be placed in a shielding enclosure to maintain C_1 constant. The capacitance C_3 from point A to ground merely loads the source and does not affect the measurement otherwise. If the corona leakage is being corrected by the aquadag shielding method, the distributed capacitance along the resistor column will at the same time be minimized with reduction of the electric fields, and practically no distributed capacitance currents will flow.

The above method is sometimes useful in connection with the high-gain feedback amplifier previously discussed, where we desire to minimize the distributed-capacitance currents flowing off the 10^{11} ohm resistor; the aquadag shielding method suffices.

For rough measurements, an *electrostatic voltmeter* may provide the best solution to the problem. The accuracy obtainable is 5% or better, and such instruments can be used up to the 200 kv range. In an electrostatic voltmeter, the voltage is determined by measurement of the electrostatic force on a disk surrounded by a guard ring (Fig. 9a).

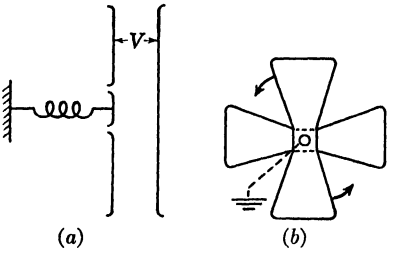


Fig. 9. Electrostatic voltmeters.

A *generating voltmeter* is often used for measurements of direct voltages from 100 kv to several million volts. It consists of a pair of vanes in the electrostatic field, one of which rotates and is grounded (Fig. 9b), alternately shielding and exposing the stationary vane to the field. The alternating voltage induced in the stationary vane is proportional to the field and can be either rectified and displayed on a d-c meter, or fed into an a-c amplifier.

The potentiometer and standard cell arrangement for direct potential measurement customarily gives an accuracy of 10^{-4} (i.e., the standard cell potential can be relied upon in the fourth figure)

provided the cell has been certified at the Bureau of Standards within a year or so and has received good treatment. An accuracy of 10^{-3} can be obtained with practically no worry about standards, and 10^{-5} is possible but requires the care of a small research problem.

MEASUREMENT OF HIGH CURRENT

The usual method for d-c measurement employs a shunt resistor with a millivoltmeter or galvanometer. For high currents, the galvanometer terminals must be located far from the other terminals so as to avoid the potential drop in the high current connections. Commercial shunts are always constructed in this way (Fig. 10).

In the measurement of pulsed currents, an oscilloscope is used in place of a galvanometer; the response is adjusted to the order of 100 v by appropriate choice of shunt resistance. The inductance of the

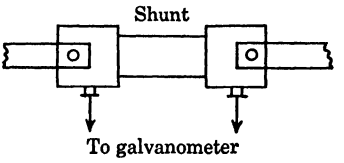


Fig. 10. Shunt.

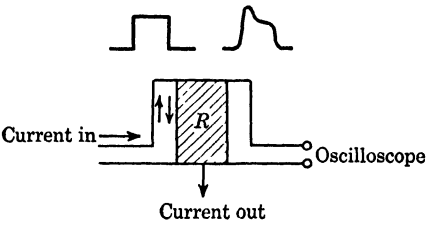


Fig. 11. High frequency shunt.

shunt, or more specifically its time constant L/R, must be short relative to the pulse length, and this factor presents a problem in the case of very short pulses. The distortion due to the inductance takes the form of an overshoot at the beginning of the pulse, as shown in Fig. 11. This condition is corrected by forming a low characteristic impedance coaxial transmission line about the shunt resistor as center conductor. The time constant can be reduced to less than 0.01 μsec by such a device.

When an oscilloscope is used as an indicator for pulse current or voltage measurements, the signal should be applied directly to the vertical deflection controls of the oscilloscope, and the oscilloscope should be connected to the alternating current through a line voltage regulator to insure that the deflection is independent of line voltage fluctuations during a measurement.

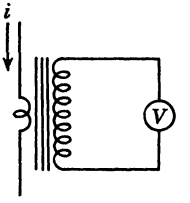


Fig. 12. Current transformer.

For pure a-c measurements, a current transformer is usually preferable to a shunt, and the commercially available transformers can be relied upon to 0.5%.

AMPLIFIER PICKUP AND INSTRUMENT SHIELDING

The problem of pickup by electronic instruments, etc., of radiation from extraneous sources frequently is a formidable one. Shielding of the equipment alone is not sufficient, since the undesired signal can be carried into the equipment via the power lines. The usual method of inserting a multiple section filter in the incoming power leads is satisfactory for filtering out the lower frequencies; for the high frequencies, however, the inductive reactance of the filter ground lead *GG* might

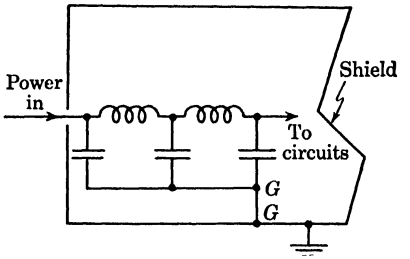


Fig. 13. Multiple section filter and shield.

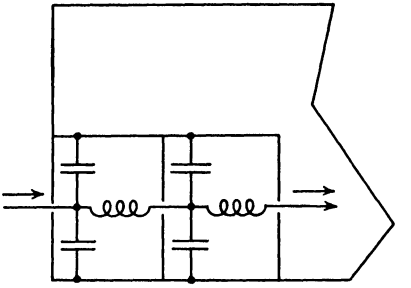


Fig. 14. Shielded divided filter.

spoil the filtering action. This inductance, for example, may easily amount to 0.02 μh (microhenry) which for a 10 megacycle signal presents a 1 ohm impedance; hence it can be seen that merely having a low capacitive filter reactance is not sufficient. The filter sections must be individually shielded with no common ground return paths, to prevent coupling between them, as in Fig. 14.

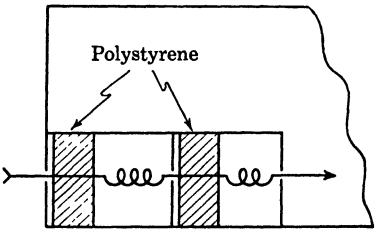


Fig. 15. Filter with low inductance capacitors.

At still higher frequencies the capacitors, together with their lead inductances, have series resonant frequencies ranging from 20 megacycles upward (for commercial units) and it becomes advisable either to buy or to construct of polystyrene, as in Fig. 15, low inductance capacitors. Capacitors so constructed become resonant only at frequencies for which their radial dimension approaches an electric quarter wavelength, and these are ordinarily far higher than those which it is desired to filter.

Analogously, inductors have distributed capacitance, and thus they too have natural resonances. Since larger inductors have lower resonant frequencies, the most desirable filter network is one con-

structed of many sections, each of which has a relatively smaller inductance and hence a higher resonant frequency.

It frequently happens that a piece of sensitive apparatus receives its signal or trigger pulse from a separate generating device connected by a shielded or coaxial cable as in Fig. 16. There are several points to be mentioned about the grounding of such an arrangement. If both pieces of equipment are physically grounded (at *A* and *B*), the cable

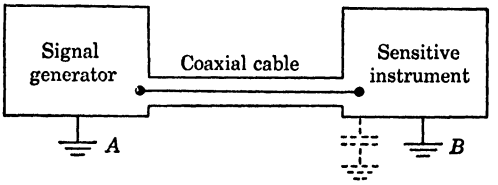


Fig. 16. Typical experimental arrangement.

shielding shares the return current with a miscellaneous ground path, thereby radiating trigger pulses throughout the vicinity to interfere with other equipment. Also if the cable shielding is thin relative to the skin depth at some low frequency it becomes essential that there be only one ground in the system, in which case only electrostatic pickup is possible. Multiple or thicker cable shielding is of course helpful. Sometimes even ungrounding the equipment at all but one point will

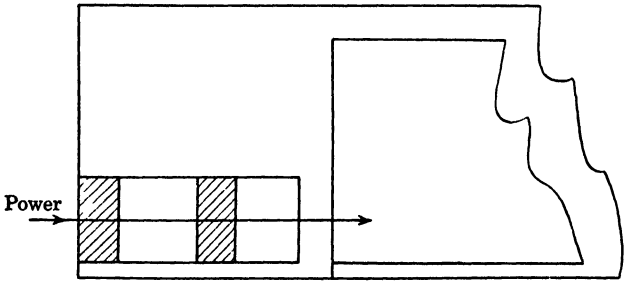


Fig. 17. Arrangement of double shield.

not suffice to prevent pickup through a cable shield if the capacitance to ground of a part of the equipment is large.

If grounding the equipment in more than one place, either physically or capacitively, cannot be avoided, we must resort to building the equipment with double wall construction, taking care to ground the two casings internally at only one point as in Fig. 17.

BEAM INTEGRATION

In the integration of beam current it is common to measure the potential developed on a known capacitor with the accumulation of

the collected charge. This method of course depends critically on the reliability of the capacitor. Many capacitors have various hysteresis effects and it is necessary to use those made of polystyrene, of certain types of mica, or to use a vacuum type. The time constant for the polystyrene type is about a week and for the mica type about a day.

The accuracy can be improved if an electronic circuit is employed to discharge the capacitor automatically when a certain threshold voltage is reached, thereby reducing the integration to essentially a counting method. This in turn depends on the accuracy with which the threshold voltage can be defined, and for vacuum tubes this is approximately 10 mv over short periods, 100 mv over periods of a day or so, and about 1 v with the change from one tube to another. It is necessary therefore to allow a capacitor to be charged to 100 v in order to realize an accuracy of 1%.

For the integration of direct current a d-c watt-hour meter is reasonably good.

BIBLIOGRAPHY

Elmore, William C., and Sands, Matthew, *Electronics; Experimental Techniques*. New York: McGraw-Hill Book Co., Inc., 1949.

Halliday, David, *Introductory Nuclear Physics*. New York: John Wiley & Sons, Inc., 1950.

Harris, Forest K., *Electrical Measurements*. New York: John Wiley & Sons, Inc., 1952.

Johnson, J. B., "Thermal Agitation of Electricity in Conductors," *Phys. Rev.*, *32*, 97 (1928).

Montgomery, C. G., *Technique of Microwave Measurements*. New York: McGraw-Hill Book Co., Inc., 1947.

Stout, Melville B., *Basic Electrical Measurements*. New York: Prentice-Hall, Inc., 1950.

Terman, F. E., *Radio Engineers' Handbook*. New York: McGraw-Hill Book Co., Inc., 1943.

2

MAGNETIC MEASUREMENTS

This section deals with some of the methods of measuring magnetic fields available to the physicist. Emphasis has been placed on methods that can give accurate answers in a reasonably short time with little inconvenience, and that are not often described in textbooks. One of the most common instruments in use for magnetic measurements, the ballistic galvanometer, is not included because it is discussed in nearly every textbook on electricity. We will not discuss magnetometers, of which there are innumerable varieties, nor permeability measurements, which are discussed in engineering texts.

Methods of measuring magnetic fields may be divided into two rough classifications: those that depend on particular properties of matter, and those that depend on Ampere's and Faraday's laws.

METHODS DEPENDING ON SPECIAL PROPERTIES OF MATTER

THE BISMUTH SPIRAL. A device utilizing the property of bismuth that its specific resistance varies with applied magnet field, has been made by Professor G. S. Smith[1] of the University of Washington. Professor Smith casts the bismuth into spirals about 1 cm in diameter and uses two of them in diagonally opposite arms of a Wheatstone bridge; the spirals, together with other resistors, are then built into a probe with one of the spirals at the end. In this way changes in resistance resulting from variations in temperature are balanced out. The probe is made long enough so that all elements except the probing spiral are in effectively zero field. The device is calibrated by least squares adjustment of the readings given in known fields. The accuracy obtained from data interpreted from these calibration curves is about 2% for d-c fields. Smith also studied a-c fields by using an oscillograph instead of the galvanometer and obtained calibration data in this case also, although a-c measurements were less accurate than d-c measurements.

The advantages of this device are the small size of probe needed, making possible measurements of fields of small extension, and the rapidity with which the device follows a varying field. Its chief disadvantage is the fact that the whole effect is temperature dependent,

[1] Smith, G. S., *Trans. AIEE, 56,* 441 (1937) and *58,* 52 (1939).

so that although the device is insensitive to temperature at *no* field, a separate calibration curve must be run for each temperature.

THE HALL EFFECT IN GERMANIUM. When a current-carrying strip of metal is placed in a magnetic field, a potential difference may be found across two points in the metal that are in a line perpendicular to the current and to the field; this effect is known as the Hall effect. It is due to the fact that the current-carrying elements in the metal are drawn into circular paths by the field and thus, instead of going in the direction of the applied emf, tend to bunch over to one side. The potential difference produced by the Hall effect is called the Hall voltage, and is

$$V_H = \frac{\rho i H}{t} \tag{1}$$

where V_H is the Hall voltage, i and H are, respectively, the applied current and field, t is the thickness of the metal, and ρ is the Hall coefficient of the metal. The direction of the Hall voltage depends on whether the current-carrying elements are electrons or "holes" (in the electron energy states).

The Hall effect has been used to measure magnetic fields,[2] using germanium, which has a large Hall coefficient and a not-too-large resistivity. Germanium may be either type N (electrons) or type P (holes), depending on impurities. In practice type N is used. A typical circuit is shown in Fig. 1. The battery circuit is set to deliver a specified current, and the germanium strip is placed in the field and perpendicular to it; the galvanometer may then be read directly in gausses, using Eq. (1). The readings are accurate to within 2% for fields between 1 and 8000 gausses; above 8000 gausses Eq. (1) is not accurate because the resistivity of germanium varies with field, and at high fields ρ itself is a function of field strength. However, with suitable calibration the method is good up to 20,000 gausses.

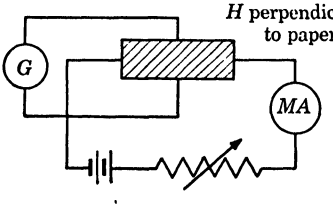


Fig. 1. Hall effect apparatus.

The advantages and disadvantages of this method are about the same as for the bismuth spiral. The germanium strip is small (about $\frac{1}{2}$ cm long) but is very temperature dependent.

MISCELLANEOUS SMALL OERSTED METERS. Several attempts have been made to make small portable meters that could be put in the field

[2] Pearson, G. L., *Rev. Sci. Instr. 19*, 263 (1948).

and read directly in gausses or oersteds. One such meter consists of two small bars of soft iron placed parallel and a short distance from each other. When the meter is placed in a field, poles are induced in the iron, and since the bars are parallel they will be magnetized in the same direction and will therefore repel each other (see Fig. 2). One bar is fixed, the other can move against a spring, and its motion is magnified by levers to drive the indicating needle. This device may be considered a magnetic analogue of the gold-leaf electroscope. It is good for a-c as well as d-c fields, and is also provided with a surrounding coil so it can be used as an a-c ammeter.

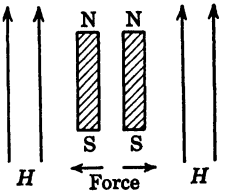


Fig. 2. Simple magnetic field meter.

Another meter consists of a small permanent magnet supported by a torsional fiber. The meter is placed in the field and slowly turned so that the torque of the fiber opposes the magnetic torque. As long as the maximum possible magnetic torque is greater than the fiber torque, the magnet and indicating needle remain stationary, but when the meter is turned a little too far the fiber torque takes hold and the magnet is flipped back to zero. The needle reading at the moment of flip gives the field strength. This device is made in several sizes, for maximum fields of 1000 gausses, 5000 gausses, etc. Care must be taken that the field being tested is not above the maximum, or the fiber may be damaged and/or the magnet strength changed. The two devices described above are not very accurate but are useful for quick, qualitative checks.

THE FLUX GATE COMPASS AND PEAKING STRIP. In a transformer, the voltage induced in the secondary depends both on the applied ampere turns and on the effective μ of the core. If the core is saturated by an external d-c magnetic field the effective μ will decrease and the output voltage will be below that produced when the core is unsaturated. This principle is used in the *flux gate compass*, which is essentially a transformer with core made of an easily saturable alloy such as permalloy. In use, the flux gate compass is provided with a bias coil carrying direct current and providing an accurately known field inside; the direct current is varied until there is a sharp peak in the output voltage, which indicates that the bias field is exactly equal and opposite to the field being measured. The core and windings can be made very small, less than $\frac{1}{2}$ cm in length, but the bias coil must be at least ten times as long as it is wide in order that the bias field be uniform. To avoid variations in output due to fluctuations in a-c input it is best to devise an output meter that compares the ratio between output and input.

The flux gate compass is especially good for measuring weak fields, down to 10^{-5} gauss; however, greater accuracy is obtained where the direct current can be read most accurately. This method has been used extensively for mapping the earth's field by air, particularly in exploration for oil. The method is not practical for fields of 1000 gausses or over because of the large number of ampere turns needed for biasing.

Another device working on the same principle is the *peaking strip.*[3] It consists of a permalloy core with a single winding, placed in a d-c bias coil if necessary. As the varying external field passes through zero and reverses direction, the magnetization of the core suddenly flips over, producing a pip of output voltage in the winding. Thus the peaking strip can be used to determine the instant a rapidly varying field goes through zero, or through any other value of biasing. For example, consider a synchrotron producing 10^5 ev electrons in a 1 m radius. The field at the edge of the magnet is 9 gausses at time of electron injection with $dH/dt \approx 1000$. Here a peaking strip 1 by 20 mils by $\frac{3}{4}$ in. wound with 400 turns of No. 49 wire provides the triggering for the electron injection and for timing other operations. Also, a stationary peaking strip may be used to trigger an oscillograph sweep while a movable strip operates the vertical deflection, thus measuring the field at different places. The half width of the peak is about 0.05 μsec.

METHODS INVOLVING ELECTROMAGNETIC INTERACTIONS

THE SEARCH COIL AND INTEGRATING DEVICES. The most common method of measuring magnetic fields makes use of a search coil and some method of integrating the emf developed across the coil. Let $A = \Sigma A_i$ be the total effective area of the coil and let $\phi = BA$ be the total flux through the coil. Then the induced emf is $E = d\phi/dt$ (absolute units used throughout this discussion), and

$$\int_0^t E\,dt = \int_0^t \frac{d\phi}{dt}\,dt = (\phi_t - \phi_0) = A(B_t - B_0)$$

Therefore
$$A(B_t - B_0) = \int_0^t Ri\,dt = RQ \tag{2}$$

For a flip coil, $B_t - B_0 = 2B$, so that

$$B = \frac{RQ}{2A} \tag{3}$$

It now becomes necessary to consider ways of integrating the induced emf. Consider the circuit elements in Fig. 3. If we could apply

[3] Gregg, E. C., *Rev. Sci. Instr. 18*, 77 (1947).

a voltage between terminals 1 and 2, looking back into an open circuit, we would have

$$V_{12} = \frac{Q}{C} = \frac{1}{C}\int i\,dt$$

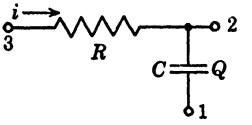


Fig. 3. Integrating circuit.

which would provide the desired integration. Actually, this arrangement is not possible and we must connect the flip coil between terminals 1 and 3. In this case we must satisfy the differential equation

$$R\frac{dQ}{dt} + \frac{Q}{C} = E \tag{4}$$

(R includes resistance of flip coil.) If the voltage pulse is in the form of a delta function at $t = 0$, the solution is

$$V_{12} = 2\frac{AB}{RC}e^{-t/RC} \tag{5}$$

The actual pulse shape is approximately half a sine wave, so the solution is

$$V_{12} = \frac{AB}{RC}\int_0^{\pi/\omega} \sin \omega\tau e^{-(t-\tau)/RC}d\tau \qquad (t > \tau)$$

which, after the pulse is completed, is

$$V_{12} = \frac{AB}{RC}\frac{1 + e^{\pi/\omega RC}}{1 + 1/(\omega RC)^2}e^{-t/RC} \tag{6}$$

Suppose the flip is made in 0.1 sec. A calculation shows that if $RC =$ 1 sec, (6) differs from (5) by 5%; if $RC = 10$ sec, the difference is 0.5%, etc. We would like to use a large value of RC, both for the sake of accuracy and to get a better approximation to the integral, but V_{12} also has RC in its denominator; we are forced to compromise in order to obtain a readable voltage.

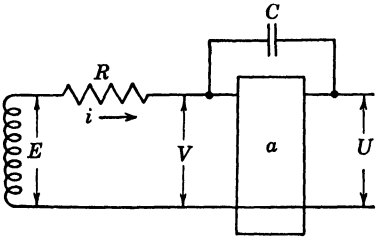


Fig. 4. Feedback integrating circuit.

A true integration, however, can be obtained by use of an amplifier. Consider Fig. 4. Let a be the amplification.

In this circuit $$E - iR = V = U - \frac{Q}{C} \tag{7}$$

Since $U = aV$, $$V - aV = -Q/C$$

so that $$V = \frac{Q}{C}\frac{1}{a-1}, \quad U = \frac{Q}{C}\frac{a}{a-1} \tag{8}$$

Combining (7) and (8),

$$E = iR - \frac{Q}{C} + \frac{Q}{C}\frac{a}{a-1} \tag{9}$$

The cause of difficulty in obtaining an integrating circuit so far has been the tendency of the capacitor to discharge. We see from formula (9) that we can prevent this discharge if we make $a/(a-1) = 1$, i.e., $a = \infty$. Amplifiers with infinite amplification are not uncommon, they are amplifiers just on the verge of oscillating and are called "superregenerative" amplifiers. A voltmeter placed across C (in the amplifier if necessary) will give a direct integration.

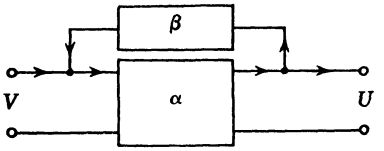


Fig. 5. Feedback integrating circuit.

It may be interesting to build a physical picture of how such an amplifier can act as an integrating element. Consider an amplifier of gain α and with positive feedback β (Fig. 5). For an input V,

$$\text{Output} = U = \alpha V(1 + \alpha\beta + \alpha^2\beta^2 + \ldots)$$
$$= \frac{\alpha V}{1 - \alpha b} = \frac{V}{1/\alpha - \beta} \qquad \text{for } \alpha b < 1$$

so that $a = \dfrac{1}{1/\alpha - \beta}$ and $a \to \infty$ in limit as $\beta = 1/\alpha$. Picture the loop $\alpha \to \beta \to \alpha$ as a delay line. A signal $V(t)$ entering α becomes amplified $\alpha(1/\alpha)$, i.e., is unchanged, and returns at $t + \Delta t$ to be added to the next input $V(t + \Delta t)$. Thus the delay line adds successive inputs, and in the limit of infinitesimal Δt, integrates the input.

The above methods of integration are purely electrical. We will now discuss integration by means of a fluxmeter, or torsionless galvanometer. The equation of motion of a galvanometer is

$$I\frac{d^2\theta}{dt^2} + D\frac{d\theta}{dt} + K\theta = T \tag{10}$$

where T is the applied torque, D is the damping constant, and K is the torsional constant. Let A_g and B_g refer to the coil area and field inside the galvanometer; then $T = B_g A_g i$, or $T = B_g A_g\, d\phi/dt(1/R)$. ($R$ is the resistance of entire circuit.) The damping emf is $E_D = B_g A_g\, d\theta/dt$ and the damping torque is $D\, d\theta/dt = (B_g A_g E_D/R)$, so that $D = B_g^2 A_g^2/R$. Equation (10) now becomes

$$I\frac{d^2\theta}{dt^2} + \frac{B_g^2 A_g^2}{R}\frac{d\theta}{dt} + K = \frac{B_g A_g}{R}\frac{d\phi}{dt} \tag{11}$$

In the fluxmeter I is made small compared with the damping term, so that inertial effects are negligible about 0.6 sec after the pulse. Equation (11) then becomes analogous to Eq. (4), and by continuing the analogy we can obtain an integrating device if we make $K = 0$. In that case we have simply.

$$\frac{B_g A_g}{R}\frac{d\phi}{dt} = \frac{B_g^2 A_g^2}{R}\frac{d\theta}{dt} \tag{12}$$

or, applying (2),

$$B = \frac{A_g B_g}{2A}(\theta_t - \theta_0) \tag{13}$$

In the fluxmeter we make $K = 0$ as follows: On the bottom of the galvanometer coil we place a small magnet (Fig. 6) so as to oppose the field and turn the coil away from zero. The suspension has a restoring torque $T_s\theta$, and the magnet has an opposed torque $T_m = mH \sin\theta \approx mH\theta$, so that to make $K = 0$ it is necessary only to set T_m equal to T_s. This can be done by adjusting the iron pads, a and b, which divert part of the main galvanometer flux to the magnet. If the coil is accidently turned through a large angle, T_s is greater than T_m, so the system is inherently stable.

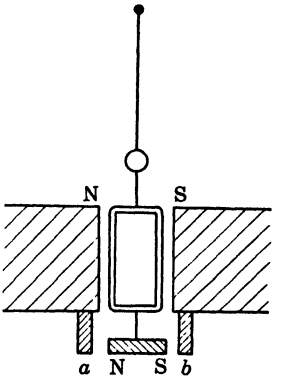


Fig. 6. Fluxmeter galvanometer.

Let us take another look at Eq. (13). If we write it in the form $\Delta\theta = 2AB/A_gB_g$ we see that the fluxmeter is most sensitive for the smallest A_gB_g. But sensitive galvanometers are characterized by large A_gB_g, so we reach the conclusion that the best fluxmeters are made of the poorest galvanometers! The reason is, of course, that a weak instrument produces less back emf during the motion, and therefore moves faster.

The above discussion on the flip coil took $\int d\phi/dt\, dt = 2AB$. If the coil, instead of being flipped, is pulled out of the field, $\int d\phi/dt\, dt = AB$, and subsequent formulas can be changed accordingly.

THE MAGNETIC BALANCE. When the fields to be measured are strong, steady, and uniform over a large area, very accurate measurements can be made by measuring the force exerted on a known current. The accuracy obtainable by this method depends only on the accuracy with which the current is known and on the construction of the balance. An example of such a balance designed for use in a vertical cyclotron field is given by Chang and Rosenblum.[4] They balance the

[4] Chang, W. Y. and Rosenblum, S., *Rev. Sci. Instr. 16*, 75 (1945).

horizontal force on a 10 cm length of wire by weights hung over a low friction (glass on glass) pulley. The wire itself is suspended from a knife edge; its motion is measured by a light spot reflected from a mirror. The accuracy claimed for this device is 1 gauss in 10^4 gausses.

Chang and Rosenblum also describe an indicating coil used for measuring the same cyclotron field. The coil is supported by horizontal torsion threads; it is set horizontal in zero field and has maximum sensitivity when the current is adjusted so that the coil is vertical in the given field. A slight nonvertical tilt can be exaggerated by hanging a small weight from the top of the coil. The accuracy claimed for this instrument is also 1 gauss in 10^4 gausses, but the instrument must be calibrated. A similar device can be made using the armature of a microammeter.

MEASUREMENT BY PROTON MOMENTS. Measurement of nuclear moments demands accurately known magnetic fields. As the values of the more important moments are now quite accurately known, and as measuring techniques are becoming more and more standardized, the use of the proton moment to determine the value of the field is being used quite frequently.[5,6] Techniques for magnetic field measurements using nuclear moments are discussed in Chapter 27.

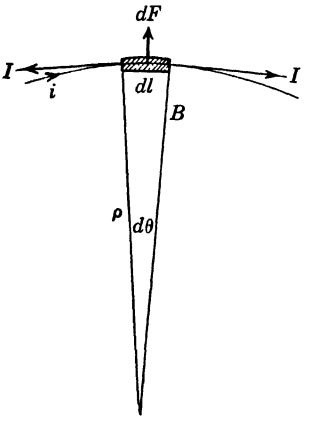


Fig. 7. Plotting of particle orbit.

ORBIT PLOTTING. Knowledge of the orbits of particles in a magnetic field is often required when experiments involving magnetic deflection of fast particles are performed. Calculations of orbits are complicated and tedious when the magnetic fields are not constant. Experimentally, the shape of the orbit can be determined by stretching a wire in the field and passing a current through it. The tension in the wire and the current must be adjusted to correspond to the momentum of the particle under consideration. To make this idea quantitative, let us consider a small element of the wire in a magnetic field (Fig. 7). Suppose that a magnetic field B exists perpendicular to the paper and directed into it, and that a current i is in the wire. There will then be a force dF on the element dl in the direc-

[5] Hopkins, N. J., *Rev. Sci. Instr. 20*, 401 (1949).

[6] Pound, R. V. and Knight, W. D., *Rev. Sci. Instr. 21*, 219 (1950).

tion shown, and $dF = (Bi/10)dl$. Now in equilibrium, $dF = T\,d\theta = (Bi/10)\rho\,d\theta$. Therefore

$$T = B\rho\,\frac{i}{10} \tag{14}$$

Thus if we place a wire in a field B, pass a current i through it, and apply a tension T to it, the wire will assume a curved shape with radius ρ given by the formula. If B is in gausses, i is in amperes, and ρ is in cm, T will be in dynes. The current must be sent in the opposite direction to the charged particles, as can be seen in Fig. 7; a beam of positive particles in the direction of i would bend upward.

Since Eq. (14) holds for any element of the wire, if the field B changes over the length of the wire, the radius ρ will change correspondingly if T and i are constant. Thus the wire will take the exact shape of an orbit of a particle with a given magnetic rigidity $B\rho$.

The experimental procedure is as follows. A very flexible wire is placed in a magnetic field, with a tension T applied to it by means of small weights. A known current i is passed through the wire and the shape of the wire is observed.

Suppose, for example, we wish to determine the orbit of 100 Mev π-mesons as they come out of the field of a cyclotron after having been created in the target. The $B\rho$ for 100 mev π-mesons is about 6.5×10^5 gauss-cm. If we choose $i = 5$ amp, a convenient current,

$$T = 3.25 \times 10^5 \text{ dynes} = 330 \text{ grams.}$$

In magnetic fields that are quite inhomogeneous, care must be taken that transverse forces on the wire do not force the wire out of the plane of the desired orbit.

BIBLIOGRAPHY

Harnwell, Gaylord P., *Principles of Electricity and Electromagnetism.* New York: McGraw-Hill Book Co., Inc., 1949.

Harris, Forest K., *Electrical Measurements.* New York: John Wiley & Sons, Inc., 1952.

Moullin, E. B., *The Principles of Electromagnetism.* New York: Oxford University Press, 1950.

Page, Leigh, and Adams, Norman I., *Principles of Electricity.* New York: D. Van Nostrand Co., Inc., 1949.

Webb, C. E., *Principles of Electric and Magnetic Measurements.* New York: Prentice-Hall, Inc., 1937.

3

VACUUM TUBES AND AMPLIFIERS

In the following remarks, emphasis will be placed principally on amplifiers applied to the problem of raising weak signals to a level that can be conveniently measured or used to actuate triggering devices for ultimate reading by electrical or mechanical means, or for display on a cathode ray oscilloscope.

Most frequently encountered in such amplifiers are receiving tubes, generally triodes or pentodes. Use of the static load line[1] provides a convenient graphical analysis of the tube and circuit characteristics. Families of plate current vs. plate voltage curves at fixed grid bias voltages for commercial tubes appear in handbooks issued by various vacuum tube manufacturers, such as RCA, General Electric, Western Electric, Raytheon, Victoreen, etc. To draw a load line for a typical circuit (Fig. 1) we have

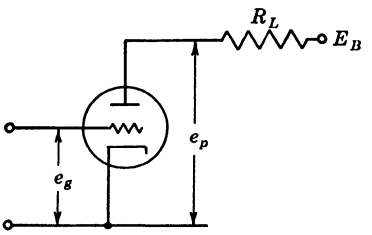


Fig. 1. Triode circuit.

$$e_p = ir_p, \quad iR_L + e_p = E_B; \quad \text{thus} \quad i = \frac{E_B - e_p}{R_L}$$

The static load line (Fig. 2) cuts the voltage axis where $e_p = E_B$, and cuts the current axis where $e_p = 0$.

It can be seen in Fig. 2 that for a given load resistance R_L, the grid voltage determines e_p and i. This method enables one to make a suitable combined choice of load resistance and grid bias so as to obtain, for example, linear output response over the widest possible input range. Linearity is shown here by the regularity of spacing of the grid lines along the load line. Moreover, the choice may be subject to the condition of maximum plate dissipation of the tube. The grid voltage excursions produced by the input signal cause the operating point to move back and forth along the load line about the point of fixed grid bias, called the quiescent point, and map the corresponding output voltage and current along the corresponding axes.

[1] Preisman, A., *Graphical Construction for Vacuum Tube Characteristics.* New York: McGraw-Hill Book Co., Inc., 1943.

For alternating currents, however the complex impedance of the load must be taken into account, and one must know the previous history to determine the output waveform. (For square waves and step pulses this process is more complicated.) For more general load impedances the path of operation is not confined to the load line but is in general an ellipse centered on the quiescent point whose major axis

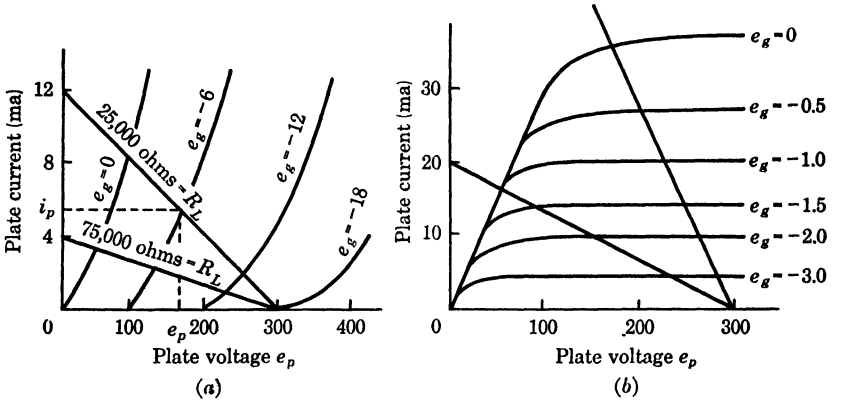


Fig. 2. Plate characteristics of 6SN7 (a) and 6AC7 (b).

makes an angle with the current axis whose tangent is Z_L, and which cuts the load line corresponding to the resistance component of Z_L at its points of horizontal tangency. If Z_L is predominantly reactive the motion is clockwise, if capacitative it is counterclockwise. Operation of the pentode in the region to the left of the grid family results in grid current being drawn, and is avoided in most applications.

The load resistance may be arranged to change with frequency by the use, for example, of a pair of resistors shunted by a capacitor. At low frequencies R_L is 60,000 ohms, but at high frequencies the capacitance gives zero impedance, and so R_L = 30,000 ohms.

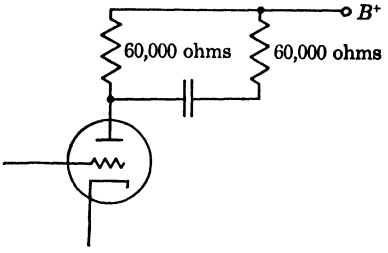


Fig. 3. Frequency compensation.

In addition to facilitating suitable choice of operating conditions, the load line may be used to analyze nonlinear (amplitude) distortion by a point-by-point plotting of the relation of input to output. Moreover, the load line analysis may be used to determine some of the tube constants. For example, the plate voltage amplification factor $\mu = (\partial e_p/\partial e_g)_{i_p}$ may be estimated from the horizontal spacing of the grid lines in the neighborhood of the desired quiescent point. It is clear from

Fig. 2b that the pentode is a high-μ tube, and consequently plays a dominant role in high-gain amplifiers. The transconductance $g_m = (\partial i_p/\partial e_g)_{e_p}$ may be similarly estimated from the vertical spacing of the grid lines in the neighborhood of the desired quiescent point. Since $g_m = \mu/r_p$ we have also that r_p, the plate resistance of the tube, is the slope of the grid line (which fact is the basis of the load line analysis). The voltage gain is readily seen to be $G = g_m R_L$, and although G varies as R_L the amplifier becomes unstable as $R \to \infty$.

For some purposes, of course, it is desirable to have a more analytical approach. A very serviceable first order theory assumes the grid

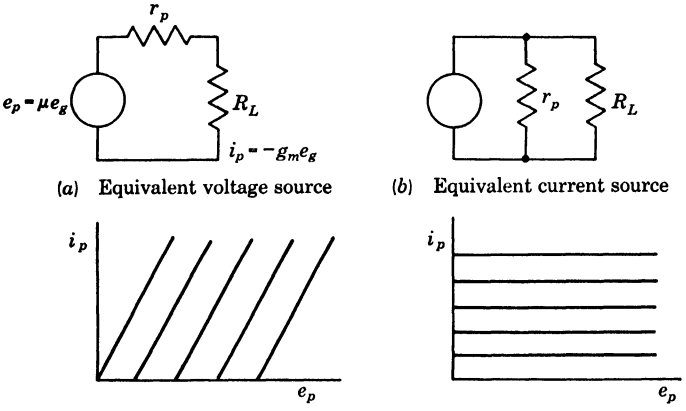


Fig. 4. Equivalent circuits and characteristics.

bias curves to be straight and parallel and is conveniently used in conjunction with equivalent circuits. In the voltage source equivalent, the tube is replaced by a generator generating a voltage $-\mu e_g$ across the resistance $r_p + R_L$. The current source equivalent replaces the tube by a pure current source $-g_m e_g$ which flows through R_L shunted by r_p. Because of the behavior of the pentode bias curves, the latter is particularly suited to pentodes.[2] Both schemes are represented in Fig. 4.

DIRECT CURRENT OR DIRECT COUPLED AMPLIFIER. This variety of amplifier may be arranged in a number of ways, all of which suffer from varying difficulties with drift and instabilities. Potentiometer coupling is simplest, (Fig. 5a), since it may use common B^+ and C^- supplies. The signal is passed along in the ratio R_2/R_1 so that $R_2 \gg R_1$, to get over-all gain. Directly coupling the plate of one to the grid of the next (Fig. 5b) requires piling up plate voltage supplies and is clumsy.

[2] Harnwell, Gaylord P., *Principles of Electricity and Magnetism*. New York: McGraw Hill Book Co., Inc., 1949, Chap. 7.

Both arrangements suffer serious drift problems. The plate voltage must be well regulated, requiring bulky equipment. A better method of overcoming drift due to changes in B^+ voltage is use of a circuit in which the plate current is exactly proportional to plate voltage, for then the voltage developed between the plate of the tube and a proper tap on a B^+ bleeder supply will remain zero as the B^+ supply voltage varies (within limits). A more troublesome cause of drift is the effect of filament temperature, which in turn effectively changes the grid bias and at the same time the internal plate resistance r_p of the tube. Cathode compensation can be achieved by paralleling tubes in such a way that one tube tends to hold the other back.[3] Compensation of plate

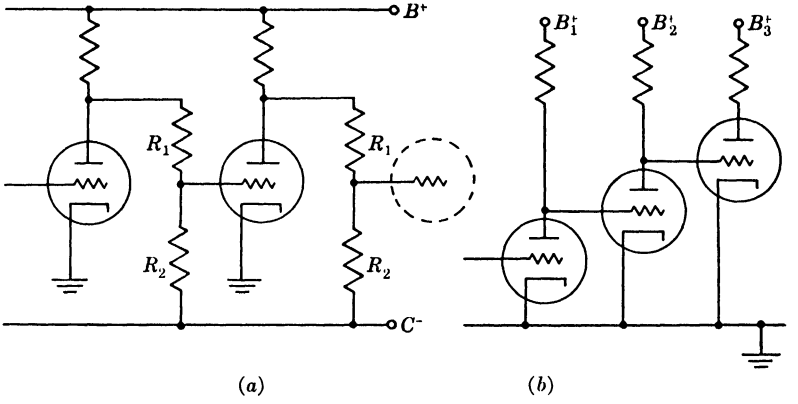


Fig. 5. Direct current amplifier circuits.

and filament supply variations can be effected in some applications by employing 100% negative feedback to obtain a current amplifier with approximately unity voltage gain, as discussed in Chapter 2. Still another method of obtaining drift correction is employment of one of a variety of bridge balance arrangements in which all variables in the amplifying arm are balanced against similarly induced changes in a second arm.[3] Drift can be partly overcome or can be made more nearly constant with time if the amplifier is left on for long periods, always allowing a warm-up period of at least half an hour after turning on. When turned on, the cathode voltage should always be applied first and the cathode allowed to come to equilibrium temperature before the plate voltage is applied.[4] The direct coupled amplifier can be made to have a flat frequency response down to direct current, and finds

[3] Artzt, Maurice, *Survey of D. C. Amplifiers, Electronics*, Aug. 1945, p. 112.

[4] Elmore, William C. and Sands, Matthew, *Electronics, Experimental Techniques, National Nuclear Energy Series, Manhattan District Technical Section, Division U*, Vol. 1. New York: McGraw-Hill Book Co., Inc., 1949, Chaps. 2, 3.

extensive use around the laboratory in electrometers, high-impedance voltmeters, and oscilloscope deflectors.

The electrometer tube circuit may be used to measure currents below 10^{-8} amp electronically. The problem is made difficult by grid currents, which in ordinary tubes are the same order of magnitude. Grid current is principally due to positive ions produced in the residual gas or emitted by the filament, photoelectrons, and leakage paths between grid and other terminals. To minimize these effects, a vacuum tube may be operated as shown in Fig. 6.[4]

The low plate voltage prevents production of positive ions in the gas. The screen absorbs positive ions emitted by the cathode, which is run at reduced temperature. Lighttight shielding reduces photoelectric effects. Special electrometer tubes, properly operated, can reduce grid currents to the order of 10^{-17} amp, and permit currents of the order of 60 electrons per second to be measured. The G.E. FP54, for example, with input resistance of 10^{11} ohms will give a current gain of $2.5(10)^5$, and using a galvanometer of sensitivity 10^{-10} amp/mm in the plate circuit, a sensitivity of $2.5(10)^5$ mm/volt can be obtained. If the tube is to be used for currents of the order of 10^{-13} or less it should first be washed with clean alcohol and ether and the glass near the control grid coated with ceresin wax to reduce surface leakage. In use, the air surrounding the tube should be dried or removed altogether. Grid leads should be short and provided with guard rings. A more rugged and less expensive tube for less exacting work is the Victoreen V-124B4.

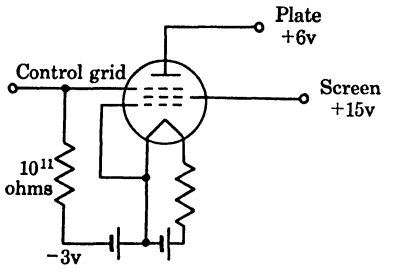


Fig. 6. Electrometer circuit.

MAGNETIC AMPLIFIERS. These amplifiers show promise of competing with other d-c amplifiers in very low resistance applications. In a magnetic amplifier[5] coils are wound on a magnetic core made of permalloy or mu-metal, a material which changes its magnetic characteristics rapidly with small flux changes. Thus very small direct currents can be used to control larger alternating currents in coils wound on the same core. Now used mostly in servo-mechanism and in some computing applications, magnetic amplifiers are also used in physics research. For example, a magnetic amplifier is used to detect changes in

[5] Butcher, F. E., and Willheim, R., "Some Aspects of Magnetic Amplifier Technique," *Proc. IRE, 40,* 261 (1952).

resistance of the components of a large solenoid coil used in a magnetic field cloud chamber apparatus.

RESISTANCE-CAPACITANCE COUPLED AMPLIFIERS. Direct coupling is never used when frequency response down to direct current is not essential. Putting aside the vast numbers of inductively coupled circuits as being generally of limited bandwidth, the most common laboratory applications employ RC coupling, Fig. 7. Since the signal divides between R_1 and C, the gain is proportional to the grid resistor R_1, which one wishes to be large. But, for example, in the dual triode

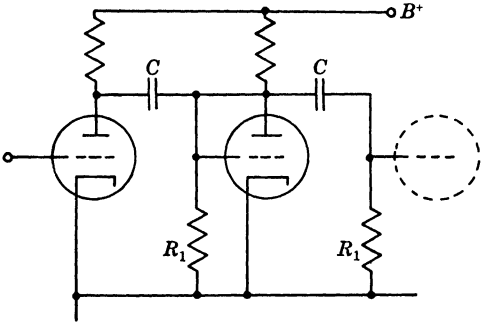


Fig. 7. Capacitance coupled amplifier.

6J6 the grid current is of the order of 10^{-8} amp, and hence for R_1 of the order of 20 megohms, the grid current through R_1 begins to produce a significant change in the grid bias. On the other hand, a high gain pentode such as the 6AU6 has higher grid current, since the grid is placed close to the cathode and is heated by it. Moreover, pieces of cathode emitting material are often carried over to the grid, promoting grid emission, which is still further enhanced by positive ion bombardment, a consequence of the residual gas in the tube. With sufficiently high grid resistance the grid current can produce decreasing negative bias, which still further increases the grid current until the grid destroys itself. For pentodes, the maximum safe grid resistance is of the order of 10^6 ohms.

VIDEO AMPLIFIERS. These amplifiers generally imply frequency response from audio well out into the megacycle range. H. A. Wheeler[6] has shown that there is a fundamental limit, however, to the maximum uniform gain G that can be obtained over a wide frequency band $\Delta\omega$ by means of a single vacuum tube in a stage: $G = 2g_m/\Delta\omega C$ or $G\Delta\nu =$

[6] Wheeler, H. A., "Wide Band Amplifiers for Television," *Proc. IRE, 27,* 429 (1939).

$g_m/\pi C$. This gain bandwidth product, which has already been referred to in the first chapter, depends on C, which is at best (i.e., neglecting all stray wiring capacitances) the geometric mean of the grid to cathode and plate to cathode capacitances of the tube itself. In all high-μ tubes this is always at least of the order of a few micromicrofarads. Since amplification of fast transients or pulses requires a wide bandwidth as a consequence of the fundamental Fourier product $\Delta\nu\Delta t = 1$, the gain bandwidth product is an important limitation of this type of amplifier. To see how the capacitance limits the high-frequency response by shunting high-frequency components, consider the tube as a switch (Fig. 8). A sudden drop in grid voltage is equivalent to opening the switch. When the switch is closed the voltage on C (equivalent to the plate voltage e_p) is $e_p = V_{B+}R/R_L$. If the switch is opened instantaneously the voltage on C will rise exponentially to V_{B+} according to $e_p = V_{B+}(1 - e^{-t/R_LC})$ (Fig. 8b). For fast response, then, the capacitance C should be small, and R_L also small (the latter condition limiting the gain). On the other hand, when the switch is closed it is the tube resistance R that limits the speed of response, since C must discharge through R. Thus tubes with high mutual conductance

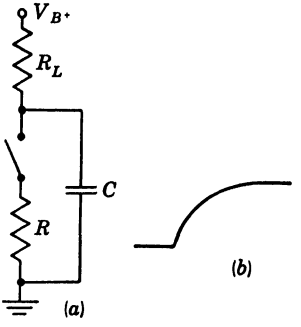


Fig. 8. Equivalent circuit.

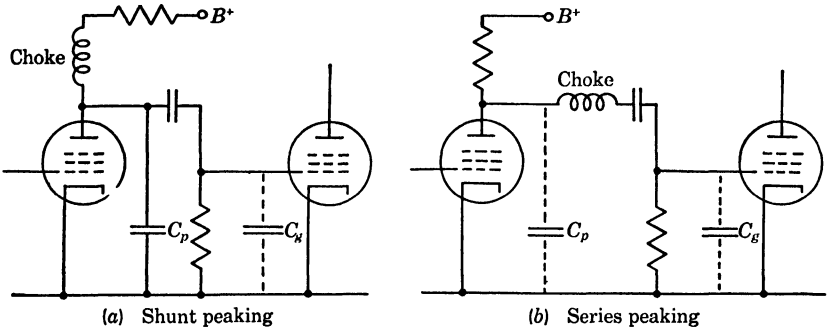


Fig. 9. (a) Shunt peaking circuit. (b) Series peaking circuit.

must be used. Since the speed of response is closely related to the bandwidth, we can define the "bandwidth index," a useful figure of merit for a tube to be used in video amplifiers, as the band width corresponding to unit gain: $G_{max} = 2g_m/\sqrt{C_pC_g}$. The RC rise can be improved somewhat by introducing a choke in shunt with the coupling (Fig. 9a). By using twice as big a choke as would give critical damping, a small

overshoot is produced, which is not objectionable in pulse amplifiers and gives a better fit to the step pulse. The choke may be introduced in series as in Fig. 9b. This combination acts as a transmission line. The effect of shunt peaking in improving normal response (dotted line) may be seen in Fig. 10, as the dashed line.

If the video amplifier is to be used to display transients on an oscilloscope, the requirements include, in addition to flat frequency response out into the megacycle range, also a constant time delay over the frequency range, or equivalently, the phase shift must be a linear function of frequency. To achieve these ends rather complicated interstage coupling networks must be introduced. The general analysis is most conveniently carried out in terms of filter theory.[6] Filter sections used for coupling may be chosen to produce desired compensation in gain or phase shift. It is good practice to use high-gain stages with compensation first until the signal is well above noise introduced by subsequent stages, followed by wide response low-gain stages. Finally a power stage is used to drive the output load. If the output feeds a line such as a coaxial cable, this output stage is generally a cathode follower. (See below for discussion.) The 6AC7, for example, has a good gain bandwidth index (90), but has a maximum plate dissipation of 3 w. On the other hand, the 6L6 has a poor gain bandwidth rating (45), but can dissipate 19 w. This latter tube operated at unity voltage gain is a satisfactory output cathode follower. An excellent chart prepared by Whyte[7] shows the characteristics of available pentodes in broad-band amplifiers.

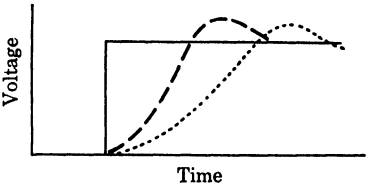


Fig. 10. Waveforms with and without peaking.

Fast pulse amplifiers most frequently encountered, for example in nuclear research, are designed somewhat differently from the transient amplifier, for here the important requirements are fast rise time, and linearity between input and output voltages under conditions in which the pulses follow one another randomly in time and at a very high rate (high counting rates). In this case, the circuit requirements are far less exacting. Indeed, it is necessary to employ circuits to alter deliberately the shape of the input pulses. Typically, the input pulses are characterized by short rise time of the order of 10^{-6} sec or less, followed by a long RC tail, sometimes a hundred times longer than the initial rise. At high random counting rates, the pulses tend to pile up by virtue of their

[7] Whyte, J. R., "Choosing Pentodes for Broad-band Amplifiers," *Electronics*, April 1952, p. 150.

long decay time, so that methods must be employed to "clip" the pulse after its initial rise in order to prevent considerable random fluctuations of the input signal about its mean. Such fluctuations, if uncorrected, would effectively alter the bias and destroy the input-output linearity. The problem of pulse shaping[4] by clipping may be achieved by reducing the gain in the region below the mid-band frequency, together with the use of a short time constant coupling between only one pair of stages. Another method makes use of the returning echo in a delay (transmission) line short-circuited at one end and placed across the plate impedance of the tube. The plate impedance is chosen to terminate correctly the other end. The delayed echo returns with inverted voltage polarity and effectively clips by subtraction from the signal voltage.

DISTRIBUTED AMPLIFIER.[8] We have already seen that the ordinary RC coupled video amplifier employing single tubes in the stages is subject to the gain bandwidth product, which is determined by the

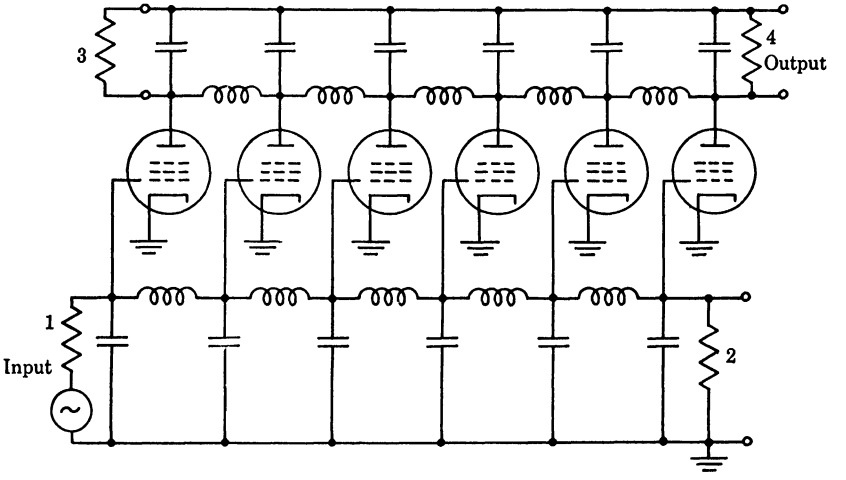


Fig. 11. Distributed amplifier.

interelectrode capacitances of the tubes. That this limitation can be overcome by the distributed amplifier principle has already been indicated in the first chapter, but it is appropriate to make a few additional remarks about it, since it is rapidly coming into wider use in the laboratory. (This is particularly true as a consequence of recent advances in fast crystal counter techniques.)

By paralleling tubes in such a way that the tube capacitances are incorporated as elements of artificial transmission lines it is possible to

[8] Ginzton, Edward, Hewlett, W. R., Jasberg, John, "Distributed Amplification," *Proc. IRE, 36,* 965 (1948).

obtain, for example, flat frequency response from direct current up to several hundred megacycles. The limitation on bandwidth now appears to be by grid loading effects in present tubes, in which some improvement may be anticipated by improved design. A single stage of such an amplifier is shown in Fig. 11. The grid line 1-2 is arranged to have the same propagation velocity as the plate line 3-4, and the lines are each terminated at 2 and 3, respectively, so as to produce no echoes. The input signal is introduced at 1 as indicated by the generator, and causes a wave to travel to the right in the grid line. As the wave passes, each grid in turn excites waves traveling to the right and left in the plate line. With perfect termination of the latter at 3, the leftward traveling waves are completely absorbed, leaving only those traveling to the right, all of which add in phase across the output termination 4. The gain of the stage is thus proportional to the number of tubes in the stage. There is an optimum grouping of tubes, i.e., a least number for a given over-all gain; this is obtained when the gain of the stage is $e = 2.72$. Thus for n sections cascaded m times, we have $m \times n$ tubes.

Traveling-wave amplifier. Another type of broad-band amplifier, based on a different principle, has been developed by J. R. Pierce[9] and his associates. In the traveling-wave tube a beam of electrons flowing down a tube is affected by an electromagnetic wave in the same tube. A weak electromagnetic wave is amplified in interaction with the electrons, and the output signal is taken from the modulated electron beam. These tubes are used as wide-band amplifiers in the ultra high frequency range. For example, an amplifier has been made for 3000 megacycles, with a bandwidth of 500 megacycles, and a gain of 20 db.[10]

Noise.[1] The problem of noise was discussed in the first chapter. However, there are some special problems of noise and spurious signals in amplifiers which it would be well to consider. Hum and audio pickup can be reduced in video amplifiers by providing as high a low-frequency cutoff as possible. This method is particularly convenient in pulse amplifiers when reduced low-frequency response is used for clipping. However, excessive clipping decreases the signal-to-noise ratio because of the resulting reduced pulse height. Appreciable loss in S/N occurs if $\omega_1/\omega_2 \approx 0.01$ where ω_1 and ω_2 are, respectively, the lower and upper half-power frequencies of the amplifier. Pentodes used in such amplifiers are inherently very susceptible to hum, pickup, and microphonics (2–10 kc) because of the close spacing of electrodes. For

[9] Pierce, J. R., *Traveling Wave Tubes.* New York: D. Van Nostrand Co., Inc., 1950.

[10] Dodds, Peters, and Kaisel, "New Developments in Traveling-Wave Tubes," *Electronics*, Feb. 1953, p. 130.

this reason it is sometimes helpful to use a triode in the first stage, and in any case, the first stage should be well shock mounted. High-frequency transients finding their way in via the filament heaters can be reduced by employing electrostatically shielded heater transformers. High-frequency parasitic oscillations frequently occur when the elements of a dual triode such as the 6J6 are used in parallel. This type of oscillation can be prevented by placing 50 ohms between the grids.

In nuclear research it is frequently necessary to locate the main body of electronic equipment at some distance from the signal source. In this case a preamplifier of moderate gain must be placed near the source. The output stage of such an amplifier is a cathode follower, which drives the coaxial cable. In such an arrangement, multiple grounds are generally unavoidable, and although the electromagnetic pickup can be reduced somewhat by multiple shields on the cable, etc., no adequate solution of this problem has been found. The extra spurious background counts must be accepted as part of the experimental error.

Hum present in plate supply can be reduced by distributing isolation filters between stages, Fig. 12. This also gives a feedback effect in increasing the likelihood of motor-boating. The best solution is to use more capacitance in the plate supply. In d-c amplifiers it is necessary to use separate power supplies. Oscillations can also occur at certain frequencies through accidental coupling of output and input by proximity of the output and input capacitors. In general, care must be exercised in locating components and wiring in wide-band amplifiers to avoid unwanted coupling and additional shunt capacitance to ground in signal-carrying leads. For example, leads carrying signals at low level must be kept away from a-c heater leads.

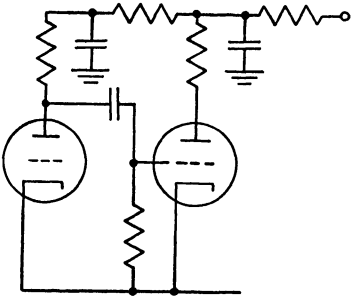


Fig. 12. By-pass filters in plate circuits.

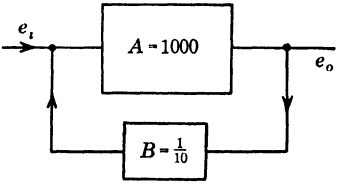


Fig. 13. Feedback amplifier.

FEEDBACK. In considering the requirement of gain stability it is at once evident that in all the preceding, one is at the mercy of scores of possibly drifting parameters particularly associated with the operation

of vacuum tubes. Tube parameters change with age, with vibration, temperature changes, supply voltage fluctuations, and on the replacement of tubes. Inverse or negative feedback has as one of its principal virtues the minimizing of such effects. Suppose in an amplifier with a voltage gain of $A = 1000$, a fraction $B = \frac{1}{10}$ of the output is fed back into the input so as to oppose the applied voltage, Fig. 13.

The real input is now

$$e_i' = e_i - e_o B$$

and the output is

$$e_o = A(e_i - e_o B)$$

or

$$(1 + AB)e_o = Ae_i$$

Hence

$$A' = \frac{e_o}{e_i} = \frac{A}{1 + AB} = \frac{1}{1/A + B}$$

Now since $A \to \infty$, the over-all gain becomes independent of A. Now B can be merely a couple of resistors, the effect of variation of tube characteristics if practically eliminated, and the gain stability problem is reduced to the much simpler problem of the stability of resistors. It is common practice in video amplifiers to obtain gain through a succession of feedback loops of the above type. Commonly, each feedback loop consists of two ordinary stages with a cathode follower whose output is fed back into the input in the ratio B.

Cathode follower. This is an important circuit element in video amplifiers. It is essentially a one-tube power amplifier in which the load is placed in the cathode circuit, Fig. 14a. The voltage e_o developed

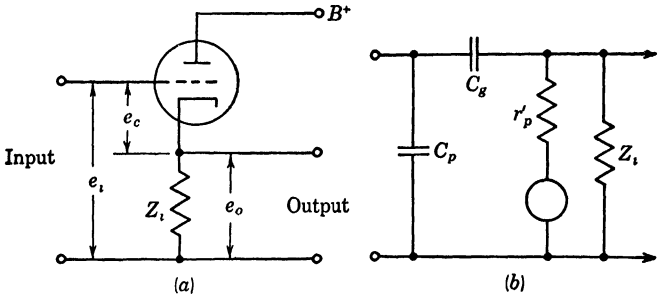


Fig. 14. Cathode follower and equivalent circuit.

across the load is related to the input voltage e_i at high frequencies when C_g, C_p can be neglected by

$$e_o = g_m e_i Z_{eq} = \left(\frac{\mu}{1 + \mu}\right) e_i \frac{R_L}{R_L + r_p/(1 + \mu)}$$

where Z_{eq} is Z_t shunted by $r'_p = r_p/(1 + \mu) = e_c/i_p$. We see from the figure that e_o opposes e_i; that is, we have a case of essentially 100% negative feedback, and consequently $e_o \leq e_i$, that is, the voltage gain is nearly equal to but never greater than unity.

The cathode follower has a high input impedance, but its very low output impedance, of the order of 200 ohms essentially regardless of load, makes it ideally suited as an impedance matching device to couple an amplifier to a coaxial cable, for example.

SELECTION OF TUBES. Tubes are generally selected on the basis of many considerations. Some of the more important are:

1. Choose receiving tubes if possible, since transmitting tubes are large and costly.
2. Linearity (triodes are usually more linear than pentodes).
3. Six-volt filaments (6 v filament transformers are more readily available).
4. Gain and mutual conductance (pentodes have higher gain than triodes).
5. Multiple purpose.
6. Gain bandwidth product.

The following tubes are useful types and are listed with comments taken from the *Tube Handbook:*

6H6	Twin diode: detection, low voltage rectifier.
6AL5	Twin diode: miniature.
6J6	Twin triode: miniature, medium gain, medium g_m.
6SL7	Twin triode amplifier: high gain, low g_m.
6SN7	Twin triode amplifier: low gain, low g_m.
6AC7	Video pentode amplifier: high g_m, high-gain voltage amplifier for high frequency.
6AG7	Video power pentode: high g_m, high output power.
6AK5	Pentode amplifier: miniature, high g_m, low interelectrode capacitances.
6AH6	Pentode amplifier: miniature, high g_m.
6AU6	Pentode amplifier: miniature, high g_m.
12AU7	Twin triode amplifier: miniature counterpart of 6SN7.
12AT7	Twin triode amplifier: miniature, high g_m, high gain.
6L6	Beam power tube: amplifier output stage, low voltage gain but high power.

CRYSTAL DIODES. These diodes are also widely used because of their good electrical properties, small size, and the fact that no heater current is required. Made now mostly with germanium, crystal diodes are available in many different types, according to the characteristics required.

TRANSISTORS. Transistors, which are also made from germanium crystals, have two or more contacts to the crystal so that they can be

used as amplifiers. Since their invention a few years ago, transistors have been developed rapidly and promise to make revolutionary changes in electronic techniques. As transistors become standardized and commercially available, they will be used widely in physical experiments. An issue of *Proc. IRE*,[11] devoted entirely to transistors and their applications, provides an excellent source of information in this field.

BIBLIOGRAPHY

Anner, George E., *Elements of Television Systems*. New York: Prentice-Hall, Inc., 1951.

Elmore, William C., "Fast Pulse Amplifiers for Nuclear Research," *Nucleonics*, September 1949, p. 48.

Elmore, William C. and Sands, Matthew, *Electronics: Experimental Techniques*. New York: McGraw-Hill Book Co., Inc., 1949.

Harnwell, Gaylord P., *Principles of Electricity and Magnetism*. New York: McGraw-Hill Book Co., Inc., 1949.

Preisman, A., *Graphical Construction for Vacuum Tube Characteristics*. New York: McGraw-Hill Book Co., Inc., 1943.

Ryder, John D., *Electrical Engineering Principles*, 2d ed. New York: Prentice-Hall, Inc., 1952.

Valley, G. E., and Wallman, Henry, *Vacuum Tube Amplifiers*. New York: McGraw-Hill Book Co., Inc., 1948.

Zworykin, V. K. and Morton, G. A., *Television*. New York: John Wiley & Sons, Inc., 1940.

[11] *Proc. IRE*, November 1952.

4

PULSE CIRCUITS

INTRODUCTION

In this chapter a number of commonly used circuits will be briefly discussed. Many physical experiments present information in the form of pulses of various sizes, shapes, and lengths, often random in time. The electronic problem is to amplify and detect these pulses, retaining as much as possible of the information content. Often the original pulse is inconvenient in size or shape, and must be changed to a more suitable form. Pulse-forming circuits are also used for comparison and testing. Discriminating circuits, counting circuits, and coincidence-pulse detectors are all commonly used in modern physics experiments. When very short pulses are to be detected, wide-band high-speed amplifiers must be used.

LINEAR OPERATIONS

Some operations on waveforms, such as addition, subtraction, inversion, differentiation, and integration may be done with *linear*

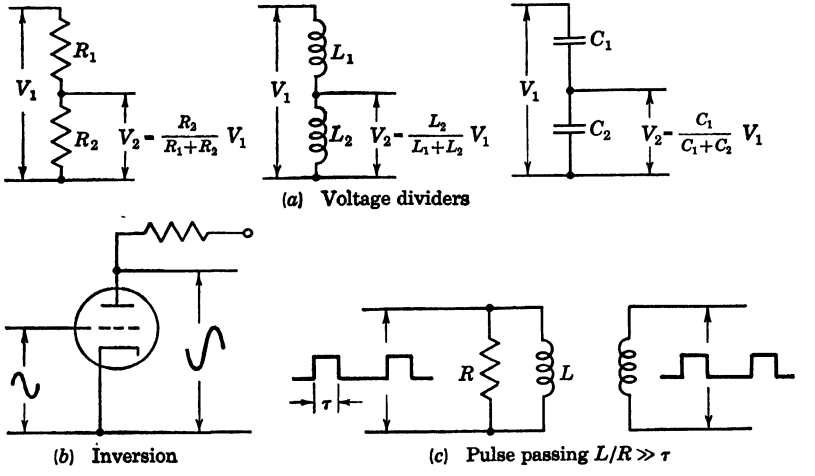


Fig. 1. Linear circuits.

circuit elements alone. These include resistance, capacitance, inductance, linear amplifiers, and delay lines. When circuits are composed entirely

of linear elements, the final result on the waveform is independent of the order of operation. However, operational nonuniformities may occur because of temperature variations affecting R and C, dielectric losses in C, copper and iron losses in L, or changing tube characteristics.

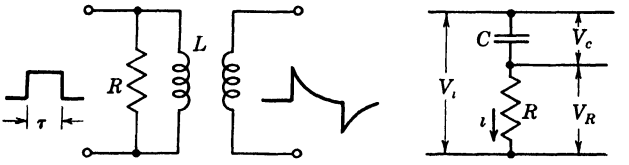
Differentiation

$$V_c = \frac{Q}{C} = \frac{\int i\,dt}{C}, \quad Q = cV_c$$

$$V_R = iR = \frac{dQ}{dt}R = RC\frac{dV_c}{dt}$$

if $$R \ll \frac{1}{\omega c}, \quad \text{i.e.,} \quad RC \ll \frac{1}{\omega}, \quad V_i \approx V_c$$

then $$V_R = RC\frac{dV_i}{dt}$$



(a) Differentiation $L/R \ll \tau$ (b) Differentiation $RC \ll \tau$

Fig. 2.

Output is proportional to derivative of input.

Integration

$$i = \frac{V_R}{R}$$

$$V_c = \frac{\int i\,dt}{C} = \frac{1}{RC}\int V_R\,dt$$

if $$R \gg \frac{1}{\omega c}, \quad \text{i.e.,} \quad RC \gg \frac{1}{\omega}, \quad V_i \approx V_R$$

then $$V_c = \frac{1}{RC}\int V_i\,dt$$

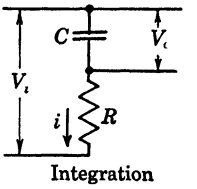


Fig. 3.

Output is proportional to integral of input.

Let us consider the output voltage from an RC circuit (Fig. 2b) when a unit step function is applied to the input,

$$\int_0^t \frac{i\,dt}{c} + iR = V_i = 1$$

$$i = \frac{V_i}{R}e^{-t/RC}$$

$$V_i = 1 = V_R + V_C$$

As the RC product decreases, the output voltage across R approaches the real derivative. As the RC product increases, the output voltage

across C approaches the real integral. This kind of circuit (linear) can never increase the steepness of the wavefront.

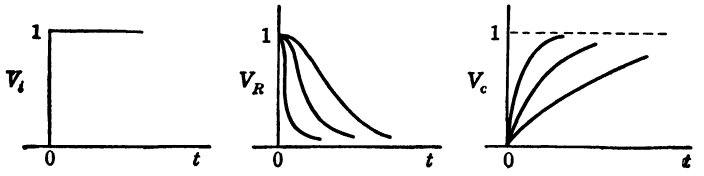


Fig. 4. Waveforms in differentiating and integrating circuits.

NONLINEAR OPERATIONS

Some operations on waveforms, such as amplitude selection or comparison, frequency discrimination, and sampling (indication of instant of equality of two signals with regard to amplitude, frequency, or phase) may be done with *nonlinear circuit elements*, either alone, or in combination with linear elements. The nonlinear elements include

1. High-vacuum diodes, or triodes connected as diodes.
2. Contact rectifiers.
3. Multigrid tubes, both high-vacuum and gas-filled.
4. Feedback circuits.

The choice of element to be used is quite broad; the manufacturer's rating data gives the necessary information.

SINE WAVE GENERATORS

Classical circuit theory has been largely concerned with circuits in which I, V, and Q are sinusoidal functions of time. The sine wave is unique, in that its wave shape is unaffected by any linear operation. Sine waves are easily generated, and are often "shaped" in nonlinear circuits to give timing pulses, for example. For a discussion of various sine wave generators, see the chapter on oscillators.

PULSED SINUSOIDAL OSCILLATIONS

The most common use of this waveform is the production of accurate timing markers. When two events are not periodic, the starting time is not predictable, and synchronization between the events is achieved by allowing the earlier event to start the timing waveform. The required waveform for this purpose must be

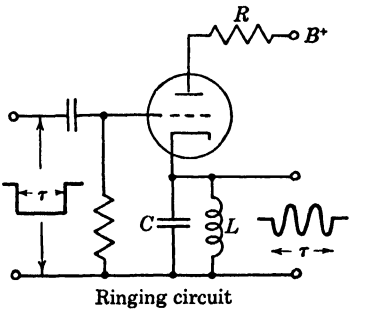
1. Zero until the beginning of the interval.
2. Sinusoidal during the interval.
3. Zero until triggered again.

The oscillations must always start at the same phase, and the starting transient must be of as short duration as possible. In order to

start an oscillator with no transients, all voltages and currents in the quiescent state must have the values that they will have in the steady state, at some instant. This is achieved by using *as few circuit components as possible.* One such circuit is the ringing circuit.

The negative "gate" on the grid must be large enough to cut the tube off completely, no matter how far negative the cathode oscillations go.

The output of the ringing circuit is slightly damped; if the desired interval τ is very large, the negative gate is used instead to pulse a Hartley or a simple crystal oscillator.



Ringing circuit

Fig. 5.

FAST WAVEFORMS

Sine waves are the "smoothest" kind of oscillations. For use in timing circuits, and similar applications, an *abrupt* waveform is desired. We will now consider waveforms of short total duration, and

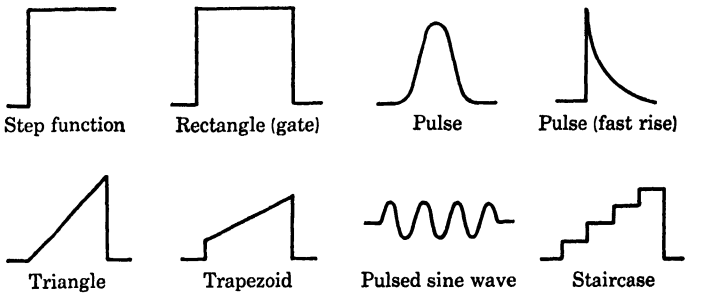


Fig. 6. Some commonly used fast waveforms.

those with short rise and fall times. The generation of fast waveforms is usually accomplished by

1. Steepening and squaring of sine waves.
2. Multivibrators.
3. Blocking oscillators.
4. Delay line pulse generators.

SHAPING CIRCUITS

Circuits that remove by electronic means one extremity or the other of a waveform are called *clippers* or *limiters.* Since a diode con-

ducts appreciably only when the plate is positive with *respect to the cathode*, we can use it for limiting.

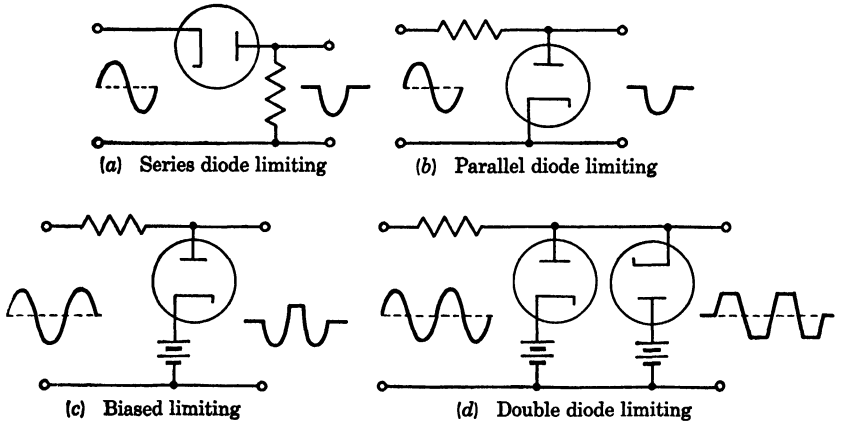


Fig. 7. Limiting circuits.

In each case, the resistance is large compared with the resistance of the tube while conducting, so that the amplitude of the output is essentially that of the input, except, of course, when the limiter is biased by means of a battery or small bias resistor to any desired level of output. This kind of circuit, ordinarily used to shape sine waves, can also be used as a simple electronic switch or to hold a series of various alternating voltages to some desired maximum amplitude.

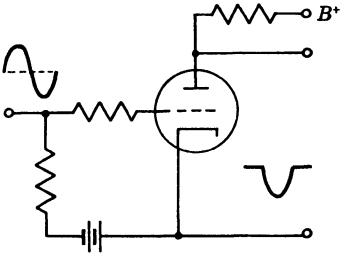


Fig. 8. Grid limiting.

By inserting a series grid resistor, as shown in Fig. 8, *any* tube can be used as a *grid limiter*. When the grid goes positive with respect to the cathode, grid current flows; if the series resistor is large, the grid-to-cathode potential remains low, hence the plate draws little current. When the grid goes negative (not enough to cut off the tube) the plate current swings follow the signal voltage on the grid. Here also the cathode may be biased to any desired value with respect to the grid.

SATURATION LIMITING. Saturation limiting may occur whenever the plate tries to draw a current greater than the ratio of the plate supply voltage to the load resistance. Thus the level of saturation limiting is easily controlled. *Cutoff limiting* occurs whenever the tube

is biased so that a negative grid swing cuts off the tube. If the grid bias is adjusted so that this takes place, and in addition the load resistance is chosen so that saturation limiting occurs on the positive swings, a square wave results directly from a sine wave input; we have an *overdriven amplifier.*

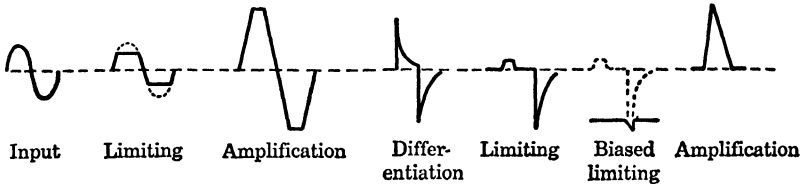


Fig. 9. Pulse forming from a sine wave.

Figure 9 shows an example of the production of a positive peaked pulse from a sine wave.

AMPLITUDE DISCRIMINATION

If it is desired to pass signals of only a certain minimum amplitude, we use a diode biased to this level of potential (Fig. 10).

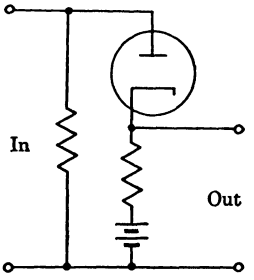


Fig. 10. Amplitude discrimination, positive signals.

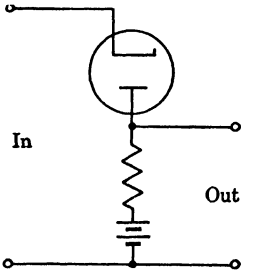


Fig. 11. Amplitude discrimination, negative signals.

To pass negative signals, we reverse the polarity of the diode (Fig. 11).

No amplification occurs in either circuit. If amplification as well as discrimination is desired, a triode may be used. However, the action is not so stable over a period of time, since triode tube characteristics change. Another disadvantage is that a large positive signal on the grid will cause it to draw current and thus bias it to an undesired value.

In the circuit shown in Fig. 12, the diode discharges the capacitor if the grid goes positive. The diode is called a "level setter." With

these types of circuits, although the amplitude may be controlled, the signals may be distorted by integration or differentiation; rather, such circuits are used to "trigger" fast waveform generators.

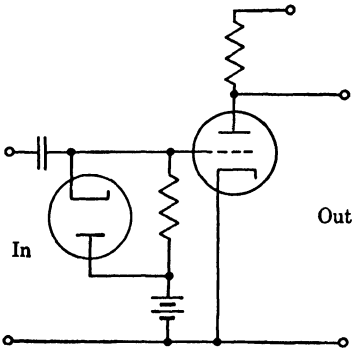


Fig. 12. Level setter.

A circuit which holds either amplitude extreme to a given reference level is called a *clamping circuit*, or *d-c restorer*, or *baseline stabilizer*.

The action of the clamping circuit depends on the fact that the input capacitor cannot change potential instantaneously. The output reference level can be anything, as shown, and the input voltage then *fluctuates around this level*. The diode clamper shown is a positive clamper; negative clamping is effected by reversing the diode.

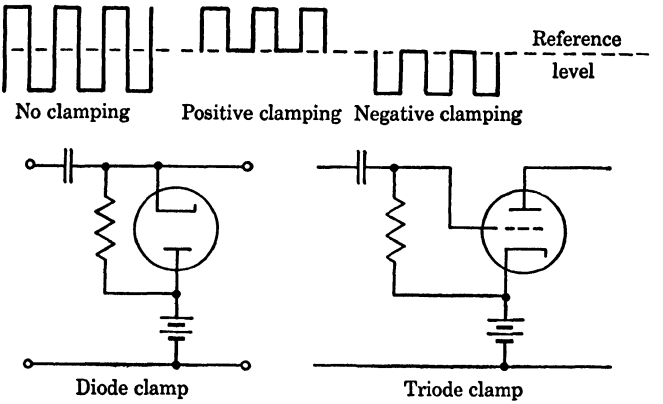


Fig. 13. Clamping.

PRODUCTION OF SAWTOOTH WAVES. This waveform is most simply produced by a *neon tube sawtooth generator* (relaxation oscillator). The d-c supply charges the capacitor through the resistor (Fig. 14). When the voltage across C is large enough, the neon tube in parallel with C fires, discharging C. As the voltage drops, the gas tube deionizes, current flow ceases, and C recharges through R in a time proportional to RC. The *thyratron sawtooth generator* operates in a similar manner except that a grid in the tube is used to control the firing potential.

Triangular waveforms can be generated with duration ranging from 10^{-7} sec to 10 min. As long as the capacitor charges or discharges at a constant rate, the voltage will be directly proportional to the time.

This condition holds best at the beginning of charge or discharge cycles.

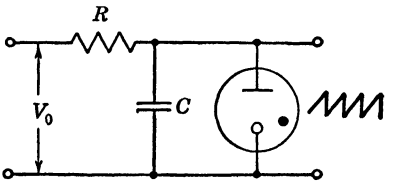


Fig. 14. Neon tube sawtooth generator.

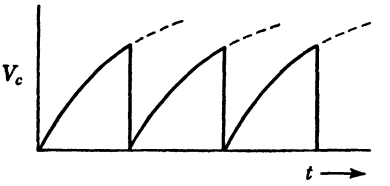


Fig. 15. Waveform from sawtooth generator.

The voltage across C is $V = \frac{1}{C}\int_0^t i\,dt$

If $i = I_0$ = constant, $V_c = \frac{I_0}{C}t$

STAIRCASE WAVEFORMS. A circuit that produces a staircase waveform can be used to add voltages, or as a frequency divider, with a

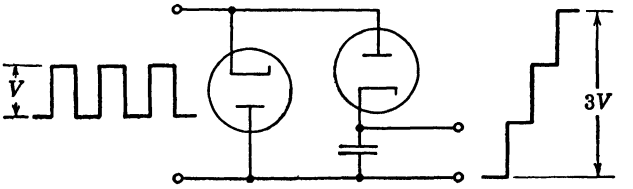


Fig. 16. Staircase generator and waveform.

"blocking oscillator" (see below) which is biased to fire only when its input voltage is nV.

MULTIVIBRATORS

Fast waveform generators can be classified as follows:

1. Astable—free-running, requiring zero trigger pulses to complete a cycle.
2. Monostable—(flip-flop circuit) requiring one trigger pulse to complete a cycle.
3. Bistable—(scale-of-two circuit) requiring two trigger pulses for each complete cycle.

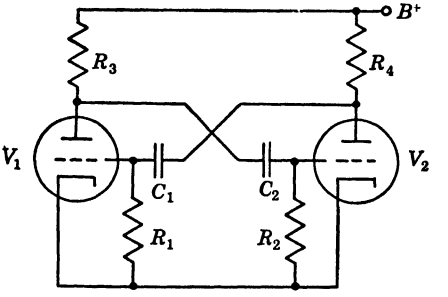


Fig. 17. Free-running multivibrator.

FREE-RUNNING MULTIVIBRATOR. This is basically a two-stage RC coupled amplifier, with the output of the second stage fed back to the input to the first stage.

When the high voltage is turned on, current flows in both tubes, but a perfect balance is impossible. If the current through V_1 is slightly greater, this tube draws more current through R_3; hence its plate voltage is less, so that the grid voltage of V_2 decreases slightly. This makes *its* plate current less, its plate voltage higher, and hence the grid of V_1 more positive. This causes V_1 to draw still more current; the effect accumulates until the voltage on C_2 cuts off V_2. Then C_2 discharges through R_2 until V_2 again conducts. As the current through V_2 increases, the current through V_1 decreases, until V_1 cuts off. This switching action repeats rapidly and continuously, with first one tube and then the other conducting. The frequency is determined by the discharge times of the coupling capacitors, and hence can be accurately controlled.

ONE-SHOT MULTIVIBRATOR. Both the neon tube generator, and the free-running multivibrator described above were astable circuits, requiring no trigger pulses to produce a repetitive output. The circuit described here is monostable, requiring one trigger pulse to execute each cycle. It is essentially a two-stage RC coupled amplifier, with one tube normally cut off, and the other conducting. Tube V_1 is initially cut off by the voltage drop across the cathode resistor, due to the plate current of V_2, which is conducting. A positive trigger is applied to the grid of V_1; its plate voltage drops, which lowers the grid voltage of V_2, which in turn passes less current. Thus the accumulated effect is to build up the voltage on C until V_2 is cut off. Then C discharges through R, until V_2 starts to conduct again, which rebiases V_1 gradually to cut off, and thus a cycle is completed. The frequency is determined by the trigger frequency, and by the discharge time of the RC combination.

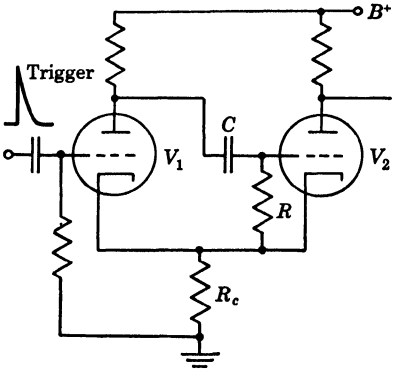


Fig. 18. One-shot multivibrator.

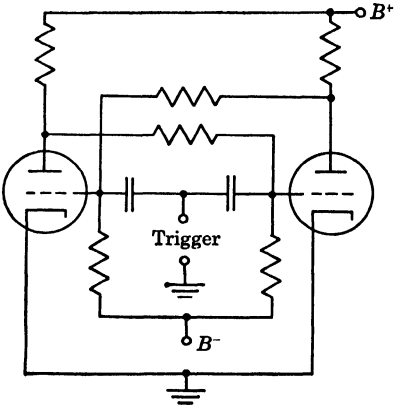


Fig. 19. Flip-flop circuit.

ECCLES-JORDAN FLIP-FLOP CIRCUIT. This is a bistable circuit which requires *two* trigger pulses to complete a waveform cycle. As in the free-running multivibrator, any slight unbalance will cause one of the tubes to cut off. There is no natural recovery here, through a capacitor discharge, since *direct coupling* between plates and grids is employed. The condition of equilibrium has, say, V_2 cut off and V_1 conducting. A positive trigger pulse is injected onto both grids, which makes no difference to the on-tube, since it is already conducting. The off-tube, however, starts conducting, which decreases its plate voltage, which in turn decreases the grid voltage and hence the plate current of the on-tube. Things progress in this direction until V_1 is cut off, and this represents a stable condition, so still another trigger pulse is needed to flip the circuit once more to bring the original conditions back. The flopping from one stable state to the other is practically instantaneous, that is, the *transition* is fast.

Negative triggers can also be used on this circuit, the action occurring is merely the cutting off of the on-tube, rather than the "turning on" of the off-tube. The frequency is just half the pulse repetition rate in either case.

PROBLEMS IN "FAST" GENERATORS

The suggestions offered below are applied to multivibrators, but the general considerations apply equally well to any fast waveform generator circuit.

OPERATION FOR FAST TRANSITION (rise-and-fall times). With the following precautions, transition times of 0.5 μsec are easily obtained, and times of 0.05 μsec are possible.

1. *Stray capacitances* (grid-to-ground, plate-to-ground, and grid-to-plate) *should be kept to a minimum.* Miniature tubes have (in general) interelectrode capacitances only half those of big tubes, and pentodes have smaller grid-to-plate capacitances than triodes. *Wiring should be kept short.* Ultimate limits are probably the subminiature solder-in tubes, and "printed" circuits. Any circuit attached to a fast generator should present as low a capacitance as possible. This is true of the trigger source, which should be isolated by triggering through a diode. If the output circuit also has a high capacitance, it should be fed through a cathode follower.

2. *Plate resistances should be kept small,* so that tubes will operate at high currents, and hence at high transconductance. Magnitudes of 3000 to 20,000 ohms are satisfactory as plate resistors. Inductive compensation in the form of series inductance is also sometimes used.

3. *High g_m tubes should be used.* These include triodes 6J4, 7F8, 6K4, 6J6, and 6SN7. In pentodes, 6AC7, 6AK5, and 6AG7 are best.

FAST RECOVERY TIME. After every abrupt change of state in any circuit, some of the capacitors are "left behind," since they cannot change their potentials instantaneously, and require a certain minimum time to recover. Thus two or more incoming pulses will not be resolved by the circuit, if their separation is smaller than the recovery time. The speed of recovery depends directly upon the current through the recharging capacitor, and inversely upon the size of the capacitor.

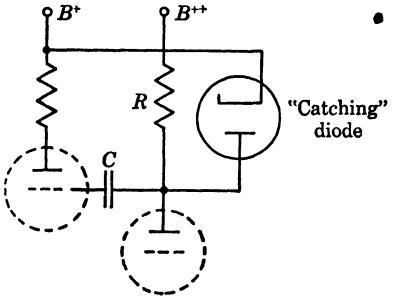


Fig. 20. Fast recovery circuit.

1. *Make R as large and C as small* as other considerations will permit.

2. In a two-tube circuit, *return one plate lead to a higher voltage supply than the other.* (This is called "catching.") This will increase the charging current through half the circuit, and this means that only the fast initial part of the charging occurs, so that the recovery time is reduced.

3. *The plate resistor may be made smaller.* This increases charging current also.

4. *Couple* the plate of one tube to the grid of the other *through a cathode follower.*

TRIGGERING. The simplest way to inject a trigger pulse into a circuit is to *couple it with a small capacitor* (50 $\mu\mu$f). However, if a positive trigger is used, and the positive rise turns a tube "on," then, as the pulse decreases, there is a tendency for this negative change to "untrigger" the same tube. The ideal triggering arrangement would include a switch that opens after the initial voltage step of the trigger has passed, so that the reversing tendency has no effect on the circuit. Use of *a series diode is the best way* to accomplish this decoupling, because its interelectrode capacitance is small, and because by inserting the diode in the proper direction, triggers of either polarity can be injected. The same diode can be used both for injecting the trigger, and for catching the plate below plate supply, as described in the last section. *Triodes are sometimes used* for triggering, but must be biased below cutoff, so as not to load the generator between trigger pulses.

GENERAL CONSIDERATIONS FOR "FAST CIRCUITS." For very short pulse times, multivibrators are not the best circuits; blocking oscillator and gas tube circuits are superior because of their ability to deliver

very large peak currents during the instants when the circuit is changing states (flopping over).

New kinds of fast generators have recently been developed in England. They are called "phantastrons" or "sanatrons"; they are essentially multivibrators, but their timing waveforms are linear (rather than an exponential *RC* rise) and are generated by a special external circuit called a Miller sweep circuit.

BLOCKING OSCILLATORS

Blocking oscillators are feedback circuits in which plate current is permitted to flow for some portion of a cycle, after which cutoff bias is imposed, usually by the accumulation of negative charge on the grid capacitor.

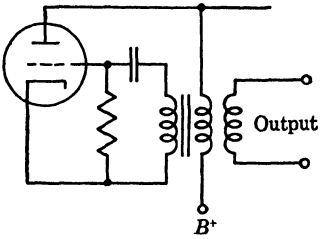


Fig. 21. Blocking oscillator.

Assume that the tube (Fig. 21) is initially cut off by the charge on the grid capacitor. As this charge leaks off, there comes a point where the tube again becomes just conducting. As plate current begins to flow, the grid is driven more positive, and the tube draws still more current. This is accomplished by arranging the polarity of the transformer windings properly. However, as the grid goes positive, grid current is drawn, which charges the capacitor and eventually cuts off the tube, and this completes the cycle.

Triggering is very important in this kind of an oscillator, since it determines the shape of the output pulse. A separate triggering diode should be used, because:

1. It causes minimum interference to oscillator operation.
2. It eliminates the reaction of the oscillator on the trigger source.
3. It reduces time delay between triggering and output pulse.
4. The trigger amplitude can thus be conveniently controlled.

USES OF BLOCKING OSCILLATORS.

1. To actuate a diode or triode switch.
2. Pulse frequency division (counting).
3. Positive or negative sawtooth generator.
4. Square wave generator.

When used for counting, the triggers should be of a uniform amplitude and equally spaced in time; the blocking oscillator is biased so that it fires on every nth trigger. For stable operation, n should be 5 or less.

The steepness of rise of a blocking oscillator pulse is dependent on the rate of rise of the trigger pulse. The shortest and most rapidly ris-

ing pulses have been produced by using a blocking oscillator to produce a fast trigger for *another* blocking oscillator. In this way, rectangular pulses of 0.1 μsec duration, with rise and fall times of 0.03 μsec are regularly obtained.

DELAY-LINE PULSE GENERATORS. These are used to produce rectangular waves from step functions or to terminate a pulse from a multivibrator, or to duplicate an existing pulse at a later time. The delay line has a certain time constant; when the pulse appears in the plate circuit, energy travels down the line and is reflected back *in phase* because of the short-circuited end, so that the line has twice its effective length. During this time of transmission and reflection the voltage across R is constant; when the reflected energy reaches R, which is of the correct size, it is absorbed by this "characteristic impedance" and the voltage drops to zero until the next pulse comes along. The net result is a square wave equal in duration to twice the "length" (time constant) of the line. The pulse shapes are very good, and lumped delay components can be added at will.

COUNTING

If the pulsed input is regular in amplitude and spacing, it can be most easily counted by a staircase generator driving a blocking oscillator. In nuclear physics, however, counting is complicated by the fact that there is a random time distribution of pulses, making it necessary to count large numbers to improve the statistical average. Rather than count for long periods of time, it is more efficient to count at high rates. Hence "scaling" circuits (dividing circuits) are used to reduce the counting rate so that it can be followed by a mechanical register.

The lower limit for counting is determined by the recovery time of the system. As a rough guide, it is not, in general, worth making recovery times of fast pulse circuits shorter than about 5 μsec, since the inherent recovery time of the counting chambers is longer than this.

The upper limit for counting is infinity if the states of the scaling circuits are stable, as in the Eccles-Jordan circuit. This kind of circuit is often used in cascade (one unit following another) to form a "*scale-of-two*" counter. The outputs of such a system are in the ratio 1:2:4:8:16:32:64:128: etc. The final output pulse, which comes once for every, let us say, 256 counts, trips a relay which drives a mechanical register.

A *scale-of-ten* counter can be built from cascades of scale-of-two circuits, by feeding back the properly numbered pulse to the third or fourth cascade unit, so that only 10 pulses cause a complete recycling and thus produce one output pulse.

SCALE-OF-*n* RING COUNTERS. Such counters are also used for counting, since each successive state is stable. They are composed of n identical units, with the output of the nth triggering the input of the first. The pulses to be counted are applied to all the units simultaneously, and the circuits are so designed that a given unit will fire only if the unit preceding it is on. Only one unit can be on at any given time. Thus the first pulse turns on the first tube, the second fires the second tube, and turns the first one off. The third pulse fires the third tube and turns off the second tube. The "on" tube thus effectively is driven around the ring by the successive pulses. The output is taken from the plate of the $(n - 1)$th unit, when the nth pulse drives it from on to "off."

These ring counters, like "sequence types" (Eccles-Jordan) and "energy-storage types," will not resolve coincident pulses; special circuits have been designed for this purpose.

COINCIDENCE COUNTERS. The most satisfactory of these types is the so-called "Rossi circuit," shown to record 3 coincidences, although a practical limit of 12 coincidences is possible. All tubes are normally

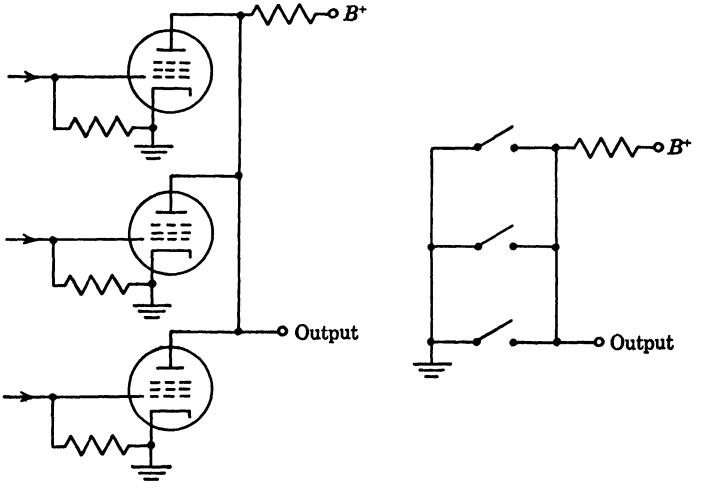


Fig. 22. Coincidence circuit and equivalent circuit.

conducting and the output voltage is constant. If one, or two, of the counters trigger their respective tubes off, the output voltage is not affected greatly, since the tubes are in parallel; if all the tubes go off simultaneously, the output voltage rises sharply to that of the supply, and an output pulse thus results for each threefold coincidence.

In general, operating characteristics require use of pentode tubes. If positive pulses are to be counted, both the grid and screen grid are

biased so that the tube is normally nonconducting. A positive pulse fed to the screen will fire the tube, but this makes the screen draw current, which is undesirable. Rather, the suppressor grid is driven in parallel with the grid, firing the tube, and causing a negative pulse in the output circuit. A special tube, the 6AS6, which is just a 6AK5 with a large suppressor, was designed for this purpose.

This circuit is used for coincidences down to 0.1 μsec. A consideration of gain-bandwidth product is not in order, since here gain is not of primary interest. If desired, following amplifiers may be used, to increase the output signal amplitude.

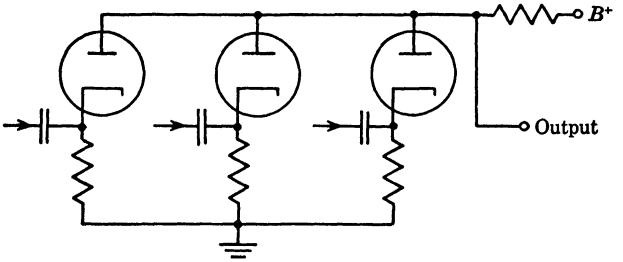


Fig. 23. Diode coincidence circuit.

Diode coincidence circuits[1] are now also widely used. Either vacuum tube and crystal diodes can be used, the latter having the advantages of small size, low capacitance, and no heater current. A basic diode coincidence circuit is shown in Fig. 23.

BIBLIOGRAPHY

Chance, Britton, *Waveforms*. New York: McGraw-Hill Book Co., Inc., 1949.
Elmore, William C. and Sands, Matthew, *Electronics: Experimental Techniques*. New York: McGraw-Hill Book Co., Inc., 1949.
Puckle, O. S., *Time Bases*. London: Chapman and Hall, Ltd., 1951.
Terman, Frederick E., *Radio Engineers' Handbook*. New York: McGraw-Hill Book Co., Inc., 1943.

[1] Howland, B., Schroeder, C. A., and Shipman, J. D., "Electronics for Cosmic-Ray Experiments," *Rev. Sci. Instr.*, *18*, 551 (1947).

5

RECTIFIERS AND POWER SUPPLIES

RECTIFIER TERMINOLOGY.

1. Maximum average current: the maximum direct current that can be supplied by a particular rectifier–limited generally by heat dissipating capacity of tubes or other rectifying units used.

2. Maximum surge current: the current that can never be exceeded without damage to emitter or rectifier element.

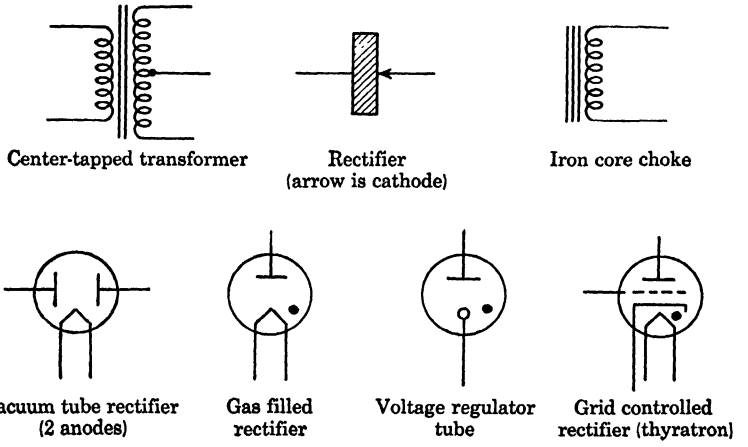


Fig. 1. Circuit symbols.

3. Maximum inverse voltage: the maximum voltage that can be applied in the nonconducting direction to a rectifier element without causing electrical breakdown.

4. Heating time: the time that must elapse between application of heater current and plate voltage to a rectifier tube to avoid damage to cathode.

5. Fractional ripple: the ratio of rms a-c component to direct voltage of rectifier output.

6. Utility factor: the ratio of volt-ampere rating of transformer to d-c power output of rectifier.

The most commonly used symbols are shown in Fig. 1.

RECTIFIERS. A rectifier is a device that has the property of passing current much more readily in one direction than in the other. The characteristic curve of an ideal rectifier is shown in Fig. 2, and that of a real rectifier in Fig. 3. The latter is a smooth curve and nonvanishing for negative applied voltage. As a result, small (10–100 mv) alternating voltages will not be rectified.

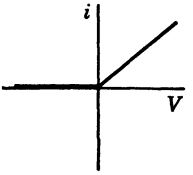


Fig. 2. Ideal rectifier characteristic.

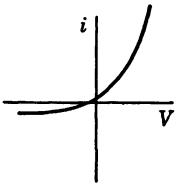


Fig. 3. Real rectifier characteristic.

HIGH-VACUUM THERMIONIC TUBES. This type is a diode consisting of a heated cathode and an anode with large surface. These elements are enclosed in an evacuated shell. High-vacuum rectifiers may be used within wide voltage and current ranges: from 200 to 10,000 v and from a few milliamperes to amperes. Oxide cathodes can be used below 1,000 v; above this value they are destroyed, and thoriated tungsten must be used. At very high voltages it is necessary to use pure tungsten.

Limitations. Pure tungsten requires high power. Maximum average current and maximum peak current must not be exceeded or the tube will be damaged. The peak current is determined by the maximum electron emission that the cathode can supply and still maintain complete space charge around itself. Inverse voltage limits must not be exceeded.

GAS-FILLED THERMIONIC TUBES. The gas is ionized by collisions with electrons from the cathode, forming positive ions which neutralize the electron space charge around the cathode. Consequently full cathode emission current can be drawn to the anode with a potential drop in the tube of only 10–15 v. This low voltage drop permits the use of oxide coated cathodes, which give high emission and are not damaged by positive ion bombardment at voltage drops of less than about 22 v. Argon, mercury vapor, or mixtures of these gases are commonly used in gas-filled tubes.

Limitations. Mercury vapor tubes are temperature sensitive. Below about 20°C the ionization is not sufficient to neutralize the space

charge, and the resultant rise in field strength near the cathode causes its destruction. At too high temperatures arc breakdown occurs.

COPPER OXIDE AND SELENIUM. A thin film of cuprous oxide on metallic copper gives a rectifying action, as does a film of processed selenium on a metal such as iron.

These rectifiers are very stable, the only effect of long use being a slight increase of resistance in the conducting direction during the first

TABLE 1

RECTIFIERS AND POWER SUPPLIES

Volts	10^{-3}	10^{-2}	10^{-1}	1	10	100
			Amperes			
100000	High-voltage high-vacuum diodes in cascade multiplier circuits 250R 100R 34R	Very high-voltage high-vacuum rect. KC4 250R	High-voltage high-vacuum KC4 etc.			
10000	High-vacuum diode multiplier ckt. 1B3, 8013, 8016 High vacuum diodes 1B3, 8016, 2X2A	High-voltage high-vacuum 8020 35R	Larger mercury vapor 872A 3-phase full-wave 872A single-phase	Mercury vapor 857B 872A below 3 amp	Mercury vapor 870A High-voltage ignitrons	
1000	High-vacuum 2X2A Selenium plates in series High-vacuum 5Y3	High-voltage high-vacuum 5R4GY 5Y3, 5U4, 5V4	Mercury vapor 866, 872A, 3B25 Selenium rectifiers below 250 v	Mercury vapor 5561 Selenium rectifiers below 250 v	Mercury vapor 870A Ignitrons	
100	Germanium diodes (1N44) High-vacuum diodes 6H6, 6AL5	High-vacuum low-resistance 5V4, mercury vapor 83, selenium rectifier stack	Selenium rectifiers	Selenium rectifiers	Selenium rectifiers	
10	Small selenium rectifiers Germanium diodes Copper oxide rectifiers	Selenium or copper oxide rectifiers	Selenium or copper oxide rectifiers	Selenium rectifiers Tungar type gas tube 4B26	Selenium rectifiers	
1						

10,000 hours of use. The copper oxide type is the less efficient. The allowable inverse voltage for a selenium rectifier is about 10 v.

GERMANIUM OXIDE. Rectifying action takes place in contact of a small wire ("cat's whisker") with a crystal of germanium oxide. This will handle 100 v at 35 ma.

COLD CATHODE GAS-FILLED RECTIFIER. In a gas at low pressure, the current flow between two electrodes is proportional to the area of the cathode. Thus by using electrodes of greatly different area, a rectifying action is obtained. This type has only limited use, mostly in vibrator power supplies. It needs no heater current, but it does need a fairly high voltage drop and it produces high frequency transients.

High-vacuum tubes, and selenium and copper oxide rectifiers may be operated in parallel. This cannot be done with tubes involving ionization, since if one ionizes first, the voltage across it will drop, and the others will not ionize. In the case of hot cathode mercury vapor tubes, parallel operation can be attained by use of a resistance in series with the plate to produce 10 to 30 v drop at peak current, or by use of equalizing reactors. Thus even though one starts first, the voltage on the others does not fall enough to prevent their starting.

The table on page 53 lists various types of rectifiers and the ranges of voltage and current at which they are useful. Note in particular the new tube types 34R, 100R, and 250R, which will stand any voltage below that at which breakdown takes place over the surface of the glass envelopes.

BASIC RECTIFIER CIRCUITS. The simplest common rectifier circuit is the half-wave type shown in Fig. 4 with its output wave form. The

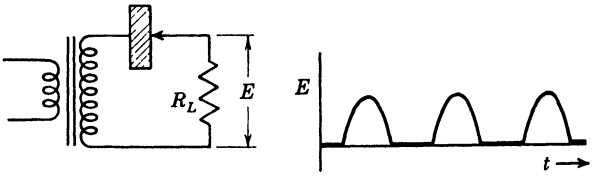


Fig. 4. Half-wave rectifier and waveform.

output direct voltage is the lowest for any circuit, being $1/\pi$ times the peak input voltage. It has a ripple of 1.21 and a utility factor of about 2.

The next more complicated and by far the most widely used circuit is the single-phase full-wave type (Fig. 5). This circuit is useful up to

about 1000 v at 200 ma. The average direct voltage is twice that of the half wave circuit, being $2E_m/\pi$. The ripple is 0.48 and the utility factor is 1.5.

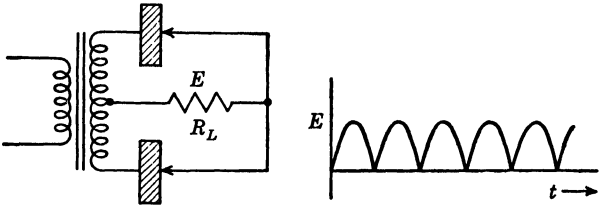


Fig. 5. Full-wave rectifier and waveform.

A third type of rectifier used with single-phase supply is shown in Fig. 6. This is the full-wave bridge circuit. Its output wave form is the same as that of the full-wave rectifier above.

Of the single-phase rectifier circuits, the full-wave type with center-tapped transformer is the most common one in use with tube elements, and the full-wave bridge is the most widely used with copper oxide and selenium elements. It will be noted that the rectifier elements in the bridge circuit operate at half the inverse voltage of those in the other full-wave circuit. In addition, the bridge gives greater d-c power for a given transformer volt-ampere rating. If used with thermionic tubes, the bridge circuit requires three separate well-insulated heater supplies, unless tubes are used that have well-insulated cathodes.

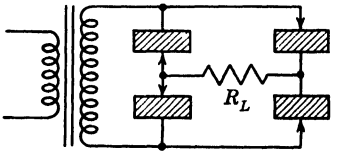


Fig. 6. Full-wave bridge.

The half-wave circuit is used only to supply low currents, as the current always flows in the same direction through the transformer core and tends to saturate it.

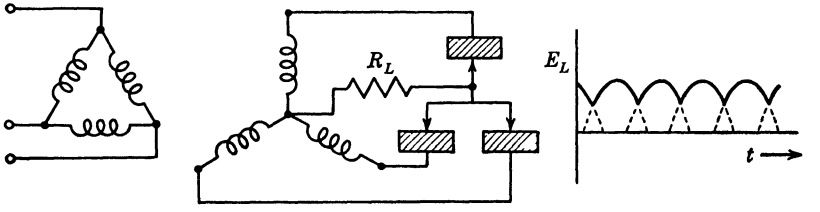


Fig. 7. Three-phase star (half-wave), ripple factor 0.17.

The circuits of Figs. 7, 8, and 9 are polyphase types, generally used to give high power at moderate to high voltage. That of Fig. 7 is the

most commonly used three-phase rectifier, and is the only one of the three types shown where there is no d-c component in the transformer core.

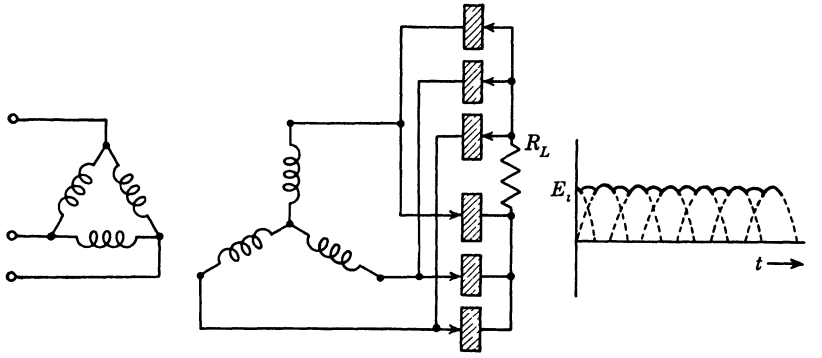


Fig. 8. Three-phase full-wave, ripple factor 0.04.

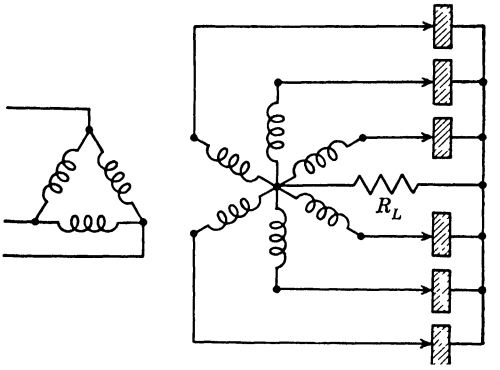


Fig. 9. Polyphase rectifier.

The single- and three-phase circuits shown in Figs. 10 and 11 are used where more than one output voltage is desired.

FILTERS. The outputs of rectifier circuits usually contain fairly large a-c components, which must be removed in most applications. Filters for this purpose contain combinations of inductors and capacitors, and fall into two main types: series-inductor input and shunt-capacitor input.

These are somewhat interchangeable, except that capacitor input must be used with the single-phase half-wave circuit.

The inductance input utilizes the tubes and transformer more effectively than the capacitor input and hence is used where large amounts of power are involved. It also gives better voltage regulation.

The simplest filter consists of a capacitor in parallel with the load. It is charged while the tube is conducting and partially discharges into the load when the rectifier voltage falls off. It can be seen that as the

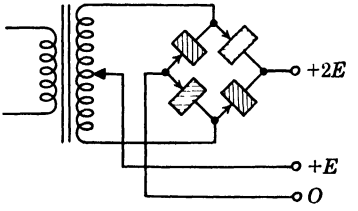


Fig. 10. Double voltage rectifier.

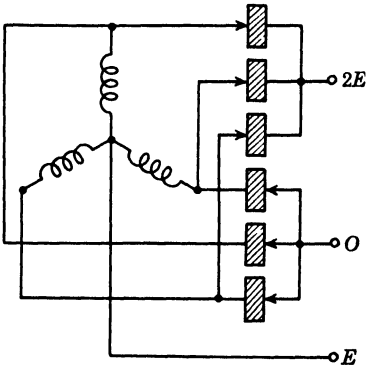


Fig. 11. Double voltage three-phase rectifier.

load approaches zero, the output voltage approaches the peak transformer alternating voltage. See Fig. 12.

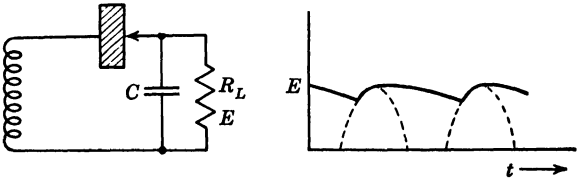


Fig. 12. Capacitor filter and waveform.

Figure 13 shows a choke-input filter, which permits the tubes to carry current over a greater phase angle. This type is not satisfactory for the half-wave filter, as it reduces the output voltage too severely.

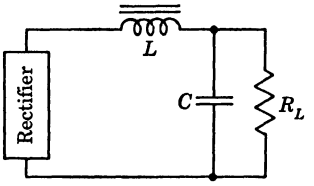


Fig. 13. Choke input filter.

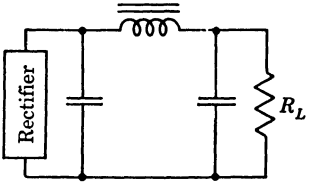


Fig. 14. Capacitor input filter.

In this circuit the minimum satisfactory value of the inductance is about $\frac{1}{1130}$ of the load resistance, for a 60 c single-phase full-wave rectifier.

Figure 14 shows the most commonly used filter for currents up to 200 ma at from 2 to 400 v. This filter has some of the bad features

characteristic of capacitor input, but gives a higher output voltage and lower ripple than the choke input. Capacitor input results in a much higher peak current, and in a much shorter conduction period, necessitating the use of high-vacuum tubes and resulting in a higher ratio of rms transformer current to output direct current.

VOLTAGE REGULATION. For many applications, it is necessary or at least desirable to regulate the voltage output of a power supply in

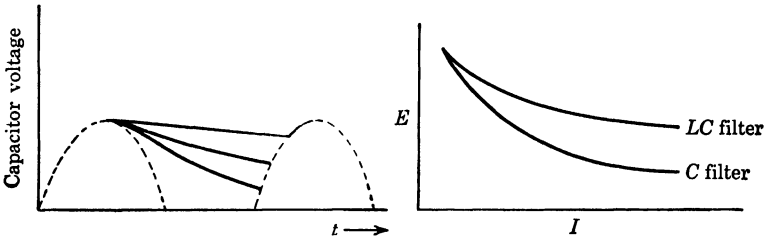


Fig. 15. Filter waveforms.

some way so that fluctuations in the input voltage or changes in the load resistance will produce little effect on the output voltage. The half-wave rectifier with capacitor input filter has very poor inherent regulation, the effect of increased load being to increase the rate of discharge of the capacitor (see Fig. 15). The *LC* filter has somewhat better regulation, which can be improved by adding a bleeder resistance across the load. This moves the operation further to the right on the *E* vs. *I* curve shown, away from the part of steepest slope.

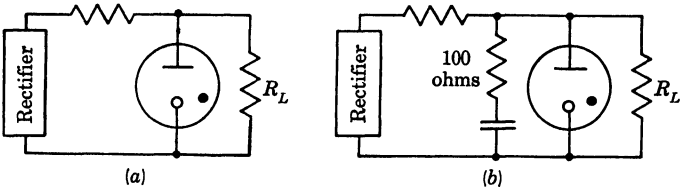


Fig. 16. Voltage regulator circuits using gas tubes.

The regulation characteristics of a choke input filter are improved by the use of a so-called swinging choke, one with a small enough air gap in the core that a fairly small bleeder current will cause appreciable saturation of the core and resultant drop in inductance.

Voltage regulator tubes may be obtained which, when conducting, have fixed voltage drops of 60, 75, 90, 105, or 150 volts. These are connected as shown in Fig. 16, the circuit to the right being provided with a capacitor and damping resistor to prevent relaxation oscilla-

tions. These are good up to 30–40 ma. A minimum current of about 5 ma must pass through the tube for good regulation.

A typical electronic voltage regulator is shown in Fig. 17. A change in the output voltage E_0 causes a change in the grid voltage E_1 applied to T_2, which is amplified in T_2 and applied to the grid of T_1 in such a way as to oppose the change. The voltage which this circuit attempts to hold can be changed by adjusting R_2.

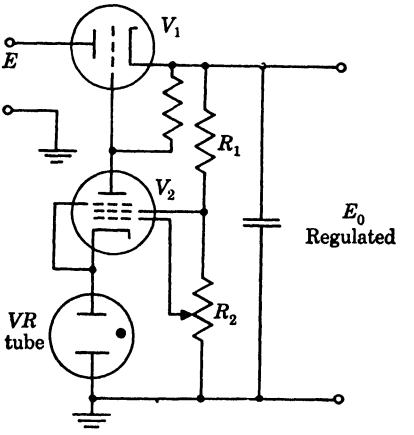


Fig. 17. Voltage regulator circuit.

Electronic regulation to 1 part in 1000 is difficult to obtain; to 1 part in 10,000, very difficult.

VOLTAGE MULTIPLIERS. In some cases we wish to avoid the use of a transformer, either because the direct voltage to be attained is so high as to make construction of adequate transformers difficult, or merely for convenience in obtaining a source of direct current at several hundred volts from a 120 v line for a low-power application. Rectifier tubes can be connected to form a voltage multiplier, giving a d-c output at some multiple of the peak input voltage as shown in Fig. 18.

The conventional doubler at the left operates by charging first one capacitor, then the other, in such a way that their voltages add to

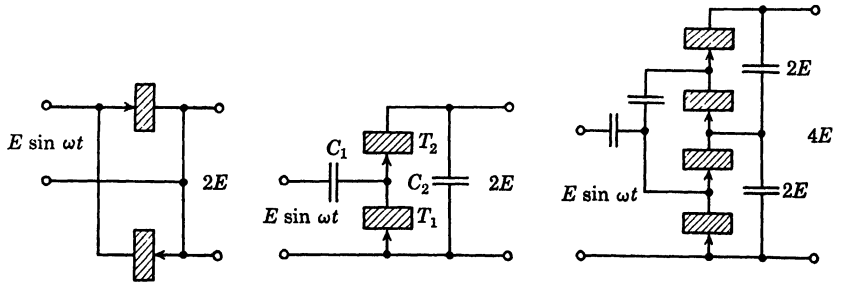


Fig. 18. Voltage multipliers.

give twice the peak a-c input. The cascade doubler first charges C_1 through T_1, then puts it in series with the line voltage to charge C_2 through T_2. The quadrupler acts in a similar manner, and in fact such a system can be extended indefinitely. A multistep cascade multiplier

was used by Cockcroft and Walton to obtain high voltage for their accelerators.

OTHER RECTIFIERS. Vibrators connected to a low-voltage d-c source are provided with two sets of contacts, one to convert the direct current to alternating current, which is raised in voltage by a transformer, and the second to convert it back to direct current. These are used mainly in portable equipment run by batteries.

Motor generators known as dynamotors are used to supply direct current where high power at moderate voltage is desired.

METER PROTECTION. Current meters associated with the power supplies can stand about ten times their full scale reading before they

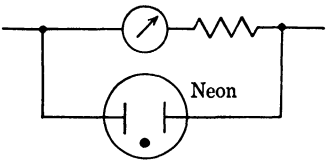


Fig. 19. Meter protection circuit.

fail. To prevent an overload, a neon bulb can be connected so that it will break down and shunt the current around the meter when the meter current reaches some given value. Thermocouple meters have a considerably smaller safety factor and should always be provided with meter protection.

BIBLIOGRAPHY

G. E. Tube Manual of Industrial Types.

Millman, Jacob, and Seely, Samuel, *Electronics.* New York: McGraw-Hill Book Co., Inc., 1951.

RCA Tube Handbook.

Ryder, John D., *Electronic Engineering Principles,* 2d ed. New York: Prentice-Hall, Inc., 1952.

Seely, Samuel, *Electron-tube Circuits.* New York: McGraw-Hill Book Co., Inc., 1950.

Terman, Frederick E., *Radio Engineering.* New York: McGraw-Hill Book Co., Inc., 1947.

Terman, Frederick E., *Radio Engineers' Handbook.* New York: McGraw-Hill Book Co., Inc., 1943.

6

OSCILLATORS

An oscillator is a device that converts d-c power to a-c power. Frequencies obtainable range from 0.01 cycle per second to 30,000 megacycles per second, and power from fractions of 1 to more than 1000 kw.

Oscillators may be classified in three main groups: feedback, negative resistance, and relaxation. The feedback oscillator is essentially an amplifier arranged to get exciting voltage of the right phase and magnitude from its own output. It derives its characteristic frequency

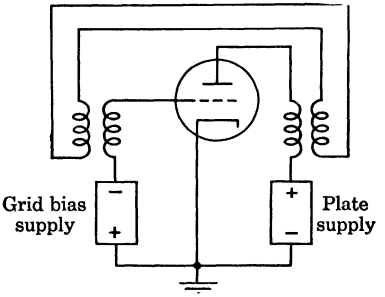


Fig. 1. Feedback oscillator.

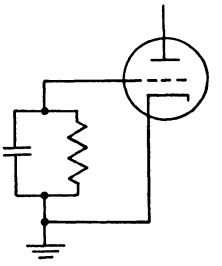


Fig. 2. Grid leak bias circuit.

from a resonant or "tank" circuit, usually a capacitor and inductor in parallel, which tends to eliminate other frequencies. Once a signal is started, it increases in magnitude until losses in the tank circuit and load become equal to the power developed in the tube.

As with amplifiers designed to work at one frequency, most efficient operation is attained when the tubes are in class C operation. In this type of operation, plate current flows over a phase angle of only around 100°, so that rather short, sharp pulses are delivered to the tank.

A simplified schematic of a feedback oscillator is shown in Fig. 1.

Fixed grid bias is not ordinarily used; instead, bias is derived from the d-c component of grid current by putting a resistor or "gridleak" between grid and ground. The a-c components are by-passed by means of a shunt capacitor (Fig. 2). This system has great advantages. When

the oscillator is first turned on, there is no bias, and consequently a high g_m, so that any small noise disturbance is sufficient to start oscillations. The bias then adjusts itself to the proper value, as grid current starts to flow. The bias is also able to adjust itself to changes in the load.

Fixed bias oscillators may not be self-starting if the operating point is close to or beyond cutoff. In such a case, it is necessary to reduce the bias until oscillations start and then return it to the proper operating value, or to induce a large transient in the circuit by some means.

Some of the most common feedback oscillator circuits are shown in Fig. 3, with bias and plate supplies omitted.

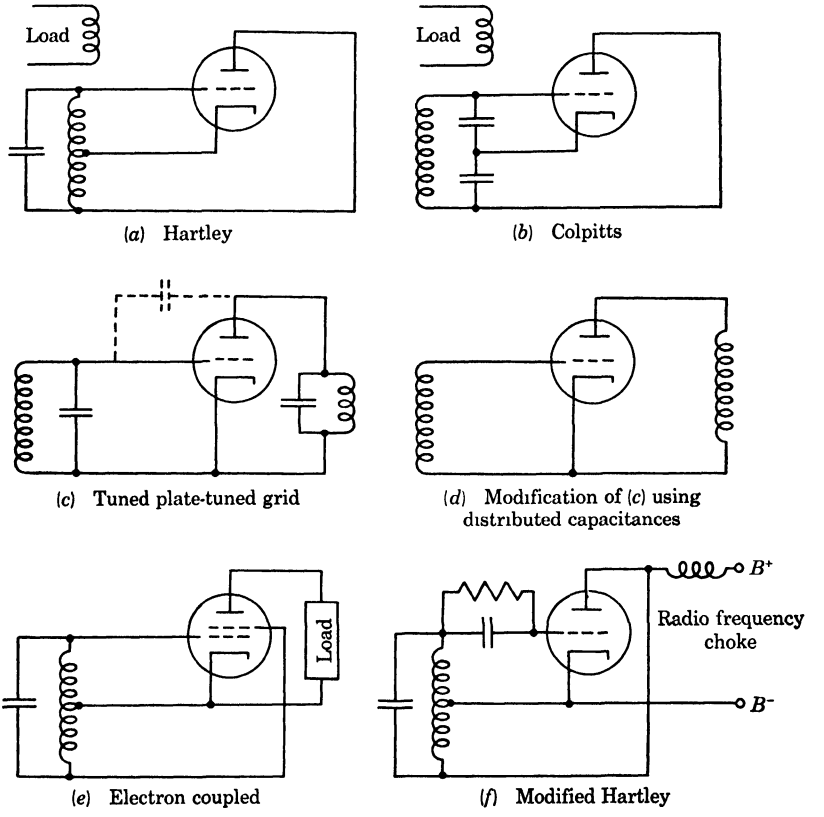


Fig. 3. Oscillator circuits.

In the electron-coupled oscillator, the screen grid acts as an anode for the oscillator. A large part of the electron current from the cathode goes on to the plate, however, and it is this current, modulated by the grids, that delivers a-c power to the load. The frequency of this oscillator is quite stable with regard to changes in the load.

Grounded grid type of oscillator to feed power into a transmission line is seen in Fig. 4.

As oscillators are of the order of 60% efficient (or less), much heat is dissipated in a tube handling large amounts of power. Water cooling is often necessary.

The efficiency of an oscillator does not depend much on the μ of the tube. However, if μ is as low as 10, the tube will draw a destructive amount of plate current if the oscillations stop. One must bear this in mind if he is using a variable load, which may become great enough to stop the oscillations.

HIGH FREQUENCIES. At very high frequencies, the distributed capacitances of vacuum tubes and associated circuits become very important. Great difficulty is introduced when they become resonant.

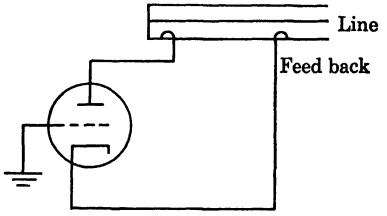


Fig. 4. Grounded grid oscillator.

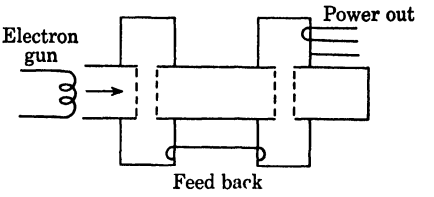


Fig. 5. Klystron oscillator.

In order to generate frequencies higher than about 100 megacycles it is generally necessary to make use of harmonics of the fundamental circuit frequency, and at 500 megacycles it is difficult to generate more than 1 kw in a single tube. However, with various devices such as the kystron or the magnetron, it is possible to obtain many kilowatts at frequencies up to 1000 or 2000 megacycles.

To generate ultrahigh frequencies, it is necessary to use small tubes in order to minimize lead inductances and interelectrode capacitances, and the transit time of the electron current in the tube. The small size often necessitates water cooling if much power is to be dissipated. Transmission lines or resonant cavities are generally used as tank circuits.

The klystron is an important device for generating high frequencies. In it, a beam of electrons passes successively through two resonant cavities tuned to the same frequency. The alternating electric field in the first cavity has no effect on those electrons which pass through when the field is zero. However, those which arrive somewhat earlier are slowed down, and those which arrive later are speeded up. This produces bunching at a series of positions along the path of the par-

ticles. The second cavity is placed at such a position, so that the bunches of electrons passing through deliver power to it, causing strong oscillations to be set up in it. Enough power is then piped back to the first cavity to maintain its oscillations. In the reflex klystron, which is also commonly used, the beam is reflected back through the bunching cavity, which then itself becomes the "second" cavity.

Another ultrahigh frequency generator makes use of electron (Barkhausen) oscillations in triode tubes. The electrodes are operated at voltages that permit electrons to oscillate back and forth through the grid, delivering energy to it.

There are two types of magnetron oscillators; one makes use of oscillations of electrons in crossed electric and magnetic fields, producing an effect similar to that of the Barkhausen oscillations. The other type makes use of a negative resistance effect produced by the cycloidal motions of electrons in crossed electric and magnetic fields. Electrons attracted by an electrode at high potential actually arrive at an electrode of low potential.

FREQUENCY STABILITY. The frequency of an oscillator is almost, but not quite, the resonant frequency of the tank circuit. The difference is due to perturbations resulting from the rest of the circuit being coupled on. It is evident that changes in the load, as well as temperature and other effects in the various circuit elements, will result in changes in the oscillator frequency. It has been found that for maximum frequency stability, the ratio of the energy stored in the tank to that stored in the various tube capacitances should be as large as possible. The lower the resonant frequency, at least down to a few hundred cycles per second, the easier this condition is to obtain. For this reason, low frequency oscillators are used with frequency multipliers when good stability is desired. Making the tank Q as high as possible is also necessary in maximizing the above ratio.

One means of minimizing the effect of load changes on frequency is use of an amplifier between oscillator and load. The electron-coupled oscillator has already been mentioned as a type in which the coupling between load and oscillator is small.

An oscillator which is to have a high degree of frequency stability generally makes use of a piezoelectric crystal in place of the LC tank circuit. The crystal vibrates mechanically at a definite frequency and is coupled to the circuit by means of the piezoelectric effect. The use of such a crystal makes frequency regulation to 1 part in 10^7 possible. The fundamental frequency of a crystal is generally not over 10 megacycles. To reach higher frequencies, frequency multipliers must be used.

OSCILLATOR DISEASES.

1. *Relaxation oscillations resulting from RC time at grid leak.* The circuit may oscillate for a few thousand cycles, then block, wait for a short time, and repeat the process. This happens when the grid leak time constant is too high for the bias to be able to follow random effects in the circuit rapidly enough. The capacitance should be made smaller.

2. *Blocking.* The phenomenon of blocking occurs when the oscillations stop suddenly, and the grid current reverses. The plate current increases to a large value and often results in the destruction of a high power tube. This difficulty occurs when the secondary electron emission at the grid becomes greater than the current from the grid to the cathode. The current through the grid leak is reversed, causing positive bias and very high plate current.

This trouble is usually started by attempting to put too great a load on the oscillator. It can be prevented by treating the tube grids to prevent excessive secondary emission, or by putting a rectifier in series with the grid leak, eliminating all possibility of current reversal.

3. *Parasitics.* Undesirable oscillations at various frequencies, originating in miscellaneous distributed capacitances and inductances in circuits and associated objects, are known under this name. Besides distorting the desired output voltage, they cause sparks, flashover, and heating effects.

Parasitics are investigated by moving a neon light around the oscillator circuit, or merely by drawing sparks to wire on a stick. The points of highest voltage are found, the frequencies are measured, and the equivalent circuit is deduced. It is usually found that single loops of wire are acting as inductances in the equivalent circuits, producing the parasitic oscillations of high frequency. Small interelectrode and interelement capacitances also become very important. Radio frequency chokes may bring about low-frequency parasitics. When the causes of the parasitics have been analyzed, an attempt can be made to eliminate them without causing others to start.

BIBLIOGRAPHY

Brainerd, John G., *Ultra-High Frequency Techniques.* New York: D. Van Nostrand Co., Inc., 1942.

Martin, T., *Ultra-High Frequency Engineering.* New York: Prentice-Hall, Inc., 1950.

Reich, Herbert J., *Theory and Applications of Electron Tubes.* New York: McGraw-Hill Book Co., Inc., 1944.

Sarbacher, Robert I., and Edson, William A., *Hyper and Ultra-high Frequency Engineering.* New York: John Wiley & Sons, Inc., 1943.

Terman, Frederick E., *Radio Engineering.* New York: McGraw-Hill Book Co., Inc., 1947.

Terman, Frederick E., *Radio Engineers' Handbook.* New York: McGraw-Hill Book Co., Inc., 1943.

7

MAGNET DESIGN

ELECTROMAGNET DESIGN

INTRODUCTION. The design of an electromagnet will be outlined according to the steps that one might follow in the design of a cyclotron magnet. The final dimensions of such a magnet will be a com-

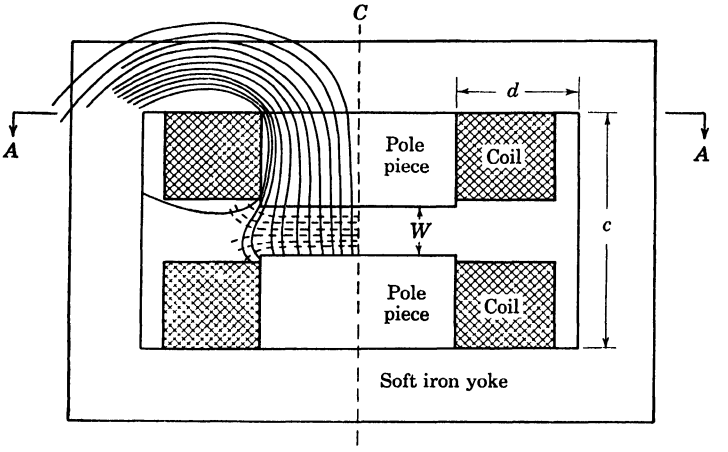


Fig. 1. Cyclotron magnet.

promise among many considerations. For simplicity we will assume that the width W in Fig. 1 of the magnet gap has been fixed by space requirements for accelerating equipment, and that the dimensions of the inside of the yoke, c and d, have been set by the space requirements for auxiliary equipment and access room. Of course the ideal design would have no space inside the yoke around the pole piece and coils, as shown in Fig. 2. This idealized design will be approached in the Bevatron magnet, since the auxiliary equipment will be located in the straight sections.

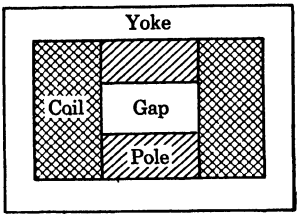


Fig. 2. Ideal magnet.

FUNDAMENTAL RELATIONS. The fundamental equations used in electromagnet design are:

$$\oint H \cdot dl = \text{ampere turns}/2.02 = \text{magnetomotive force} \tag{1}$$

where H is in oersteds and l is in inches.

$$B = \mu H \tag{2}$$

where B is in gausses and H is in oersteds. The slope of the soft iron magnetization curve is μ. In the soft iron yoke and pole piece

$$\phi = B \times \text{total area} \tag{3}$$

Equation (3) defines the flux. The equations above, plus a plot of the magnetization curve, B against H for soft iron, are used to calculate the performance of the magnet. In addition to these, an estimate must be made of the leakage flux.

ESTIMATION OF LEAKAGE. The leakage coefficient is defined as

$$\text{leakage coefficient} = \frac{\text{flux in pole}}{\text{flux in gap}} \tag{4}$$

It is naturally desirable to make this coefficient as small as possible. The leakage flux can be estimated by a graphical method. A scale drawing of the pole piece and yoke, as in Fig. 1, is made. The induction in the gap is assumed to be approximately constant over the area of the gap and equal to its value at the center. A scale is arbitrarily chosen and parallel lines of B are drawn, as in Fig. 1. Lines of constant magnetomotive force (mmf) are drawn normal to these induction lines and equally spaced in this uniform field region. This grid of straight lines can extend only to a point perhaps 90% of the radius out from the center of the pole piece. The fringing of the field can be estimated by continuing to draw the induction and mmf lines, always making them normal to each other at each intersection. Although over most of its area the pole face is a surface of equal mmf, the cylindrical sides of the pole pieces are not equal mmf surfaces; therefore the induction lines leave them obliquely, as shown. If one has a little experience and takes care to maintain the orthogonality of the lines, the cross section of a field can be estimated with sufficient accuracy so that the flux leakage so estimated will be accurate to 10%. The three-dimensional leakage flux is calculated by counting the induction lines, each multiplied by its radius from the center, that cross the merdian plane in the straight field zone, and dividing this count by the flux that passes through the pole of the magnet near the center of the coils, at section

AA in Fig. 1. The leakage coefficient in the magnet shown is probably 2.2 to 2.7.

If the gap width is reduced the flux leakage is also reduced. There is also flux leakage on the outside of the yoke, but this leakage is an advantage, since with a given number of ampere turns more flux can go through the pole piece with less iron in the yoke if there is outside leakage. The geometrical center of the cross section of the coils should be as close to the pole sides as possible to reduce flux leakage paths in the coils. A disadvantage of flux leakage from the yoke is the stray magnetic field created in the vicinity of the magnet.

The efficiency of the magnet is defined by the relation

$$\text{efficiency of magnet} = \frac{\text{mmf in gap}}{\text{total mmf around circuit}} = \frac{2.02\int H_g \cdot dl_g}{\text{ampere turns}} = \frac{\int H_g \cdot dl_g}{\int H_m \cdot dl_m} \quad (5)$$

where g stands for gap and m for magnet. It is the object of the design to make $\int H_g \cdot dl_g$ as large as possible and thus to reduce $\int H_m \cdot dl_m$. The induction B is chosen in the gap and then, using the magnetization curve, H is evaluated along a closed path around in the yoke linking the coils. The line integral is taken, and the efficiency is evaluated. The value of H in the iron can be made as large as desired by adding ampere turns, but to carry the induction much above 21,500 gausses is expensive in current and coils. If B is 18,000 gausses in the gap, it will probably be found that the flux density is too high next to the coils in the pole pieces. It is possible to build a magnet with an induction of 15,000 gausses throughout most of its air gap, without having excessive flux next to the coils.

Magnet models. It is usually desirable to build a scale model of the proposed magnet. If the linear dimensions of the iron and the coils are scaled down by a factor η, the magnetomotive force must also be scaled down by the same factor in order to keep the field intensity H constant. Under these conditions the iron will be operating at the same point on the magnetization curve in the model as it will be in the final design, and the field configurations will be identical except for the scale factor on the dimensions. The only difficulty with this procedure is that the current density in the coil *increases* by the factor η. The reason for this is that, although the actual current in the coils is reduced by the factor η (keeping the number of turns constant), the area of the wire decreases by a factor η^2 and the current density $j = I/A$ therefore increases by the scale factor. The power consumed in the

model is decreased by the factor η. Such scale models are always made for large magnets because of the difficulty in designing for specially shaped fields.

COILS. The coils of an electromagnet are usually made of copper, although during the last war, silver was available for such purposes. Characteristic values for both copper and silver are included in Table 1. The advantage of silver over copper is slight, and here we will consider only copper coils. There are large forces, given by (8) in Table 1, exerted on the coil windings in an electromagnet. The insulation and coil fastenings must be able to withstand these loads. The coil fastenings must be designed in such a way that they can preload the coils in such a direction with the current off, that when the current and resulting field come on, the magnetic field will exert a force on the conductor that partially, but not completely, unloads the fastenings. If this preloading is done properly, the conductor forces on the insulation will decrease with increasing current but never go to zero. Electromagnets have been built for pulsed operation, which have withstood 20,000 stress changes due to application of the magnetic field without failure of the coil insulation.

The economics of the cost of copper for the coils must be balanced against the cost of electric power. Equation (10) of Table 1 shows that the tons of copper in the coils are inversely proportional to the power consumption of the coils when the same field is being produced with coils of the same dimensions, except for the cross section of the conductor. It is usually found in the consideration of a complete magnet design that it is desirable to pack as much copper, on a volume basis, into the coil space as cooling requirements will allow. Once the cross section and length of the conductor and the power to be consumed in the coils have been chosen, the voltage of the generator is set by Equation (13) of Table 1.

COOLING THE COILS. The allowable current density, which is related to the power and coil weight by Eq. (11) of Table 1, is controlled by the cooling of the coils. The power, in the form of heat, that must be dissipated from a cubic inch of conductor, 1 in. long, is 0.732 w when the current density is 1 kiloamp per in.^{2} It is unfortunate that no nonelectrolytic cooling liquid with a high heat capacity is known. Water is used in many electromagnets, and when frequent hose connections are made to the coils, it is possible to cool coils that have as high a current density as 20,000 amp/in^2. In order to prevent electrolysis, the conductors are made in the form of tubes, to keep the water out of regions having voltage gradients. In order to allow them to be

TABLE 1
MAGNET DESIGN DATA

	Annealed copper	99.9% *Silver*
Physical properties of copper and silver		
1. Modulus of elasticity (lb/in.^{2})	17.5×10^6	11.5×10^6
2. Specific weight (lb/in.^{3})	0.322	0.380
3. Resistivity at 20°C (ohm in.)	0.679×10^{-6}	0.641×10^{-6}
Resistivity at 40°C	0.732×10^{-6}	0.694×10^{-6}
Resistivity at 55°C	0.772×10^{-6}	
4. Heat conductivity at 20°C (watts/°C in.)	9.76	10.52
5. Specific heat at 20°C (watt sec/°C lb)	174.9	105.9
6. Coefficient of linear thermal expansion (°C)	16.8×10^{-6}	18.8×10^{-6}

Relations that are independent of material

7. Amp turns = 2.02 × (gauss induction in gap) × (in. across gap)
8. Lb force on conductor = $\frac{1}{1750}$ × (kilogausses normal to conductor) × (in. length of conductor)
9. Lb force between pole faces = $\frac{1}{1.735}$ × (kilogausses)^2 × (area in in.^{2})

Relations for copper and silver coils at 40°C

10. Kilowatts × Tons (Cu/Ag) = (0.118 Cu/0.131 Ag)* × (megampere turns)^2 × (in. mean turn length)^2
11. Amp/in.^{2} = (469 Cu/525 Ag)* × (kw/ton)$^{\frac{1}{2}}$
12. In.^{2} conductor area
$$= \frac{(0.731 \text{ Cu}/0.698 \text{ Ag})^* \times (\text{megampere turns})(\text{in. mean turn length})}{(\text{volts on coils}) \times (\text{parallel coil paths})}$$
13. In.^{2} conductor area = $\frac{(2.130 \text{ Cu}/1.903 \text{ Ag})^* \times (\text{kw} \times \text{tons in coils})^{\frac{1}{2}}}{(\text{volts on coils}) \times (\text{parallel coil paths})}$
14. Power in watts liberated/in.^{3} conductor 1 in. long = (0.732 Cu/0.694 Ag)* × (kiloamp/in.^{2} conductor)^2

Relations for a rectangular conductor losing heat from two edges at 40°C

15. Heating at center of conductor °C = (0.00940 Cu/0.00821 Ag)*10^{-6} (in. width of conductor)^2 × (amp./in.^{2})^2
16. Watts transferred/in.^{2} edge surface = (0.366 Cu/0.345 Ag)* 10^{-6} (in. width of conductor) × (amp/in.^{2})^2

* The / between the Cu and Ag values means "either/or" and does not indicate that the Ag value is in the denominator. However, the / does mean a fraction when used between other expressions.

packed together more compactly, the conductors are usually rectangular tubes with rectangular or round water passages. There is an additional problem with water cooling, in that the coils will sweat if the air around them is humid and the cooling water is circulated with the current turned off. The condensation may cause short circuits between adjacent coils. This can be prevented by filling all space be-

tween the coils with a plastic so that no air will be present to cause condensation. One advantage of water cooling is that only a small heat exchanger is needed, or perhaps none at all, if an ample supply of pure water is available. With water cooling and careful coil design it is possible to secure a 65 to 70% volume density in the coils.

If it is decided that water cooling offers too many problems, oil or air cooling can be used. With rapid flow oil cooling, it is possible to operate coils at 2000 amp/in.^{2}, while with rapid flow air cooling 1000 amp/in.^{2} is the upper limit, with 700 amps/in.^{2} a more reasonable figure. All these current densities are for steady d-c operation. Pulsed operation may allow these d-c values to be exceeded. With air or oil cooling it is possible to achieve a 60% volume density of copper in the coils. It is important to use as thin insulation as the voltage difference between adjacent coils will allow. There is an advantage to thick conductors, since they allow less frequent insulators and yet do not cause any increase in potential drop between coils. Oil cooling or air cooling is usually arranged so that only the edges of the conductors are cooled, there being no clearance between them in one direction. There is an additional advantage to this arrangement, since it makes inspection of the windings simple. Oil cooling presents a serious fire hazard, but leaks in the cooling system are much easier to repair than leaks in a water cooled system, where the fire hazard is absent. Oil cooling requires a large heat exchanger. Air cooling requires the use of blowers and care must be taken to eliminate dust before the air is circulated around the coils. Of course no leakage or fire problems exist with air.

Shimming. If the pole piece of a magnet is a right circular cylinder, the field will begin to fringe quite a distance in from the edge because of the curvature of the constant mmf lines inside the pole piece, as shown in the left half of Fig. 3. To prevent this, it is customary to add shims to the pole face as seen in the right side of Fig. 3 in order to extend the region of uniform decrease of field as far as possible. The shims on the pole faces of the 184-in. magnet at the University of California Radiation Laboratory are similar to those shown. Martyn Foss[1,2] has shown theoretically that a magnet with a raised outer edge shim as seen in Fig. 3 can be made to have a useful radius out to 92% of the pole radius, while this can be increased to about 96% by cutting a groove into the pole face just inside the raised edge. Foss's treatment is incomplete in the sense that no method of determining the

[1] Foss, Martyn H., "Design of Cyclotron Magnets," *ONR Technical Report*, Carnegie Institute of Technology.

[2] Foss, Martyn H., "High Magnetic Fields," *ONR Technical Report*, Carnegie Institute of Technology.

shape of the pole is developed, but the problems encountered and the results to be expected are discussed.

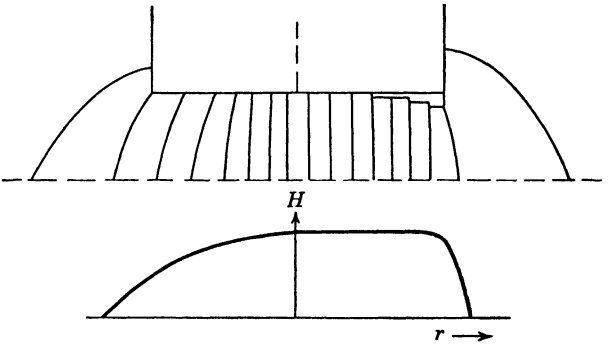


Fig. 3. Effect of shimming on magnetic field.

HELMHOLTZ COILS. If a very uniform field is desired over a limited region, Helmholtz coils can be used. Fields that are uniform and calculable to 0.15% can be produced with a Helmholtz coil arrangement with no iron. Foss[3] and Koch[4] have also worked on this problem.

Recent magnet design, as mentioned by Oliphant,[5] has favored large conductors and large currents. The large currents are produced by homopolar generators. The main problem is to bring leads out of the coil without the field of the leads distorting the main field.

PERMANENT MAGNET DESIGN

INTRODUCTION. Permanent magnets are limited in size and are initially expensive, but they possess many desirable features. They require no power supply or cooling apparatus. The lack of heat makes them especially suited to cloud chamber work. Because they require no power supply, they are ideal for portable experiments. When the aggregate cost of power, power supply, and cooling equipment is compared with the cost of a permanent magnet, there are many cases in which it is cheaper in the long run to pay the initially high cost per pound of permanent magnet material than it is to buy copper coils and soft iron. In addition, if a permanent magnet is properly stabilized, the field is unusually constant. For example, an Alnico II permanent magnet used in a wattmeter will, if it is not subjected to rough handling, have a field that will be constant to 0.1 or 0.2% over a century. The main limitation of permanent magnets is that they cannot supply

[3] Foss, Martyn H., "Cloud Chamber Coils," *ONR Technical Report*, Carnegie Institute of Technology.

[4] Koch, H. W., "Cloud Chamber Coils," *J. Appl. Phys.*, *21*, 387 (1950).

[5] Oliphant, M. L., "The Cyclosynchrotron," *Nature*, *165*, 466 (March 25, 1950).

a large field in a large gap volume. However, they can furnish a field as high as 30,000 gausses in a small volume.

DESIGN EQUATIONS FOR A PERMANENT MAGNET. We will develop two equations which, in conjunction with a demagnetization curve for the magnetic material that is used in the magnet, are all that is needed for a complete solution of the design problem. However, to reach a final design requires some experience with magnets, certain simplifying assumptions, empirical results, and then carrying out successive approximations to reach a final solution. The flux ϕ passing through part of a magnetic circuit is

$$\phi = \frac{(\int \vec{H} \cdot \vec{dl})}{R} \quad \text{where} \quad R = \frac{1}{\mu A} \quad \text{is the reluctance} \tag{6}$$

Here H is the magnetic field, l is the distance along the magnetic circuit, and A is the cross-sectional area of the magnetic circuit, along which the reluctance is R. In a permanent magnet, Maxwell's equation for H is

$$\vec{\nabla} \times \vec{H} = 0 \quad \text{and consequently} \quad \oint \vec{H} \cdot \vec{dl} = 0 \tag{7}$$

From this, if H is uniform and there is no leakage,

$$H_g l_g = H_m l_m \tag{8}$$

where the g refers to the gap and m to the magnet paths. Since there is leakage, Eq. (8) is made to correspond to an actual magnet by introducing a constant f.

$$f H_g l_g = H_m l_m; \quad \text{or, since} \quad H_g = B_g; \quad f B_g l_g = H_m l_m \tag{9}$$

Defining the permeance $P = 1/R = \mu A/l$, we have from (6) that

$$\phi = \Sigma P \int \vec{H} \cdot \vec{dl} \tag{10}$$

where the summation is carried over all parts of the circuit over which the line integral is carried. Evaluating (10) for the flux through the magnet, and multiplying the right side by $1 = A_g/l_g P_g$, which is taken from the definition of the permeance, and then substituting from (9), and then from (6): $\phi = H_g l_g A_g / l_g$, we have

$$\phi_m = \Sigma P H_m l_m = H_m l_m A_g \Sigma P / l_g P_g = f H_g l_g A_g \Sigma P / l_g P_g = \phi_g f \Sigma P / P_g \tag{11}$$

We define $F = f \Sigma P / P_g$, and from (11),

$$\phi_m = \phi_g F; \quad B_m A_m = F H_g A_g \tag{12}$$

The two design equations are then (9) and (12). Equation (9) is a linear relation that makes the magnetic potential return to its starting value in the course of one transit of the circuit, while Eq. (12) states that if due allowance is made for flux leakage, the flux through any cross section of the circuit is the same as through any other. Sanford[6] gives equations for the P's, and Evershed[7] gives P's for many arrangements met in practice.

APPROXIMATE SOLUTION OF THE DESIGN EQUATIONS. The usual design problem consists of finding the length and area of a magnet that will supply a given flux density B_g throughout an air gap of

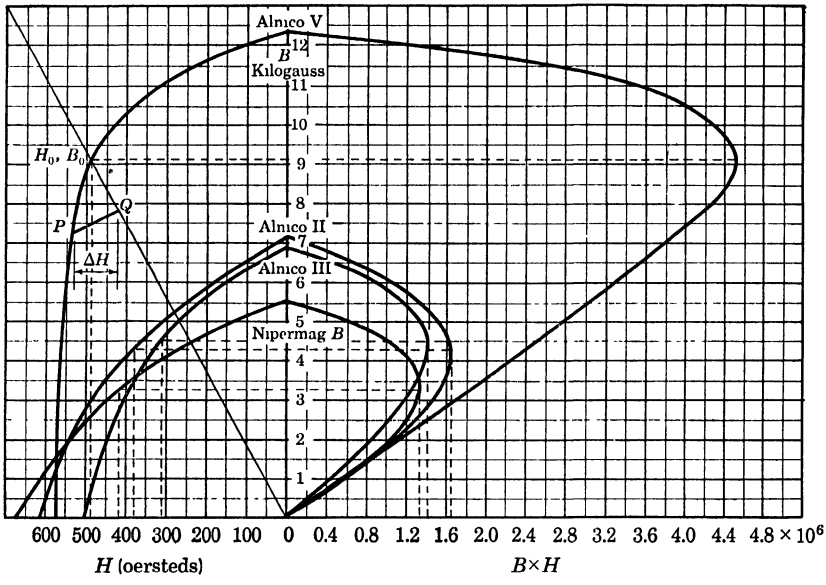


Fig. 4. Characteristics of permanent magnet alloys.

volume $A_g l_g$. Assuming that f and F in (9) and (12) are known, A_m and l_m, the desired dimensions, cannot be solved for until the values of H_m and B_m have been chosen. The values of H_m and of B_m at which a permanent magnet will operate will lie on its demagnetization curve. If the magnet is not associated with any other material and is a closed path in itself, it will be at B_r, if it is made of Alnico V for example, as shown in Fig. 4. If some air or other material such as soft iron is introduced into the magnetic circuit, the operating point will shift to the left to some value of B and H on the demagnetization curve.

[6] Sanford, R. L., "Permanent Magnets," *Publication C448*, National Bureau of Standards, 1944.

[7] Evershed, S. J., *J. Inst. Elec. Engrs.* (London), *58*, 780 (1920); *63*, 725 (1925).

By properly choosing the dimensions of the magnetic circuit, one can make the operating point lie anywhere on the curve. The optimum place for the operating point to lie can be chosen by solving (9) for l_m and (12) for A_m and multiplying these together.

$$A_m l_m = \frac{FfB_g^2 A_g l_g}{B_m H_m}; \quad \text{volume of magnet} = (\text{volume of air gap}) \frac{FfB_g^2}{B_m H_m} \tag{13}$$

In order to minimize the volume and the resulting cost of the magnet, H_m and B_m are chosen so that their product will be a maximum. This is the condition for maximum energy storage in the magnet. The proper operating points, B_0 and H_0, for various magnetic materials can be seen on the energy product curve on the right side of Fig. 4 where the BH vs. B curves have their maximum value to the right. Now that we have fixed B_g, A_g, l_g, H_m, and B_m, we can solve (9) and (12) for A_m and l_m, and the magnet is completely determined. Unfortunately, we assumed that we knew f and F. It is in the estimation of these parameters that the difficulty lies in magnet design.

ESTIMATION OF f AND F. The estimation of f and F actually involves evaluation of the leakage flux from the magnet and the air gap. In order to know the leakage, we must know the dimensions of the magnetic circuit, and in order to know these, we must know the leakage and so on. Thus the problem resolves into one of successive approximations. Underhill[8] outlines a method for determining F from empirical relations that he has established. He points out that f varies from 1.2 to 1.5, but that the extremes are seldom met in practice. If it is assumed that $f = 1.35$, the results will be satisfactory for almost any permanent magnet circuit.

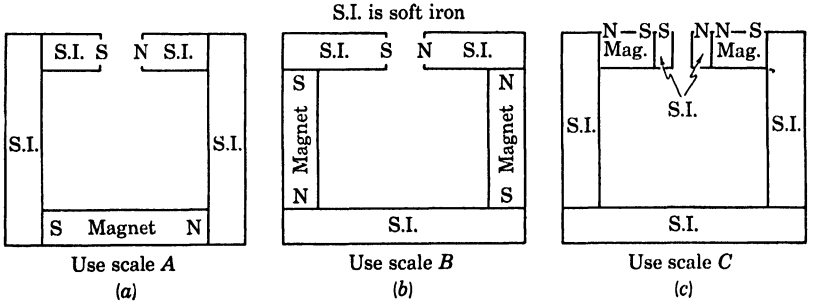


Fig. 5. Typical arrangements of permanent magnets.

On the other hand, F has a larger range. The minimum value of F is approximately 2 and the upper limit for conventional magnets is

[8] Underhill, E. M., *Electronics*, Dec. 1943, Jan. 1944, Feb. 1944, April 1944, Jan. 1948.

about 10, but there is no upper theoretical limit. In order to estimate F, a guess is made about the circuit, assuming a value for F of, for example, 5. The view of the magnet circuit that is parallel to the flux lines is laid out to scale and two circles are drawn on top of the main view as shown in Fig. 5a. The inner circle has approximately the area of the inside of the circuit and the outer circle has approximately the area of the inside of the circuit plus the area of the iron. Then R_e is the ratio of the radius of the outer circle to the radius of the inner circle.

Curve A, B, or C in Fig. 6 is chosen, depending on whether the proposed circuit is most like Fig. 5a, 5b, or 5c, respectively. The point

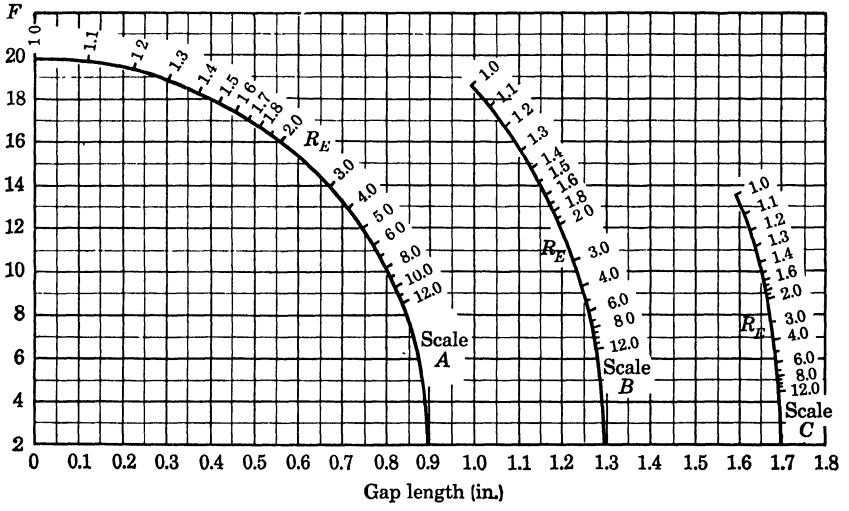


Fig. 6. Nomogram for leakage estimates.

on the A, B, or C curve so chosen, having the value of R_e measured as described above, is then connected to the origin at (0,2) with a straight line. This straight line then gives the values of F as a function of the corresponding air gap lengths. With the new value of F corresponding to the desired air gap we can now redraw the proposed magnet circuit and redetermine the value of F. This successive approach to the solution is greatly facilitated by the nomograph shown in Fig. 7. The nomograph is used by laying a straightedge across from X to X' or from Y to Y'. For example, if B_m and B_g are known, the straightedge is laid from the value of B_m on the Y' scale to the value of B_g on the diagonal scale, and the corresponding value of FA_g/A_m is found on the Y scale. Similarly, if H_m and B_g are known, the straightedge is laid from the value of H_m on the X scale to the value of B_g on the diagonal scale and the value of fl_g/l_m is found on the X' scale. The successive

approximations are thus much more rapidly made. The demagnetization curves for the four magnetic materials are included on Fig. 7 to make it easy to pick properly related values of H_m and B_m.

It should be noted that Alnico V does not always give the highest flux in a magnetic circuit. For low values of H_m, Alnico V is quite superior, but for values of H_m above 580 oersteds, Nipermag B and

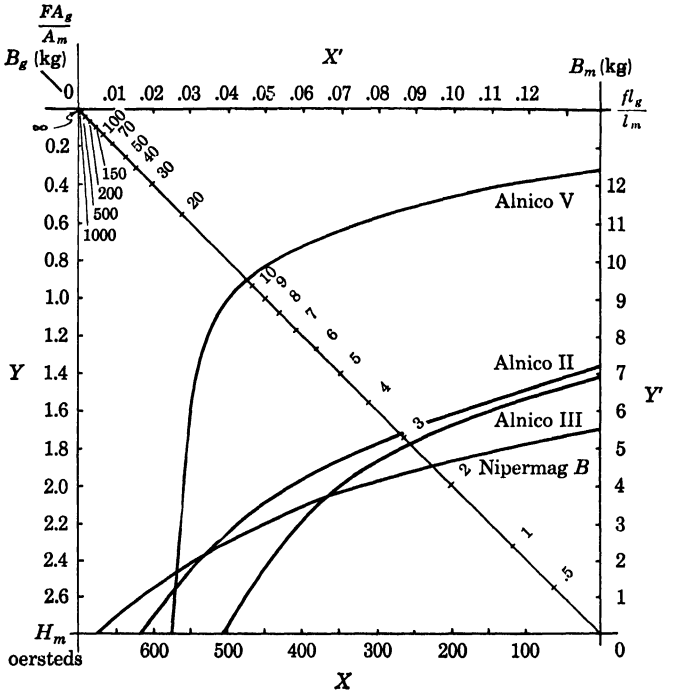


Fig. 7. Demagnetization curves.

Alnico II are the only useful alloys. However, for most purposes, Alnico V gives the highest flux.

Models. It should be emphasized that the method outlined above for determining F is based on empirical results and cannot give precise values. It is therefore important to build a model, and to measure f and F experimentally. It is difficult to get a true scale model of the Alnico sections of the magnetic circuit, since the Alnico heat treatment depends on the size of the piece. In addition, the degree of contact between the Alnico and the soft iron will in general be different in the model than it will be in the finished magnet. This may introduce errors that are difficult to correct.

STABILIZATION. After a permanent magnet is magnetized it must be stabilized against vibration, temperature change, and external fields if it is to operate in a reproducible manner under these external influences. Vibration stability can be achieved by vibrating the magnet. Temperature cycling from −200° to +200°C will make the magnet's temperature behavior reproducible. Stability against small external fields can be achieved by carrying the demagnetization to the left of (H_0,B_0) in Fig. 4 to the point P. Then minor field changes of the order of ΔH will only cause the operating point to move from P, on the demagnetization curve to Q, on the sheering line, along a minor hysteresis loop in such a way that the slope of PQ is approximately the same as the slope of the demagnetization curve at B_r.

BIBLIOGRAPHY

ELECTROMAGNETS

Foss, Martyn H., "Cloud Chamber Coils," *ONR Technical Report*, Carnegie Institute of Technology.

Foss, Martyn H., "Design of Cyclotron Magnets," *ONR Technical Report*, Carnegie Institute of Technology.

Foss, Martyn H., "High Magnetic Fields," *ONR Technical Report*, Carnegie Institute of Technology.

Koch, H. W., "Cloud Chamber Coils," *J. Appl. Phys.*, *21*, 387 (1950).

Roters, H. C., *Electromagnetic Devices*. New York: John Wiley & Sons, Inc., 1941.

PERMANENT MAGNETS

Harnwell, G. P., *Principles of Electricity and Electromagnetism*. New York: McGraw-Hill Book Co., Inc., 1949.

Permanent Magnet Design Manual. General Electric Circular CDP, 609.

Sanford, R. L., "Permanent Magnets," *Publication C448*, National Bureau of Standards, 1944.

Shutt, R. P. and Whittemore, W., "Design of Permanent Magnets for Cloud Chamber Work," *Rev. Sci. Instr.*, *22*, 73 (1951).

Underhill, E. M., *Electronics*, Dec. 1943, Jan. 1944, Feb. 1944, April, 1944, Jan. 1948.

8

VACUUM TECHNIQUES

The subject of vacuum techniques will be subdivided into the three categories: (1) evacuation; (2) pressure measurement; (3) leak hunting, etc.

For authoritative and exhaustive treatments of the many problems of vacuum techniques, the excellent books by Dushman and by Guthrie and Wakerling should be consulted.

EVACUATION

The flow of gas through pipes and orifices is governed by the equation

$$Q = F(P_2 - P_1) \tag{1}$$

where Q is proportional to the rate of flow of mass through the tube. It is given in terms of the volume of the gas, measured at atmospheric pressure, that passes a point per second, and has the units of $(cm^3\ atm)/sec$. It is often given in other units such as liter microns/sec, cu ft microns/min, etc. The quantity F is analogous to conductance and in general depends on the pressure and the geometry of the system. It is commonly measured in liters per second.

Regions of flow. *Streamline or laminar viscous flow.* In regions of about 100 microns and above, the gases follow viscous flow and F is a linear function of the average pressure in the tube.

Turbulent viscous flow. Turbulent flow need never be considered in vacuum technique as it occurs in the viscous region only with speeds of flow never attainable in ordinary evacuation methods.

Free molecular flow. In regions where the mean free path is much larger than the transverse dimension of the tube, F is independent of pressure and depends only on the geometry of the tubing. Means of determining F will be given below. A further description of the free molecular region is useful. The transition from viscous to free molecular flow occurs when

$$5.5\frac{L}{a} = 1\ .$$

And free molecular flow certainly obtains when

$$L = \tfrac{2}{3}a$$

where L is the mean free path and a is the transverse radius.

In vacuum technique the "fore vacuum" or "roughing down" region is viscous flow, whereas the high vacuum region ($<10^{-2}$ microns) is free molecular flow.

DETERMINATION OF F OF TUBES AND ORIFICES. Since F is not independent of pressure in the viscous flow region, calculation of F for various geometries would apparently be applicable only in the free molecular flow region. However, use is made of these F's in both regions by approximate formulas as follows.

F for a long tube. The value of F for a long tube of constant circular cross section is[1]

$$F = 12.1\frac{D^3}{L} \text{ liters/sec} \qquad (D \text{ and } L \text{ in centimeters; air at } 20°\text{C}) \quad (2)$$

F for an orifice. The value of F for an opening of cross section A (any shape) is

$$F = 11.6A \text{ liters/sec} \quad (A = \text{area in cm}^2\text{; air at } 20°\text{C}) \quad (3)$$

F for combinations of orifices and tubes. Since F corresponds to conductance, the total conductance of a *system* of tubes and orifices is computed in the usual manner.

$$\frac{1}{F_{\text{total}}} = \frac{1}{F_1} + \frac{1}{F_2} + \cdots \quad (4)$$

F for a short tube. For tubes which are short ($l/a < 10$) relative to their end openings, end effects make the above relation invalid. Dushman's book[2] gives a complete discussion of the end corrections.

THE EFFECT ON F OF BENDS IN TUBES. It should be noted that the conductance calculation has taken *no* account of bends in the tubing but has to do only with total length of path and geometry of the cross section. Thus in the molecular flow region F depends *only* on the total length of path and on cross sections of the pipe, but never on the type of bends.

[1] Guthrie, A. and Wakerling, R. R., *Vacuum Equipment and Techniques*. New York: McGraw-Hill Book Co., Inc., 1949.

[2] Dushman, Saul, *Scientific Foundations of Vacuum Technique*. New York: John Wiley and Sons, Inc., 1949.

CONVENIENT FORM OF F CALCULATION FOR AIR, 25°C. The F of a tube in the viscous flow region is[2]

$$F = 0.182 \frac{D^4}{L} \bar{P} \tag{5}$$

where P is in microns of mercury; D is in centimeters; D = diameter, L = length; $\bar{P}$ = average pressure.

PUMPS. Three characteristics of a pump must be considered in designing a high vacuum system:

1. The exhaust or back pressure.
2. The lower limit of pressure on the vacuum side.
3. Speed of pump.

SPEED. The speed S of a pump is defined as the conductance in the usual relation:

$$-\frac{dP}{dt} = \frac{S}{V_0}(P - P_s) \tag{6}$$

where V_0 is the volume to be exhausted; P is pressure in this volume; P_s is lower limit of pressure attainable by pump. Equation (6) when S is independent of P integrates to

$$(P_t - P_s) = (P_{t=t_0} - P_s)e^{-St/V_0} \tag{7}$$

Thus the pressure in the system falls exponentially to the lower limit of the pump.

The time for reducing the system from the original pressure (at the time t_0) to P_t (at the time t) is

$$(t - t_0) = \frac{V_0}{S} \ln \frac{P_{t_0} - P_s}{P_t - P_s} \tag{8}$$

This is actually equivalent to

$$Q = S(P_{V_0} - P_s) \tag{8a}$$

The pressure at the intake of the pump is not the pressure in the given volume V_0, since there is a drop between the volume and the pump due to gas flow through the connecting pipes. This is taken into account by making use of the fact that the pumping speed S is given by the nominal speed S_p combined with the conductance of the connecting pipes.

$$\frac{1}{S_e} = \frac{1}{S_p} + \frac{1}{F} \tag{9}$$

where S_e is effective pumping speed. Thus it is clear that large diameter connecting pipes are desirable.

The most useful equations used in selecting a pump are

$$\frac{dP_{V_0}}{dt} = \frac{S_e}{V_0}(P_{V_0} - P_s) \tag{10}$$

$$(t - t_0) = \frac{V_0}{S_e} \ln \frac{P_{t_0} - P_s}{P_t - P_s} \tag{11}$$

$$Q = S_e(P_{V_0} - P_s) \tag{12}$$

$$\frac{1}{S_e} = \frac{1}{S_p} + \frac{1}{F} \tag{9}$$

EXHAUST AND MINIMUM PRESSURES. Forepumps, of which some types will be described below, can be worked between atmospheric pressure and very low but not extreme pressures. However, extremely high vacuums are obtained by high vacuum pumps which will perform only against low exhaust pressures. For this reason evacuating equipment often consists of a forepump in series with a high vacuum pump.

EXAMPLE OF PUMPING SPEED DETERMINATION. Pump speeds for fore vacuum pump Z and high vacuum pump X are determined as follows.

1. The volume V_0 is to be exhausted.
2. Preceding the high vacuum pump are connections of conductance F_x [calculated by (2) through (4)].
3. Preceding the fore vacuum pump are connections of conductance F_z.
4. The rate at which air leaks into the system V_0 *at steady* state is Q.
5. The volume V_0 is to be at pressure P_{V_0}.

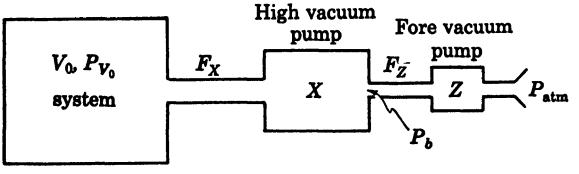


Fig. 1. Vacuum system.

The back or exhaust pressure of the high vacuum depends on the pump, but the choice is limited by the stipulation of P_{V_0}; thus P_b is determined, at least approximately.

Determination of S_p for X

$$\frac{1}{S_e} = \frac{1}{S_x} + \frac{1}{F_x}$$

and $Q = S_e(P_{V_0} - P_s)$. So, combining these two,

$$S_x = \frac{Q}{(P_{V_0} - P_s)} \left[\frac{1}{1 - \frac{Q}{(P_{V_0} - P_s)F_x}} \right]$$

Since P_s is determined by the models available in the high-vacuum region of P_{V_0}, and the other quantities are assumed known, we can solve for S_x. The complicated bracket can be estimated in practice.

Determining S_z involves the same relation with the appropriate substitutions.

PUMPING DOWN. Having thus determined the pumping speeds S_x and S_z, the time to reduce the system to P_{V_0} is now calculated by (8). The time to reduce the system to P_b, i.e., "roughing down" is

$$(t - t_0) = \frac{V_0}{S_{z\,\text{eff}}} \ln \frac{P_{\text{atm}} - P_{s_z}}{P_b - P_{s_z}}$$

The time to reduce from P_b to P_v is determined by using the same relation with the high vacuum pump values for minimum pressure and effective speed. The pumping down time for the forepump will probably be much longer than that for the high vacuum.

The rate of leak used in determining the pumping speeds can be estimated by measurement with a pump of known speed and connections of known conductance.

$$Q = S_{\text{eff}}(P_{V_0} - P_s)$$

Measurement of the pressure in the volume V_0 when steady state is reached with the pump of known speed S_e and minimum pressure P_s then gives Q_{leak}.

MEASUREMENT OF PUMP SPEED. *Speed of a forepump.* Connect a known volume to the intake of a forepump by a pipe of very large conductance. Measure the pressure in the volume by continuously reading pressure gage (e.g., Pirani gage). Then

$$(P_{V_0} - P_s)_{t=t} = (P_{V_0} - P_s)_{t=t_0} e^{-S(t-t_0)/V_0}$$

and if $P_s \ll P_{V_0}$, then S is determined.

Speed of a diffusion (high vacuum) pump. The intake of the pump is plugged except for a pressure gage and an adjustable leak with capillary tube. The leak is adjusted until the pressure gage reads the pressure at which the speed is to be measured. A slug of mercury is then put into

the capillary. The speed with which the mercury moves, together with the radius of the capillary, gives the flow rate at atmospheric pressure.

$$Q = P_{\text{atm}} \frac{V_{\text{capillary}}}{t}$$

The rate at which the air leaves the pump is

$$Q = S(P_{V_0} - P_s)$$

The pressure P_{V_0} is read on the pressure gage connected to the input. So, using both

$$\frac{V_{\text{cap}}}{t} P_{\text{atm}} = SP_{V_0}$$

Operating characteristics of diffusion pumps. Any good efficient diffusion pump has a speed S_p (when pumping to minimum attainable pressure P_s) equivalent to about 20% of the throat orifice conductance. The maximum speed would usually not exceed 40% of the throat orifice speed. Therefore the bigger the pump the faster. There appears to be no limit on the practicable size. A typical pump-down time is about one hour, which is the time necessary to reduce the pressure in a system from 100 microns to 10^{-5}.

It should be noted that maximum speed of a diffusion pump does not coincide with minimum pressure. If, for example, the speed decreases by 30%, the P_s (minimum attainable pressure) may go down by a factor of 5.

The limit to low pressure is the vapor pressure of the pumping vapor. The pressures attainable with oil diffusion pumps are about 10^{-6} mm. Such a pump, when used in conjunction with liquid air cooling brings the pressure down to about 10^{-8} mm. Mercury diffusion pumps used with liquid air cooling give the same order of vacuum: 10^{-8} mm.

Forepumps. Forepumps (or "backing pumps") are of the rotary oil-sealed type. Depending on manufacturer and model (see Martin and Hill[3] for a large listing) the minimum pressures run from 0.005 micron to 0.3 micron.

The speeds run from about 1 liter per minute to 1000 liters per minute; e.g., a Kinney pump of P_s equal to 0.2 micron pumps 1320 liters per minute at 1 atm.

Manufacturers specify volume pumped at 1 atm in rating their pumps. Another quantity used in specifying speeds is "volumetric

[3] Martin, L. H. and Hill, R. D., *A Manual of Vacuum Practice.* Melbourne: Melbourne University Press, 1949.

efficiency" which is (volume pumped)/(volume displaced). At 1 atm, volumetric efficiency is about 80%, while at $P = 0$ it falls to zero.

PRESSURE MEASUREMENT

Instruments used in measuring pressures depend for their action on the pressure itself, or on the relation between pressure and heat conductivity, electrical conductivity, density, viscosity, and so forth.

The most common gages which depend on the pressure itself are the U-tube manometer and the McCleod gage.

U-TUBE MANOMETER. The ordinary U-tube manometer is usually filled with mercury, which thus limits it to pressures above about 0.5 mm. With vacuum pump oil the lower limit can be pushed to 0.01 mm.

MCCLEOD GAGE. This gage depends on Boyle's law for pressure measurement. A sample of gas from the evacuated system is admitted to a known volume and then compressed under known pressure. The new volume of the sample gas indicates the pressure in the evacuated system. Mercury is utilized to provide the "known" pressure. When the gas sample is compressed, the nonpermanent gases may condense, and thus the gage should have a liquid air trap interposed between it and the evacuated volume.

The gage is accurate only to pressures above 10^{-5} mm usually, and certainly not below 10^{-6} mm. Since it requires manipulation it is *not* a continuously reading gage. It is, however, an absolute instrument.

Typical of gages that depend on the pressure only indirectly for their action are the Pirani, thermocouple, and ionization gages.

PIRANI GAGE. This instrument depends on the variation of heat conductivity of a gas with pressure. A constant potential is applied across a filament that is fixed in a bulb connected to the volume whose pressure is to be measured. The temperature of the bulb is held constant by means of a temperature bath. The varying pressure of the gas in the tube produces a varying heat loss from the filament. The temperature of the filament fluctuates therefore, and changes the resistance of the filament, which can be measured with a bridge circuit and calibrated in terms of pressure.

Since the heat conductivity of the gas depends on the gas, the Pirani must be calibrated against such an absolute gage as the McCleod.

The lower limit of pressure measurement with this type of gage is

about 0.01 micron, while it can be read up to about 150 microns. Its principal advantage is the continuous reading characteristic.

THERMOCOUPLE GAGE. This is another continuously reading gage operating on the same principle. The temperature of the filament is indicated by a thermojunction connected to it (Fig. 2), and the resulting emf indicates the pressure.

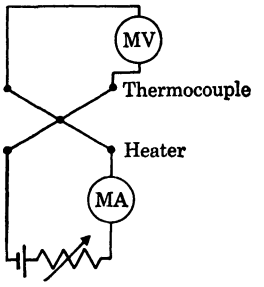


Fig. 2. Thermocouple gage.

IONIZATION GAGE. The ionization gage depends on the ionization of the gas in the system by electrons accelerated from a hot filament by a positive grid. Many electrons go through the grid and make positive ions in the space between the grid and the plate (Fig. 3a), which is held at a negative voltage. The current of positive ions is measured by a microammeter and indicates the pressure. The response of the instrument is linear with pressure variation only under certain electrode arrangements, and it must be calibrated against a McCleod gage for each gas mixture. It has the advantage of continuous measurement, and can be used, with various electrode arrangements, down to 10^{-7} micron.[4] A hot filament ionization gage is, however, subject to destruction if a sudden leak develops in the system, admitting air,

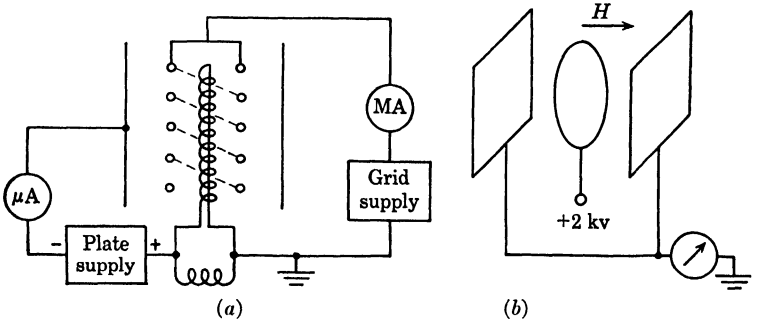


Fig. 3. Ionization gages: (a) hot filament gage; (b) Philips ion gage.

which rapidly oxidizes the hot filament. A safety device should be used that turns off the filament power when the vacuum is bad.

PHILIPS ION GAGE (Fig. 3b). This gage has no hot filament and can be left on even in atmospheric pressure. Nicknamed the PIG, this

[4] Bayard, R. T., and Alpert, D., "Extension of the Low Pressure Range of the Ionization Gauge," *Rev. Sci. Instr., 21,* 571 (1950).

gage must operate in a magnetic field, and is therefore convenient to use in magnetic field particle accelerators. The central ring, maintained at a positive potential, attracts the free electrons in the gas. The magnetic field constrains these electrons to move along the field lines, but they strike gas atoms, ionizing them and creating more electrons. The resulting positive ions are attracted toward the outer plates, and the ionization current, which depends on the pressure, is indicated in the meter. The PIG is not as sensitive as the hot filament ion gage, but is useful in the range 10^{-4} to 10^{-6} mm Hg.

LEAK HUNTING

An important point to be noted in connection with leak detection is the "relaxation time" of a vacuum system. An example will serve to explain its importance in hunting leaks. Suppose a continuously sensitive pressure gage is to be used to detect the leaks. The system is searched, say, by painting glyptal on suspicious joints; the pressure is read, and when the pressure drops we know we have plugged a leak and thus located a hole. However, the pressure gage will not show an immediate drop in pressure because of the exponential relation (8). The pressure would actually change as shown in Fig. 4. The pumping speed should be as rapid as possible and the pressure gage as sensitive and *rapidly responding* as possible for such a leak detection method. A brief description of leak detectors will indicate the general technique.

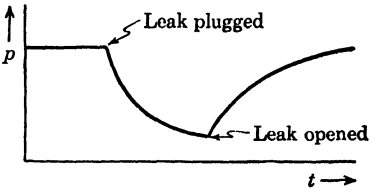


Fig. 4. Relaxation time of vacuum system.

REPLACEMENT VAPORS. This method depends on condensation in a liquid air trap of nonpermanent gases, such as acetone or water vapor. A liquid air trap is put on the input of the high vacuum line and a jet of such nonpermanent vapor is directed at a suspected leak. A continuously reading gage, such as an ionization gage, is observed as the jet is played over the area. If the vapor enters the system, it will be rapidly condensed in the liquid air trap. The trap thus acts as an exceedingly fast vacuum pump. The pressure drops, accordingly, and the hole is located. It should be noted that the relaxation time consideration mentioned above actually applies here, but the "speed" S of the "vacuum pump" (the liquid air trap) is so great that the response of the ionization gage is, in practice, immediate.

HIGH PRESSURE TESTS. These tests are made by filling the system with a particular gas under pressure above atmospheric. The leak is

located by making tests at suspected regions for the particular gas. Some of these testing devices are described below.

Halide-Freon leak detector. The system is filled with freon. The system is then searched by a torch whose air intake is a tube which serves as a probe. The probe is moved over the area of the system. If a leak exists, freon will leave the system and enter the probe through which it reaches the flame. The flame burns acetylene and is played on Cu. The freon in the flame colors it blue-green. Caution should be exercised in using the probe, to make sure that physical contact is not made with the surface, since some metal might be scraped off and a spurious coloring of the flame would result.

Soap solution. The system is filled with air under pressure, and the surface is covered with soap solution. If a leak exists, tiny fuzzy bubbles (not suds or large bubbles) will appear after a time.

Helium and mass spectrograph detector. The system is operated under the desired conditions. Some sort of helium detector, such as a mass spectrograph, is connected by a line to the system, and a helium jet is directed at the surface. (The detector registers the helium *some time* after the hole is actually hit by the jet, of course.) Helium is used, for one reason, because of its fast diffusion. The best helium detector is the mass spectrograph. If helium (or some other gas such as hydrogen, for instance) is used as indicated, the composition of the gas in the system can be analyzed by a mass spectrograph. The analysis can be recorded by an electronic device.

In conjunction with helium, a difficulty arises due to its low density and high diffusion rate. Since the helium spreads very rapidly over the surface, the location of the leak is hard to find, although the existence of a leak is indicated.

MISCELLANEOUS INFORMATION. Glyptal can be used to seal tiny leaks permanently or large holes temporarily. It can be wiped off easily with acetone, and is a good hole plugger in leak hunting.

Temporary liquid air trap. A temporary trap is installed to good advantage in leak hunting even though the system, when ordinarily running, is not affected adversely by condensable gases. The temporary trap condenses the nonpermanent gases, reducing the pressure, thus making leak hunting easier.

BIBLIOGRAPHY

Bachman, C. H., *Techniques in Experimental Electronics.* New York: John Wiley & Sons, Inc., 1948.

Dushman, Saul, *Scientific Foundations of Vacuum Technique.* New York: John Wiley & Sons, Inc., 1949.

Guthrie, A. and Wakerling, R. R., *Vacuum Equipment and Techniques.* New York: McGraw-Hill Book Co., Inc., 1949.

Heldman, J., *Techniques of Glass Manipulation in Scientific Research.* New York: Prentice-Hall, Inc., 1946.

Kohl, Walter H., *Materials Technology for Electron Tubes.* New York: Reinhold Publishing Corp., 1951.

Martin, L. H. and Hill, R. D., *A Manual of Vacuum Practice.* Melbourne: Melbourne University Press, 1949.

Strong, John, *et al., Procedures in Experimental Physics.* New York: Prentice-Hall Inc., 1938.

9

PARTICLE COUNTERS: GENERAL CONSIDERATIONS

Many different kinds of particle detectors are available to the physicist, and often one of the most important considerations in the design of an experiment is the choice of a particle detector. Most particle detectors depend on the interaction of the electric field of the moving particle with the detector to produce a count. However, neutral particles such as neutrons and neutral mesons must also sometimes be detected, and photons often require other special techniques. In general, neutral particles are detected indirectly by allowing them to produce secondary ionizing particles, which can then be counted in appropriate counters.

CHARGED PARTICLES

Fast charged particles can be detected by counters that depend on ionization, such as ionization chambers, proportional counters, Geiger counters, crystal counters, and electron multiplier tubes, or by scintillation counters, which depend on excitation of a phosphor and subsequent emission of light quanta. The particular device to be chosen depends on the experimental requirements. If the particle has only to be detected, the simplest device to use is the Geiger counter. Any fast charged particle passing through a Geiger counter gives a large pulse independent of the ionization produced by the particle, which is easily detected, being of the order of 50 v. Thus the electronic equipment associated with Geiger counters can often be quite simple.

If information is required about the *ionization* of the particle, an ionization chamber or proportional counter can be used; both of these devices give output pulses that are proportional to the ionization, but the ionization chamber requires more amplification. In each case, however, the associated electronic equipment is much more complicated than that required for Geiger counters. Another device that gives information about energy loss is the scintillation counter, which also requires rather complicated electronic circuits.

Sometimes the choice of a counter is dictated by the *speed* of *response*. The slowest counter is the high-pressure ionization chamber operating on positive ion collection, which may require a second or

more to complete collection of the charge. Ionization chambers operating on electron collection are much faster and can be used in circuits with resolving time of the order of microseconds. Pulses from Geiger counters have fairly fast rise time ($\approx$ 0.1 μsec) but have a long subsequent dead time ($\approx$ 200 μsec). Proportional counters do not have a long dead time and are therefore often used in applications where very high counting rates are expected. In cyclotron experiments, for example, the pulses of fast particles are very short. If a Geiger counter is used it may record only a single count because the dead time is longer than the pulse time. The very fastest counters are scintillation counters, the pulse length being of the order of 10^{-9} sec. Here the electronics rather than the counter is often the limiting factor.

A fast counter that has sometimes been used is the electron multiplier tube, used directly. It is not good for detecting electrons but has found some application in the detection of slow heavy particles. Other types of counters occasionally used include photoconducting crystals such as silver chloride or diamond, and Cerenkov radiation detectors.

NEUTRAL PARTICLES

Photons may be detected in the ultraviolet with a photomultiplier tube if the envelope is transparent at the desired wavelength. Nuclear γ rays are detected by their conversion into electrons by the photoelectric effect, the Compton effect, or pair production, depending upon their energy. The phoelectric effect is useful in the region of the atomic ionization potentials, a few thousand volts. The lead K shell ionization potential is 90 kv. The photoelectric effect is important in high-Z materials up to several hundred kev, but it falls off rapidly with energy. The Compton effect is proportional to the electron density of the material and is important from about 100 kev up to about 3 mev. It also falls off with energy. Pair production has a 1 mev threshold and is important in lead above about 3 mev. It increases rapidly both with energy and Z, and is much less important in the light elements. In the light elements the Compton effect is relatively predominant, but in the heavy ones the photoelectric effect and pair production are relatively more important. Of course, in all elements the energy of the γ ray determines the relative importance of the three. Heitler[1] has a good discussion of this and graphs of the data. In actual counters, high-Z walls are desirable for maximum conversion of the γ rays. The walls should be as thick as possible without getting appreciable self-absorption of the converted electrons. Scintillation counters are also useful in detecting γ rays. At a few mev, we can get 10% efficiency, but a

[1] Heitler, Walter, The Quantum Theory of Radiation. New York: Oxford University Press, 1944.

plateau is hard to obtain. Very high efficiency is obtained in the 100 mev range.

For detecting *slow neutrons* a boron fluoride ionization chamber or preferably a proportional counter is usually used; B^{10} has a very high cross section for the reaction.

$$B^{10}(n,\alpha)Li^7 + 2.9 \text{ mev}$$

Of this energy, 2.5 mev is available as kinetic energy of the fragments. For thermal neutrons, the capture cross section of B^{10} is 3000 barns. That of normal isotopic boron, however, is only 520 barns, since the B^{11} cross section is low.

$$\sigma_{cap} \approx \frac{1}{\sqrt{E}} \approx \frac{1}{v} \quad \text{for } B^{10} \text{ up to 10 ev } (v = \text{velocity of neutron})$$

The counting rate $\approx \rho\sigma v V \approx \rho V$, where V is volume of counter. Thus the boron fluoride chamber measures ρ, the neutron density, rather than ρv, the neutron flux. The fission of U^{235} may also be used to detect thermal neutrons but it has a much smaller absorption cross section than does boron. Fast neutrons may be detected with a recoil proton counter with either a hydrogen-rich atmosphere or hydrogenous walls. Recoil proton detection is useful from about 100 kev up. If we want to preserve the neutron energy, a very thin converter must be used so that the energy loss of the protons by ionization within the converter will be negligible. The maximum efficiency with a thick converter is about 0.1%. Scintillation counters are good for detecting neutrons; they have an efficiency of about 10%. High-energy neutrons may be detected either by proton recoil or by high-energy fission. For bismuth, $\sigma_{fiss} \approx 0.030$ barns at 100 mev.

OTHER DETECTORS

Of course, particles can be detected with other detectors than these counters. Cloud chambers and photographic film are useful, and a nuclear reaction produced by the radiation may also be used to measure it. For example,

$$C^{12}(n,2n)C^{11} \quad \text{and} \quad C^{12}(p,pn)C^{11}$$

are useful reactions for detecting high-energy protons and neutrons, since they have a 20 mev threshold and produce the convenient C^{11} 20.5 min 1 mev beta activity.

EXPERIMENTAL ARRANGEMENTS

Geiger counters and proportional counters are often made of a conducting cylinder with a wire down the axis and filled with a gas such as

argon, plus (in the case of Geiger counters) a small percentage of some organic gas. A typical arrangement is shown in Fig. 1.

As the voltage across the counter is increased at first, all the electrons and positive ions are collected, and the counter is acting as an ionization chamber. Then the field near the wire becomes high enough to produce electron multiplication by collision, so the pulses are larger but still proportional to the initial ionization, giving proportional counter action. Higher voltage and greater electron multiplication produce the region of limited proportionality in which lightly ionizing events receive greater multiplication than the heavily ionizing events. This is the region of limited proportionality. Still higher voltage produces multiplication so great than an electron avalanche will be initiated by even one electron, and all pulses have the same size regardless of the initiating ionization. This is the Geiger region. Increased voltage increases the pulse size until the tube goes into what is called continuous discharge but is really continuously intermittent

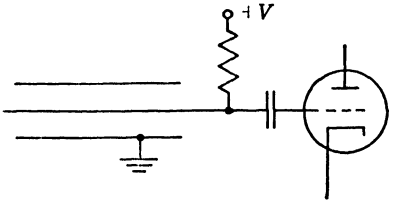


Fig. 1. Counter circuit.

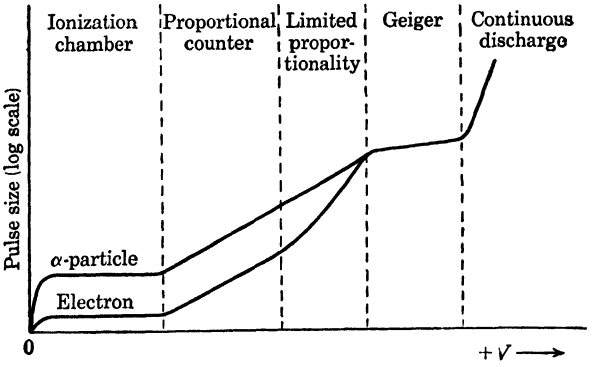


Fig. 2. Counter characteristics.

discharge. A graph of pulse size versus voltage for an α track and an electron track in a counter are shown above and illustrate this behavior.

In Geiger and proportional counters and fast ionization chambers, gases such as oxygen and water vapor that will pick up electrons to form negative ions must be carefully excluded because these counters depend on the great mobility of the unattached electrons for their action. The specific ionization from a track in a counter at atmospheric pressure varies from about 70 ion pairs per centimeter for a minimum

ionization electron to about 50,000 ion pairs per centimeter for a low energy α particle (and even higher for fission fragments). The actual value is proportional to pressure.

In the proportional counter, electron multiplications of up to about 10,000 may be attained; but the greater the multiplication, the greater the loss of proportionality with large pulses because of space charge effects near the wire. In the proportional counter, multiplication occurs only within about one wire diameter of the wire, so only the original ionization within this very small volume receives less than the full proportional multiplication. Since almost all the ion pairs are formed right next to the wire (because of the multiplication), the pulse is actually formed by the work done in pushing the positive ions away from the wire. These pulses are differentiated by an RC circuit to sharpen them. If the multiplication is not too high, further pulses may be recorded before the positive ions are swept away. Too high a multiplication produces sufficient space charge effects that the second pulse receives appreciably less multiplication than the first, thus losing proportionality. Pulses only 1 μsec apart can be counted. These pulses are only a few millivolts high, and do not appreciably change the applied potentials.

COUNTING CORRECTIONS

In Geiger counters, any ionizing event produces a discharge at the wire. Again the work done pushing the positive ions away produces the pulse. Ultraviolet photons propagate the discharge up and down the wire in about 10^{-8} sec. The space charge effect of the slow-moving positive ions produces a dead time of about 200 μsec and a recovery time with smaller pulses of about 400 μsec. This necessitates a coincidence correction to the observed counting rate such that

$$N_{\text{true}} = \frac{N_{\text{obs}}}{1 - N_{\text{obs}}\tau_{\text{dead}}} \approx N_0 + N_0^2\tau_d \quad \text{(actually used)}$$

One of the most important problems in the design of experiments involving counters is the question of the statistics of counting. If the total number of counts is N, the standard error in N for random counts is $\sqrt{N}$. The relative error is then $\sqrt{N}/N$ or $1/\sqrt{N}$. Thus in an experiment where the total number of counts is 100, the relative error or accuracy of the total number is 10%. To get a 1% error the total number of counts must be 10,000. In most counting experiments an attempt is made to reduce other errors to the point where the statistical counting error is the dominant one.

Counters are often used in coincidence to reduce the effect of radiation background. Events are then counted only when all the

counters are tripped. High counter efficiency and fast counters are necessary for this. Corrections must be made for accidental coincidence as follows.

Double coincidence $A_2 = 2n_1n_2\tau$ (n_i are average counting rates)

Triple coincidence $A_3 = 3n_1n_2n_3\tau^2$ (τ is the resolving time)

The generalizations for Mfold coincidence are obvious.

For very low coincidence counting rates, correction must also be made for the occasional cosmic ray showers that will trip all the counters. For good results, short resolving times are needed and individual counting rates should be fairly low. At the cyclotron, "telescopes" of three or more porportional counters are used for counting in a high electron background.

Here the problem is further complicated by the fact that the cyclotron delivers a pulsed beam. Since useful counts are obtained only during the short beam pulse, very fast counters are required.

BIBLIOGRAPHY

Curran, S. C. and Craggs, J. D., *Counting Tubes; Theory and Application.* New York: Academic Press, 1949.

Heitler, Walter, *The Quantum Theory of Radiation.* New York: Oxford University Press, 1944.

Korff, Serge A., *Electron and Nuclear Counters.* New York: D. Van Nostrand Co., Inc., 1946.

Rossi, Bruno, *High-Energy Particles.* New York: Prentice-Hall, Inc., 1952.

Rossi, Bruno and Staub, Hans, *Ionization Chambers and Counters; Experimental Techniques.* New York: McGraw-Hill Book Co., Inc., 1949.

Wilkinson, D. H., *Ionization Chambers and Counters.* London: Cambridge University Press, 1950.

Note: An excellent series of articles on particle detectors by various authors appears in *Helv. Phys. Acta, 23,* Supp. III (1950).

10

IONIZATION CHAMBERS

The ionization chamber is, along with the photographic emulsion, one of the oldest devices for the detection of ionizing radiation. Its use dates from 1896, when Becquerel discovered that the emanation from uranium would discharge an electroscope. Mme. Curie used a parallel-plate instrument in some of her investigations. Like the several products of natural radioactivity that are capable of ionizing gases, cosmic rays (which produce about two ion pairs per second per cubic centimeter in standard air at sea level) were discovered with a form of this device.

The operation of the ionization chamber is based on the ability of certain types of radiation to ionize gases. It is known that a particle loses around 35 ev per ion pair in standard air. Thus a 3.5 mev alpha stopping in the chamber would produce about 10^5 pairs, or a quantity of charge of each sign of around 10^{-14} coulomb. A fast cosmic ray particle, on the other hand, produces only about 300 ions in 10 cm of path.

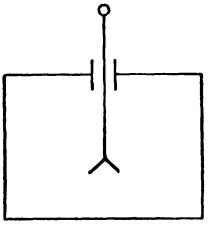


Fig. 1. Gold-leaf electroscope.

The liberated electrons may remain free for some time, attach themselves to neutral molecules, or recombine with the positive ions, although the latter effect is important only at high pressures and for very dense ionization. If the gas under consideration contains electrodes at different potentials, the electrons and other ions will be collected, the total charge acquired being a measure of the intensity of the ionization.

It is important to have an electric field strong enough to collect most of the charged particles before they can recombine. However, if the field strengths used are too large, the phenomenon of gas multiplication will occur. Ionization chambers are thus ionization measuring devices operating in the lowest voltage region. They are simple, rugged, dependable, and reproducible, and easily lend themselves to automatic recording. The earliest devices were simple gold-leaf electroscopes. More modern instruments are usually of the forms shown crudely in Fig. 2.

The choice of insulating material is not critical if guard rings are used, since these reduce the effects of leakage currents, and also electro-

statically shield the collecting electrode. The guard ring is at approximately the same potential as the collecting electrode. If the latter is not near ground potential, the guard ring should be connected to ground through a large capacitor. The potential difference between the two electrodes may be several hundred to several thousand volts.

The size and shape of the chamber is usually determined by the nature and number of the particles to be detected. For many investigations the nature or purity of the gas is not particularly important. For example, air at 1 atm is used in some instruments. But in most applications the composition of the gas is carefully controlled so that collection times may be kept small. Purity is also especially important at high pressures, where recombination must be kept to a minimum. Argon

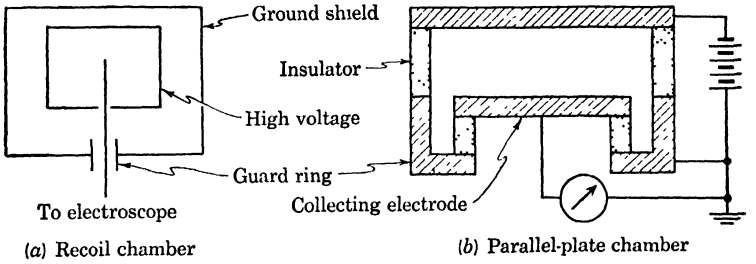


Fig. 2. Ionization chambers.

is especially good in this respect for work at pressures up to the order of 100 atm.

Particles to be counted may be produced in the volume of the chamber, or may be brought in through the walls, perhaps through thin windows. As an example of the latter, beta rays have been introduced into a chamber operating at 3 atm through a bulging 1 mil aluminum window 3 cm in diameter. The principal problem encountered in using aluminum foil for windows is that of pinholes. For small pressure differences, coat them with a few coats of lacquer. At higher pressure differences it is necessary to use thicker foils, perhaps of such substances as nickel-silver.

An important technique for radioactivity measurements is described by Segrè and Wiegand.[1] They measured the *difference* in the activities of two large samples.

As has been mentioned, an electroscope can be used as the detecting and indicating instrument for some purposes, while in more sensitive types an electrometer may be used as the recording instrument. But in most modern uses one records rapidly varying ionization; in this case

[1] Segrè, E. and Wiegand, C. E., "Experiments on the Effect of Atomic Electrons on the Decay Constant of Be^7," *Phys. Rev.*, *75*, 39 (1949).

the chamber is ordinarily connected through a leak resistor to an amplifier.

If the time constant RC is very large compared with the time of collection of positive ions we have what is called an "integrating" chamber. In a typical case, R might be 10^{15} ohms, $C = 10$ $\mu\mu$f, giving a time constant of 10^4 sec. Depending on the geometry and whether or not factors such as recombination are taken into account, the collected current can be more or less easily calculated if the ionization is constant or varies slowly with time.

Usually we are more interested in recording pulses; in this case the capacitance of the chamber and associated circuits becomes of more importance. If the time constant is large compared with the time of collection of positive ions, but small enough that individual events can

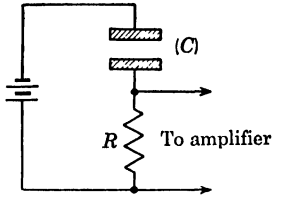


Fig. 3. Ionization chamber circuit.

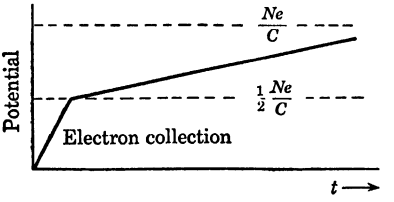


Fig. 4. Collection characteristics.

be recorded, the device is called an "ion pulse" chamber. If the number of ion pairs formed is N, the potential of the collecting electrode changes rapidly at first, then more slowly while the positive ions are being collected, reaching a maximum value Ne/C. It then decreases exponentially. The initial slope of the curve of Fig. 4 is due to electron collection, the rest is due to the slower moving positive ions. Since the time of collection of the positive ions is usually of the order of milliseconds, the time constant RC must be at least 0.01 sec. The advantage of this type of operation is that the maximum voltage is directly proportional to N. The chief disadvantage is that the long time constant requires an amplifier with a low cutoff frequency. This makes the apparatus unsuitable for fast counting and sensitive to microphonics and a-c pickup.

Microphonics are often due to changing geometry of the chamber. If the charge on the collecting electrode is constant and the capacitance varies, a voltage pulse is passed along to the amplifier. In a typical alpha pulse counter, a change in capacitance of the order of 1 part per million will give a pulse of the same order as that of an alpha. One could reduce microphonics by increasing the capacitance and/or decreasing the sweeping potentials, but both are undesirable and it is

best to build the chambers rigidly, to provide holes for the equalization of pressures inside of the chamber, and to mount the chamber on vibration free supports and surround them with sound absorbing material.

If RC is decreased so that the apparatus responds only to the fast part of the pulse corresponding to the collection of electrons, it is called an "electron pulse" chamber. Since the collection time of electrons is of the order of a microsecond, RC can be around 10 to 20 μsec, and the arrangement will allow a fast counting rate. However, in this case the pulse height depends not only on the total number of ion pairs, but also on the position of production of ions. This trouble can be reduced through the use of a shielding grid, to be mentioned later.

The output wave form depends, of course, on the transient response of the amplifier. For example, if the input pulse is of the form of curve (a), the output might look like curve (b) (Fig. 5). (A term frequently used in connection with the transient response of an amplifier is the "rise time," which may conveniently be defined as the time for the output pulse to increase from 10% to 90% of its maximum value.) With careful work, a nearly square pulse as short as 0.2 μsec can be obtained (c).

Fast ionization chambers can be experimentally tested by: (1) studying the shape of voltage pulses produced by single ionizing particles such as alphas, or (2) by investigating the time dependence of the electron current when a source of continuous ionization, such as an x-ray source, is suddenly turned off and on. The latter method is superior. Circuits are given by Rossi and Staub.[2] Absolute calibration can be accomplished through use of an alpha source of uniform and known energy. A more convenient way is to apply known voltage pulses to the collecting electrode. This method requires calculation or measurement of the partial capacitance of the collecting electrode.

The fundamental limit to the smallest amount of ionization detectable in each instrument is set by statistical fluctuations in the background counts. These counts are due to alpha particle contamination from the wall material (1 to 5 counts/cm^2/hour), cosmic radiation, and natural radioactivity of the external surroundings. The alphas from the walls can be nearly eliminated by painting the inside of the chamber with carbon black or with some very pure organic chemical, such as collodion. Grids can also be used to shield the main body of the chamber. Counts from cosmic rays or external sources of radioactivity can be reduced by the use of absorbers. For example, 10 in. of lead has been used as a shield in some cosmic ray work. The final limitation set

[2] Rossi, Bruno and Staub, Hans, *Ionization Chambers and Counters; Experimental Techniques*. New York: McGraw-Hill Book Co., Inc., 1949.

by background can be roughly estimated by the "square root rule": if the background is n counts per minute, the chamber can be used to measure intensities as low as $\sqrt{n}$ counts per minute.

If a chamber is connected to a fast amplifier and irradiated with a strong source, the output exhibits fluctuations (exclusive of background counts) that are considerably larger than those due to tube noise. These are caused by statistical fluctuations in the ionization current.

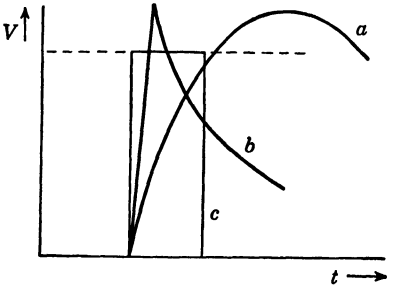


Fig. 5. Waveforms from ionization chamber circuit.

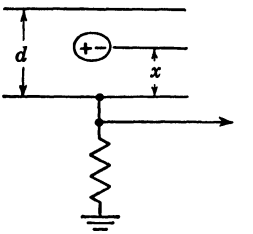


Fig. 6. Separation of charge in ionization chamber.

The rms fluctuation of the output voltage is $\Delta V/V = \sqrt{ne/I\tau}$, where I is the ionization current, n is the average number of ion pairs produced by each particle in the chamber in the time τ, which is the resolving time of the detecting equipment. From such considerations, Rutherford, Schrodinger, and others were able to calculate the number of ions produced by an ionizing particle in its passage through a chamber.

As has been mentioned, older chambers were filled with air, in which an ion moves with a speed of about 1 cm/sec in a field of 1 v/cm. Thus if the chamber were 1 cm deep and a 1000 v potential difference were established, it would take about 0.001 sec to collect the ions, and a fast amplifier would be of no use. At the beginning of this century it was discovered that although electrons readily attach themselves to oxygen and water molecules, they do not become attached to the noble gases, nitrogen, or carbon dioxide, and so may cross a chamber in a time of the order of 1 μsec.

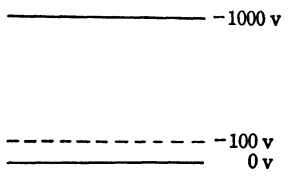


Fig. 7. Grid in ionization chamber.

Assuming the positive ion is fixed, it is easy to show that the output voltage depends on the position in the chamber at which the ion pair was formed. In fact, one can see by inspection that the output pulse supplied to the amplifier is (Fig. 6) $\Delta V = qx/Cd$.

It would be desirable in many applications to have the pulse size independent of the position of ion-pair formation. A device used to this end is the so-called Frisch grid (apparently Frisch had nothing to do with it) which consists of a fine wire screen not far from the collecting electrode. This is shown in Fig. 7.

Since the electrons are effective in producing an output signal only when they are between the bottom plate and the grid, the same change in electrode potential is obtained, regardless of where the electron was

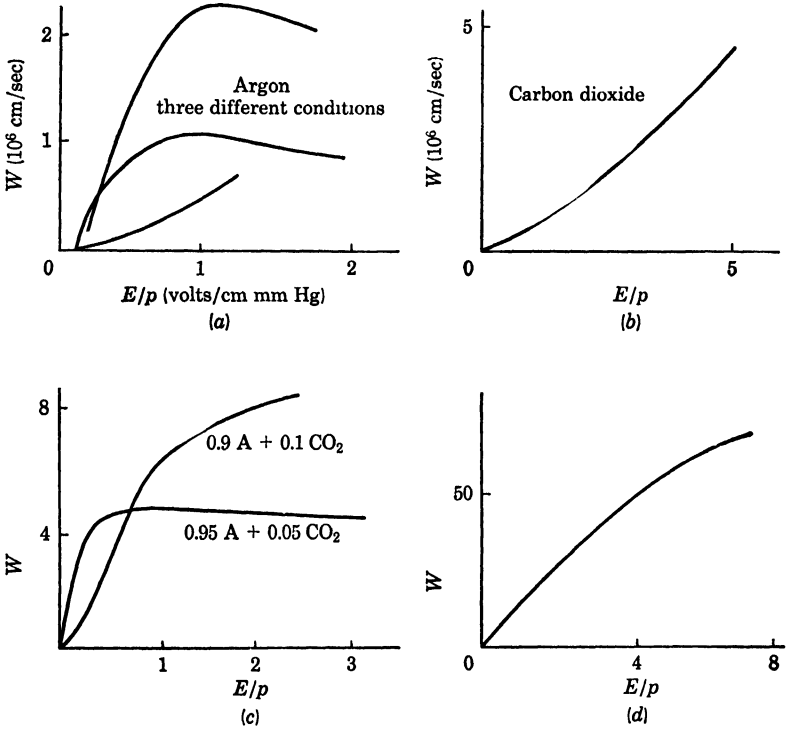


Fig. 8. Characteristics of various gases.

set free. It is thus possible to determine the number of ion pairs formed with considerable precision.

We next investigate the factors affecting the collection time, for a given geometry. The functional form, rather than the exact numerical factor, will be obtained, a more careful average being necessary to find the latter. Assume all the particles have the absolute or agitation velocity v, and designate the drift velocity in the field direction, the mean free path, and the time between collisions by w, λ, and τ, respectively. Between collisions the electron goes a distance $\frac{1}{2}Ee\tau^2/m$ in the field direction. Multiplying by $\nu = 1/\tau$, the number of collisions per

second, and using the relation $\lambda = v\tau$, one gets the drift velocity $w = \frac{1}{2}Ee\lambda/mv$. For pressures in use, the mean free path is inversely proportional to the pressure, so this can be rewritten $w = \frac{1}{2}e\lambda' E/mvp$, where λ' is now the mean free path per unit pressure.

It is seen that the drift velocity is proportional to E/p; this is experimentally verified, but the experimental results do not agree very well quantitatively among themselves, possibly because of varying degrees of purity. The curves of Fig. 8 are from Rossi and Staub.[2]

It is also seen that w is inversely proportional to v. It is difficult to maintain v at a low value in a gas such as argon, because little energy is transferred in inelastic collisions. Moreover, electrons would have to reach high absolute velocities to give inelastic collisions, because the first excitation level is 11.5 ev. However, it is found that 10% carbon dioxide will reduce the electron energy from 10 ev to 1 ev at $E/p = 1$. CO_2 is able to do this because of its many low-energy rotational and vibrational levels. The drift velocity is still limited mainly by collisions with argon.

Finally, the mean free paths of electrons with energies between 10 ev and 1 ev increase rapidly with decreasing energy (in argon) because of the Ramsauer effect. So a few per cent of carbon/dioxide acts in two ways to increase the ratio λ/v and so w. One often finds that the argon used has enough impurities that this result is obtained without addition of carbon dioxide.

The following table[3] summarizes a few of the properties of some of the more useful gases. Data are for electrons in gases at 1 atm with $E/p = 1$. Here ϵ is the average agitation energy in units of $\frac{3}{2}kT$.

Gas	$w(10^6$ cm/sec)	$\epsilon(\frac{3}{2}kT)$	$v(10^6$ cm/sec)	$\lambda/p(10^{-4}$ cm)
H_2	1.2	9.3	35	0 3
He	0 8	53	84	0 5
Ne	1.2	214	168	1 5
A	0 6	287	195	0.9
N_2	0 9	21.5	53	0.3
CO_2	0.55	1.5	14	0.06

ALPHA PARTICLE COUNTERS. Absolute counters are used to determine half-lives or the amount of active material present. The pulse size is unimportant, but the number of particles emitted within an accurately known solid angle per unit time must be determined. Usually a parallel-plate chamber is used, with the material spread on a thick plate; this is called a 2π detector for obvious reasons. It is necessary to make corrections for the finite thickness of the material and back scattering in the plate in this type of instrument, but by depositing the active material on a thin collodion film, it is possible to make a 4π detector that requires no back-scattering correction.

A parallel-plate ionization chamber containing a foil carrying the radioactive material, and filled with argon to about 1 atm, is used for range measurements, the data collected being the number of counts as a function of pressure. Range measurements on samples emitting particles of more than one range can be made with a differential chamber. Particles that go through both parts of the chamber produce effects that are nearly equal and opposite, and can be biased out. Thus the only particles recorded are those that stop in the front half of the chamber.

A pulse analyzing ion chamber (with a quick changing mechanism for the samples) has been built that requires no grid around the inside of the chamber. The collecting electrode is a small sphere, and since

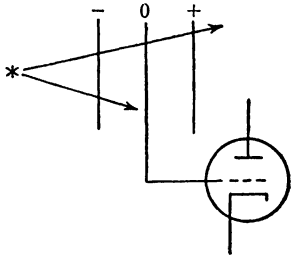


Fig. 9. Balanced ionization chamber.

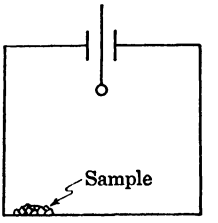


Fig. 10. Chamber geometry used to eliminate sample size effects.

most of the potential fall is in the vicinity of its surface, the pulse height is essentially independent of the position of ion-pair formation.

NEUTRON RECOIL DETECTORS. Neutrons can be detected through the production of ionizing recoils, either in the body of the chamber itself or in some material on one wall of the chamber. Hydrogen is by far the best recoil material, because it has a large scattering cross section, which is well known as a function of energy. Also, the energy of the recoils is larger than from any other element. Neutron recoil chambers are used to:

1. Measure relative fluxes of neutron beams with the same energy distribution. No special precautions are necessary.
2. Make absolute flux measurements of monoenergetic neutron beams. This requires considerable care, as well as knowledge of the geometry and materials of the chamber.
3. Determine the energy distribution of neutrons.
4. Investigate neutron scattering on light atomic nuclei.

In general, it is necessary to keep heavy materials out of the path of the neutron beam. Neutron beams are often accompanied by gamma rays; these can be biased off in a pulse chamber, but it is necessary to compare operation with that of an identical chamber filled with deuterium if the chamber is to be operated as an integrating device.

(n,α) AND (n,p) REACTION DETECTORS. Measurements of neutron beams are often accomplished through (n,α) and (n,p) reactions. For example, a chamber can be filled with boron fluoride, and neutrons of any energy can be measured by slowing them down with sufficient paraffin. It is possible to make the sensitivity nearly independent of the incident neutron energy by using a long counter surrounded by a cylindrical layer of paraffin; the source is placed on the axis. A large number of modified geometries have been used with varying degrees of success. Boron fluoride chambers can be used to make absolute flux measurements if the geometry is accurately known.

BETA, GAMMA, AND X-RAY DETECTORS. These instruments are essentially high-energy electron detectors, the gamma and x rays generating secondary electrons in the walls of the chamber or in the gas itself. Individual electrons have small specific ionization, and so produce only a few ions in the chamber, unless the pressure is very high. For this reason, gas multiplication devices are used to record single particles, and the use of ionization chambers is reserved to the integration of a large number of events. If the number of charged particles passing through the chamber per second is very large, rapid changes can be detected in the intensity of the ionizing radiation, but, as in other fast counters, sealing waxes and organic insulators must not be used in the construction of the chamber if one is to avoid contamination of the gas, with its consequent electron attachment.

Beta ray detectors must be provided with thin windows or the source must be placed inside the chamber, but gamma ray detectors usually have wall thicknesses greater than the range of the secondary electrons.

FISSION DETECTORS. These are used to:

1. Measure the rate of fission and so find the cross section of the reaction, or the flux of a monoenergetic beam of neutrons.
2. Measure the relative neutron flux through the chamber.
3. Investigate the energy distribution of fission fragments.

Chambers for this purpose are similar to those used in the study of alpha spectra.

Because the mean range of fission fragments is only about 2 cm of air, the fissionable material must be in the form of a thin foil or a gas, such as uranium hexafluoride. So far, it has not been possible to

use a gas because of corrosive effects and lack of electron collection. A fission chamber always has a high alpha background, so it is necessary that the chamber and associated circuits have a short resolving time to avoid several alphas piling up to give a fission count. It is also helpful to collimate the beam so that particles do not travel parallel to the electrode, and to make the chamber small enough so that the alphas and fission fragments spend only a small part of their range in the sensitive part of the chamber. Since fission particles ionize most heavily at the beginning of their range, in contrast to alpha particles, the latter are discriminated against. The efficiency is always less than unity because fission fragment energies range from zero up.

In general, it is necessary to apply thickness and wall corrections (see Appendix 10 of Rossi and Staub)[2] if absolute measurements are to be made.

A typical parallel-plate fission chamber for absolute measurements may operate at a pressure of 1.5 atm of argon or nitrogen with a collecting voltage of -300 v at the electrode that carries the foil, and work into an amplifier of 0.1 μsec resolving time (with pure argon the rise time can be of the order of 1 μsec). Such an arrangement will reduce accidentals due to alphas to a negligible number.

Fission chambers of high counting yield, suitable for relative flux measurements have been made by distributing a large amount of fissionable material over multiple alternately connected plates in argon at about 1 atm.

It is possible to make fission detectors of very high counting yield and of dimensions as small as $\frac{1}{4}$ in. in length and diameter by winding two pieces of metallic tape coated on both sides with fissionable material into a compact double spiral, which is then enclosed in a small container filled with argon to 5 to 10 atm. A collecting voltage of 135 v can be used. To avoid counts from several alphas that are nearly coincident in time, it is necessary to use a fairly high counting bias. Efficiencies of 80 to 90% can be obtained with backgrounds no larger than one or two counts per minute. Because of their small over-all size, these chambers can be pushed into otherwise inaccessible places to measure the slow neutron intensity.

The measurement of fission in the pulsed beam of a cyclotron presents special problems because the chamber would respond to the cyclotron pulse. Jungerman and Wright[3] describe a differential chamber which works satisfactorily under these conditions.

COSMIC RAY CHAMBERS. Ionization chambers have been used both as integrating and pulse instruments in the study of cosmic rays, but

[3] Jungerman, J., and Wright, S. C., "Kinetic Energy Release in Fission," *Phys. Rev.*, *76*, 1112 (1949).

in the past amplifier limitations restricted their use to burst measurements or the integration of the effects of many lightly ionizing events. Early integrating instruments were uncompensated, but in modern practice the average ionization current is balanced out and only deviations are recorded.

A well-known instrument of this type is the Model C ionization chamber.[4] It consists of a spherical steel bomb of 20 liter capacity, filled with argon to 50 atm and surrounded with lead shot to an equivalent thickness of 10.7 cm of lead. The spherical shape is for strength and ease of calculation of the average particle path. The anode is a fork in the center, and a uranium source in a cylinder at 250 v provides the compensating current. The spherical shell is at −250 v and the collecting electrode is at 0 v. From knowledge of the average specific ionization of cosmic ray particles, the particle flux can be determined.

With fast chambers it is possible to distinguish between various kinds of events, stars and air showers for example, by examining the shape of the pulse obtained. Measurements of this type are described by Bridge, Hazen, Rossi, and Williams[5] and illustrate the fact that use of ionization chambers is still very much an art. Using argon at 5 atm (for high sensitivity) they found it was impossible to get the pulse size to saturate for any voltage that their chambers could stand, if carbon dioxide was included to cut down collection time. They finally had to settle for pure argon and its considerably longer collection time of 7 μsec.

Their filling procedure seems to be typical: a purifier (an iron vessel filled with calcium metal chips at 300°C) was connected to the chamber and vacuum system for several hours. The chamber was then filled with tank argon through a dry ice trap, and the gas was circulated by convection through the purifier for 12 hr. This method also works for argon-carbon dioxide mixtures.

BIBLIOGRAPHY

Korff, Serge A., *Electron and Nuclear Counters.* New York: D. Van Nostrand Co., Inc., 1946.

Rossi, Bruno, *High-Energy Particles.* New York: Prentice-Hall, Inc., 1952.

Rossi, Bruno and Staub, Hans, *Ionization Chambers and Counters; Experimental Techniques.* New York: McGraw-Hill Book Co., Inc., 1949.

Wilkinson, D. H., *Ionization Chambers and Counters.* London: Cambridge University Press, 1950.

Montgomery, D. J. X., *Cosmic Ray Physics.* Princeton: Princeton University Press, 1949.

[4] Compton, A. H., Wollan, E. O., and Bennett, R. D., "A Precision Recording Cosmic-Ray Meter," *Rev. Sci. Instr., 5,* 415 (1934).

[5] Bridge, H. S., Hazen, W. E., Rossi, B., and Williams, R. W., "A Study of Cosmic-Ray Bursts," *Phys. Rev., 74,* 1083 (1948).

11

PROPORTIONAL COUNTERS

DESCRIPTION. Let an ionizing particle pass through the cylindrical counter shown in Fig. 1. The electrons produced in the ionizing process will be collected on the central wire. As the voltage V applied to the counter is increased from zero, the electrons moving in toward the wire gain more energy between successive collisions with the gas molecules until a point is reached where they pick up sufficient energy between successive collisions to ionize the gas in the counter. Since the field in

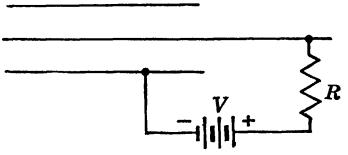


Fig. 1. Counter circuit.

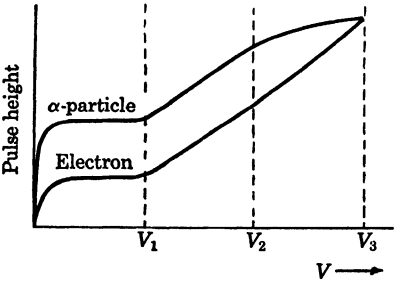


Fig. 2. Proportional counter action.

a cylindrical counter varies as $1/r$, this ionization will take place in the immediate neighborhood of the wire, where the field is highest. The secondary electrons thus produced are also collected on the wire, leaving behind them a sheath of slow-moving positive ions which move out to the counter wall, where they are neutralized.

The motion of the ions and electrons in the field produces a pulse in the voltage applied to the central wire, and if we plot the height of this pulse vs. V, a curve such as one of the two shown in Fig. 2 results. At low voltages the size of the pulse produced by an α particle is much greater than that produced by an electron because more primary ion pairs are produced by the α particle than by the electron. The region below V_1 is the ionization chamber region, that between V_1 and V_3 is the proportional region, and above V_3 is the Geiger region. In the proportional region, the pulse height is proportional to the energy lost by the ionizing primary particle in the sensitive region of the counter, so that one may distinguish between heavily ionizing particles and electrons or cosmic rays.

The gas multiplication M is the ratio of the number of electrons which arrive at the central wire to the number produced by the primary ionizing event (α particle). The value of M is unity when $V < V_1$ and may reach 10^4 or higher at $V = V_3$.

The pulse height depends strongly on M (which in turn depends on the counter voltage among other things), on the rate of energy loss of the primary particle, and on the length of the path of the primary particle in the counter, but does not depend on the location of the path of the primary particle in the counter, since all gas multiplication takes place within a very short distance of the wire. Thus if $M = 10^3$, all gas multiplication takes place in about 10 mean free paths for an electron in the gas ($2^{10} = 1024$). The mean free path for an electron in a gas at 1 atm is of the order of 10^{-4} cm, so the multiplication will occupy only the last 0.01 mm next to the wire. The equation for the pulse height is

$$dV = \frac{n_w e}{C} = M \frac{ne}{C}$$

where n_w = number of electrons collected by the wire, e is the electronic charge, C is the capacitance of the counter, n is the number of electrons produced by the primary ionizing particle, and M is the gas multiplication. Typical pulse height in a proportional counter is of the order of a few millivolts.

Counters of various sizes and shapes will work all right, but for good proportionality, a cylindrical counter should be used so that the field has cylindrical symmetry; otherwise the orientation of the track in the counter may affect the proportionality. The central wire is usually from 0.001 to 0.01 in. in diameter, supported at the ends by insulated feed-through leads. The counter is sealed so that the air and water vapor can be pumped out and the counter refilled with a gas such as argon or a mixture of gases. Gas pressures of from 10 cm Hg to 10 atm have been used. The voltage used will depend on the pressure of the gas in the counter and the exact part of the proportional region that one wishes to work in. It may be anything from a few hundred to a few thousand volts (say 500 to 3000). Good regulation is required since the pulse height depends on the voltage. The capacitance of the system is usually 50 $\mu\mu$f or less, and $R = 10^4$ ohms is a common value.

Speed. Let an ionizing particle pass through the counter in a direction parallel to the wire. The voltage change dV of the wire does not occur at the instant the charge arrives at the wire, but changes continually while the charge is moving toward the wire. If the radius of the wire is a and that of the counter is b, the field at a

point r is $E = V/(r \ln b/a)$. The energy of the system is $E = \frac{1}{2}CV^2$, so that the change in energy caused by an electron moving from a point r to the wire is $\Delta E = CV\Delta V$. We also have

$$\Delta E = \int_a^r F\,dr = \frac{Ve}{\ln b/a} \ln \frac{r}{a}$$

Equating these two expressions and using the formula for the capacitance of a cylindrical capacitor,

$$\Delta V = \frac{2e}{L} \ln \frac{r}{a}$$

However, since most of the electrons in a proportional counter originate in the immediate vicinity of the wire, the size of the pulse will not depend on the distance between the primary ionizing particle and the wire, in contrast to the case of an ionization chamber.

Consider now the voltage on the wire as a function of time. We will assume that R is infinite here so that there is no drain on the system. The voltage changes rather slowly as the electrons move in toward the wire until they reach the point where multiplication sets in, designated by the time t_0 in Fig. 3. At this time, the voltage changes rapidly to the full pulse height, the total time required after the primary ionizing particle enters the chamber being on the order of 1 μsec or less. If the path of the primary particle is not parallel to the wire, the shape of the pulse will be as that shown by the dotted curve. After all the charge has been collected, the voltage on the wire recovers to its initial value as governed by the time constant of the counter and resistor R, which is usually of the order of 1 μsec or less.

The speed of a proportional counter depends on the time necessary for an electron to drift into the region of multiplication and depends upon the same factors as those discussed in the section on ionization chambers. Small percentages of carbon dioxide are often added to increase the speed of the counter.

It should be noted here that the role of the polyatomic molecule is merely to slow down the electrons and not essentially to absorb photons and neutralize monatomic ions as is the case in the self-quenching Geiger counters, since the release of a single electron in a proportional counter does not result in a pulse of full height as it does in a Geiger counter. The photoelectric efficiency of the counter walls is sufficiently small so that few electrons are produced in any case.

The resolving time of a proportional counter (the shortest time interval between two particles that can be detected) essentially depends only on the speed of the counter. In particular there is no dead time due to the space charge effects of the positive ion cloud as in

Geiger counters, because the space charge is so much lower here than in Geiger counters. A resolving time of 1 μsec is easy to obtain in proportional counters and this can be improved by a factor of 10 or more by using small counters of 0.5 cm or less in diameter and using suitable precautions.

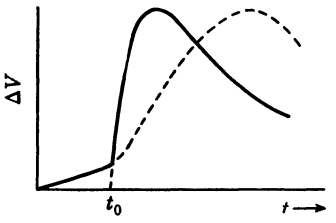


Fig. 3. Pulse shape from proportional counters.

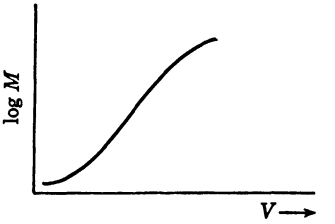


Fig. 4. Gas multiplication vs. voltage.

GAS MULTIPLICATION. The gas multiplication depends on the counter voltage V, its capacitance C, the radius of the wire a, the gas pressure p, and the type of gas. We can write the functional dependence in the form

$$M = f\left(\frac{V}{\ln b/a}, pa, \text{type of gas}\right)$$

If b is increased, the value of M can be held constant by increasing V. If a is decreased, we must change V and at the same time change the product pa in order to hold M constant. The reason is that the path length of field sufficiently strong to cause multiplication is increased when the radius of the wire is decreased.

It is found experimentally that if all factors other than counter voltage are kept constant, the multiplication is given approximately by an expression of the form $M \approx e^{kV}$. This is shown by plotting log M vs. V, as is shown in Fig. 4. It is found that if one uses about a 10 mil wire and a pressure of about 1 atm, M changes by a factor of 2 for each 100 v in the linear region. The linear region extends over an interval of several hundred volts.

Increasing the pressure is found to increase the multiplication. When mixtures are used it is found that the smaller the percentage of the polyatomic component, the more rapid the variation of the multiplication with voltage. A monatomic gas gives higher multiplication for a given V than a mixture containing a polyatomic component. This is presumably because the bombardment of the counter wall by monatomic ions will cause electron emission while polyatomic ions will not; excited polyatomic molecules decompose in a time of the

order of 10^{-13} sec, thus releasing their energy instead of transferring it to an electron in the counter wall.

FACTORS THAT INFLUENCE THE PROPORTIONALITY. In an ideal proportional counter, the pulse height should be the same for all like particles of the same energy. Thus the actual place where the particle goes through the counter should not influence the height of the pulse. Of course, the length of the track in the sensitive volume of the counter will vary as the orientation of the particles' trajectory is changed with respect to the counter, but in many applications the particles to be counted are in a beam or originate from a small source, so that this effect need not be excessive if proper precautions are taken. Geometrical factors can influence the pulse height for different track orientations. For example, a noticeable effect is produced if the wire is off center by 5%. The wire must also be accurately cylindrical or the field will be distorted locally and cause variations in the gas multiplication. Also the end effects where the wire is attached to its supports cause distortion of the field. A 5 mil wire and 50 mil end supports can produce a 5% divergence from proportionality in a counter 10 in. long.

In order to have good proportionality there must be no loss of electrons that have been created in the counter. Gases such as oxygen, water vapor, etc., which lead to recombination, should therefore be avoided. In fact, a contamination of a fraction of a per cent of these gases is undesirable, so precautions are necessary to insure that they are present in negligible amounts. The counter chamber should therefore be vacuum tight, and operation at a pressure slightly above atmospheric will help prevent such contamination should a small leak be present. The best gases to use are the noble gases and nitrogen. Boron fluoride is not so good as these but will give fairly good results.

The total gain of the system is the product of the electronic gain by the gas multiplication. Hence, if the electronic gain is halved and the gas multiplication doubled by raising the counter voltage by 100 v, we might expect the total gain to remain constant and the proportionality of the counter to be unaffected. If we plot the number of pulses vs. pulse height as in Fig. 5 for a counter operating at say 1500 v, we find the half width of the distribution to be of the order of 5%. If now the electronic gain is halved and the counter voltage increased 100 v, this distribution will be broadened, the reason being that space charge effects play a more important role at the higher gas multiplication values, and increased space charge will tend to decrease the pulse height on pulses which follow another pulse closely.

A high background of β or γ rays will also contribute to a lack of proportionality of a counter, the reason again being due to space charge effects.

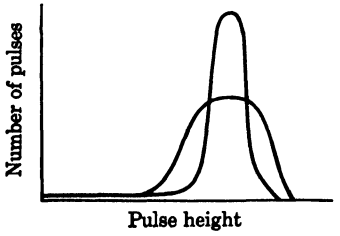


Fig. 5. Pulse height distribution.

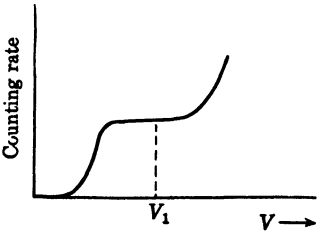


Fig. 6. Plateau of proportional counter.

Counting rate curve. In the actual use of a counter, the question as to the exact value of voltage V to use must be answered. Consider, for example, the case where one wishes to count protons in the presence of β and γ ray background. The background pulses are much smaller than those produced by the protons, so they can be eliminated by the use of a discriminator circuit—one that will pass pulses only if they exceed a certain minimum height. The counter voltage must be adjusted so that all the protons will produce pulses large enough to be counted but none of the β or γ rays will give pulses large enough to count. To do this we plot a curve of counting rate vs. V a so-called plateau curve. This is shown in Fig. 6. The counter voltage should be run at some point near the middle of the flat portion of the curve, say at V_1. This will insure that the height of the pulses that are caused by protons is sufficiently large to be passed by the discriminator circuit while those due to background will not be large enough to get through. The counting rate curve will start to rise again when the voltage is increased and some of the background pulses become large enough to get through.

Coincidence counting. In many applications it is useful to use two (or more) counters in coincidence. That is, a pulse will be registered only if the particle to be counted passes through both counters, thereby energizing them simultaneously. In any coincidence circuit there will be a certain number of "accidental" coincidences, usually small. These are produced by two independent events (presumably due to background) occurring almost simultaneously in the two counters so that the coincidence circuit is operated.

As an illustration, suppose it is desired to count electrons which are at their minimum ionization in the presence of a background of electrons of other energies. The minimum ionization counts are not

necessarily those having the smallest pulse height. There may be some low-energy electrons present in the background. One might at first try using a single counter and plotting a counting rate curve. The voltage at which the smallest pulses start coming in would then be the threshold for the minimum ionization electrons; larger pulses would be considered background. But this technique is not really good since a small pulse is not necessarily due to an electron at minimum ionization. A much better technique is that shown in Fig. 7, where the two counters are operated in coincidence with an absorber between them which is thick enough to stop low-energy electrons but not so thick as to stop the minimum ionization electrons which have energies of about 1 to 2 mev. If a plateau curve is now run, one can be assured that the smallest pulses will be due to electrons that are at minimum ionization.

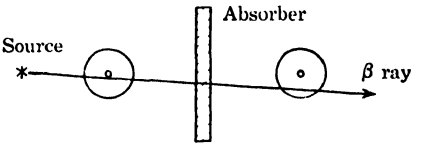


Fig. 7. Typical experiment using proportional counters.

NEUTRON COUNTERS. Slow neutron counters usually employ boron fluoride gas since B^{10} has a large cross section for capture of slow neutrons: $B^{10} + n \rightarrow He^4 + Li^7$. The energy released by the He^4 and Li^7 in the counter is about 2.5 mev, so these pulses are easily distinguishable from β or γ ray background which releases an energy of the order of 100 kev or less in the sensitive volume of the counter. Since B^{10} isotope is only about 20% of natural boron, a fivefold increase of counting efficiency is possible if one can use pure B^{10}. The counting efficiency is the ratio of the number of particles detected to the total number passing through the counter.

For counting fast neutrons, use is made of the fact that neutrons produce knock-on protons when they pass through a hydrogenous substance. The counter may therefore be filled with a hydrogen-rich gas such as pentane (C_5H_{12}) or the counter walls may be lined with a hydrogen-rich material. Sheets of polystyrene, which contains hydrogen in the form $(CH_2)_n$, where n is a fairly large number, may be distributed throughout the volume of the counter. Solid substances should not be thicker than the maximum range of the protons to be expected, for otherwise protons originating from deep inside the solid will never get out into the sensitive region of the counter.

BIBLIOGRAPHY

Corson, D. R., and Wilson, R. R., "Particle and Quantum Detectors," *National Research Council Preliminary Report 7* (January 1950); also *Rev. Sci. Instr., 19,* 207 (1948).

Curran, S. C. and Craggs, J. D., *Counting Tubes; Theory and Application.* New York: Academic Press (1949).

Korff, Serge A., *Electron and Nuclear Counters.* New York: D. Van Nostrand Co., Inc., 1946.

Landau, L., *Jour. Phys. USSR, 111,* 201 (1944).

Lapp, R. E., and Andrews, H. L., *Nuclear Radiation Physics,* 2d ed. New York: Prentice-Hall, Inc., 1954.

Rossi, Bruno and Staub, Hans, *Ionization Chambers and Counters; Experimental Techniques.* New York: McGraw-Hill Book Co., Inc., 1949.

Wilkinson, D. H., *Ionization Chambers and Counters.* London: Cambridge University Press, 1950.

12

GEIGER COUNTERS

Although Geiger counters are now commercially available, many physics laboratories continue to maintain their own facilities for making counters, partly because many of the commercial counters are quite expensive, and partly because special sizes and shapes are often required. We shall first discuss the discharge mechanism briefly so that some of the problems in Geiger counter construction can be understood.

DISCHARGE MECHANISM. The operating part of a Geiger counter consists of an outer conducting cathode surrounding an anode, which usually is a fine tungsten wire. The space between is filled with gas at about 10 cm Hg pressure. When a charged particle passes through the gas it liberates electrons by collision, and these are attracted toward the center wire. As they come close to the high electric field near the center wire (within a few diameters) they begin to pick up enough energy between collisions with gas atoms to ionize the atoms. The electrons produced in these collisions can also participate in the process, and an "avalanche" of charge is built up. Since the total charge collected is many times larger than the charge liberated by the passing fast charged particle, the output pulse size is independent of the initiating pulse, and may be of the order of 50 v. The positive ions move toward the cathode cylinder and form the output pulse by taking energy from the field. Photons are also produced in the discharge and produce photoelectrons in the gas and also, unfortunately, in the cathode.

The important problem in a Geiger counter is to quench the discharge. Since a single free electron will fire the counter, the photoelectrons from the cathode and the secondary emission electrons produced when the positive ions strike the cathode must somehow be suppressed. One way to do this is to remove the voltage from the tube during the period immediately following an avalanche. Very large series "quench" resistances or special circuits have been used to do this. A more satisfactory method, now nearly universally used, is to add a quenching gas to the counter to absorb the photons. Organic gases such as ethyl alcohol are used, whose molecules when excited are more likely to dissociate than to radiate. The quenching gas must

also have a lower ionization potential than the regular noble counter gas. The quenching gas absorbs the photons present in the gas and releases the energy by dissociation. Because of its lower ionization potential, it exchanges electrons with positive noble gas ions in collisions, so that the latter are neutralized. The quenching gas ions when reaching the cathode dissociate upon neutralization rather than producing the secondary emission that helps sustain the discharge. About 10^{10} organic molecules are dissociated to quench each discharge, so a self-quenching counter has a definite life. If heavy organic molecules are used, the dissociated products may also help quench the discharge.

The noble gases, particularly argon, are usually used to fill counters because of their high specific ionizations and low firing voltages. Argon is readily available commercially and is inexpensive. To produce counter speed a high drift velocity of the electrons along the field is desirable, and this is produced by a low average electron velocity, which requires inelastic collisions with the molecules. Since the excitation states of argon are much above the electron kinetic energies, argon will not produce these inelastic collisions. The rotational states of the organic quench gases are low enough to absorb this energy. In proportional counters and ionization chambers, carbon dioxide is added just to do this. For self-quenching Geiger counters a mixture of argon with 10% ethyl alcohol at 10 cm Hg pressure is a good filling. Since the avalanche process depends on the mobility of the free electrons, the counter must not contain gases like oxygen or water vapor, which will pick up the free electrons to form heavy negative ions.

Construction. There are two general types of construction. In the first type, the anode and cathode are mounted inside a glass envelope which serves to contain the gas. In the second type the cathode cylinder itself acts as the container. Since the cathode is very often maintained at ground potential (the center wire at $\approx +1000$ v) there is no serious objection to an exposed cathode, and the trend in modern counter technique is toward the second type. There are, however, some disadvantages. Many counters are made with brass or copper envelopes and it is necessary to use an insulator soldered to the metal case to support the central wire. Frequently Kovar-glass seals are used for this purpose, but it is quite difficult to silver-solder them without destroying the metal or glass. Soft solder is commonly used and often causes leaks.

Metal counters are fairly expensive to make. The cost of the metal and insulators is high, the inside of the cylinder must be polished before assembly, and the assembly process itself requires considerable

skill. The most important factor in the choice of the cathode material is its degree of photosensitivity. Both brass and copper are normally quite photosensitive but with proper treatment either can be used. The procedure is to oxidize the inside of the cylinder after assembly by admitting oxygen to the counter while it is hot. A thin layer of oxide forms and the photosensitivity is much reduced. Brass behaves somewhat better than copper, which tends to make a scaly oxide unless great care is taken. It is important, however, not to heat a brass counter to the point where the zinc starts to come out and get all over the vacuum system.

If it is important to reduce the cost of making Geiger counters, the cheapest counters to make are those with external cathodes. In these the envelope is soft glass, the anode is the usual fine tungsten wire sealed into the glass, and the cathode is a thin layer of graphite in lacquer that is sprayed onto the outside of the glass cylinder. The conductivity of the glass is high enough to allow the charge to leak through. Such tubes have the difficulty that their recovery time is quite long because of the high resistance of the glass and therefore they cannot be used for very high counting rates. They are very useful for cosmic ray work where counting rates are low.

After the counter is assembled it must be cleaned, evacuated, and filled with the proper mixture of gas. Cleaning is very important, and provision should be made if possible to flow cleaning fluids and solvents through the counter. All soldering flux must be removed with a tap water wash followed by distilled water, then a drying agent that leaves no residue. Acetone is satisfactory, but it must be chemically pure. Alcohol is not good. With the external cathode counters the envelope is cleaned with standard glass cleaning solution before the wire is sealed in.

The central wire is often cleaned by heating it to a dull red heat in vacuum by passing a current through the wire. Thus it is desirable to have both ends of the wire come through the insulators and out where connections can be made to them. The wire can be pretreated before assembly by heating it in a hydrogen atmosphere. The smaller the anode diameter used, the lower will be the operating voltage. Three to five mil tungsten wire is ordinarily used; smaller wire is difficult to handle. Tungsten can be sealed directly to the glass but it is more commonly spot welded to a larger piece of tungsten wire or, in the case of external cathode counters, to Dumet wire for sealing to the soft glass.

After the counter is cleaned, it is evacuated and usually baked in an oven to remove surface gases. Baking is not difficult with glass counters, but one must be careful with metal counters not to melt the soft solder. It is important at this point to make sure that there

are no leaks in the counter, because even the smallest leak will eventually affect the performance. The ultimate vacuum required is really not very high, perhaps 10^{-3} mm Hg, but it is usually necessary to get a better vacuum than this to be able to check for leaks. Mercury does not affect the performance of the counter.

After evacuation the counter is filled with a mixture of rare gas and quenching gas. For a self-quenching counter the rare gas must have a higher ionization potential than the organic quenching gas. Argon is a good rare gas to use. The amount of the quenching gas is kept to about 10% because it requires a much higher voltage for Geiger operation than does the rare gas. For example, the rise in Geiger threshold with pressure is 1.5 v/mm Hg for argon and 22 v/mm Hg for ethyl alcohol. There is considerable leeway in the choice of a quenching gas. Ether, ethyl alcohol, amyl acetate, ethylene, and methane, for example, may all be used. The lifetime of the counter will be almost proportional to the molecular weight of the quenching gas, varying for these in a 10% mixture from about 10^9 counts for ether to about 10^7 counts for methane.

If the gases are admitted to the counter separately, one must make sure that equilibrium is reached. The filling system has high inertia. The counter is filled first with the organic gas and equilibrium is reached. Then the argon is admitted slowly. One must be very careful that the organic gas is not backed up out of the counter. Premixing the gases solves this problem. Some think, however, that it is better for the organic gas to interact with the counter surfaces before the rare gas is admitted.

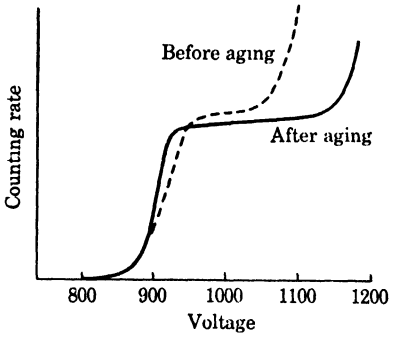


Fig. 1. Effects of aging on Geiger counter plateau.

A new counter must be allowed a break-in period of a few hundred thousand counts, probably because of surface phenomena at the anode and cathode. The counter is first tested with a standard radioactive source. It is then connected to a scaler and the lights turned on and off to check the photosensitivity. The percentage change in counting rate that can be tolerated depends on one's requirements. Brass cathodes show their superiority at this point. Photosensitivity manifests itself in multiple pulsing at high counting rates and at the upper end of the plateau. If the counter is too photosensitive, the only thing to do is to remake it. With the counter counting a standard source the

characteristic Geiger counting curve of counting rate versus voltage is taken. A typical curve, before and after aging, is shown in Fig. 1.

A counter with a very photosensitive cathode probably will not have any acceptable plateau at all. A plateau 200 v wide with a slope of 3% per 100 v is acceptable. The steep rise at the upper end of the plateau is caused by multiple pulsing where complete quenching is not occurring. The operating point V is chosen so that

$$V - V_0 = \text{voltage of pulse} \qquad (V_0 = \text{threshold voltage})$$

If the counter is operated at a higher voltage on the plateau, the phenomenon known as overshoot occurs, in which the pulses are so large that the anode drops below the Geiger threshold. The counter is ordinarily operated at a voltage at about the middle of the plateau.

As the counter ages, the plateau gets shorter and steeper. This changes the operating voltage and may change the absolute counting rate, although relative counting rates are preserved. As the quenching gas is used up, the tube wears out and the plateau gets shorter and steeper and finally disappears. Then the tube must be discarded. High overvoltage uses up the quenching gas too fast and drastically shortens the life of the counter. It is possible to destroy a metal tube by operating for even a short time at very high overvoltage because the counter goes into a continuous discharge, which uses up the quench gas very rapidly. With external cathode counters the danger from overvoltage is greatly diminished since the counter cannot go into continuous discharge because of the high resistance of the glass.

The rejuvenation of a dead counter is sometimes possible and is certainly desirable if the original cost of the counter was high. For consistently satisfactory results the counter must be rebuilt, but sometimes a thorough cleaning of the cathode and glowing of the central wire, then refilling, will make a counter usable again.

Some work has been done on counters using halogens as quench gases. These counters have essentially infinite life but have much larger slopes in the plateau and are still not in general use in experimental physics.

Note on references: Since the material in this chapter is very general and the literature extensive, rather than specific references, a list of general references is given.

BIBLIOGRAPHY

Alder, F., Baldinger, E., Huber, P. and Metzger, F., "Über die Ausbildung der Entladung in Zählrohren mit Alkoholdampfzusatz," *Helv. Phys. Acta*, *XX*, 73 (1947).

Corson, D. R. and Wilson, R. R., "Particle and Quantum Detectors," *Rev. Sci. Instr.*, *19*, 207 (1948).

Curran, S. C. and Craggs, J. D., *Counting Tubes; Theory and Application.* New York: Academic Press, 1949.

Den Hartog, H., Muller, F. A., and Verster, N. F., "Time Lags in Geiger-Müller Counters," *Physica, XIII*, 251 (1947).

Farmer, Earle C. and Brown, Sanborn C., "A Study of the Deterioration of Methane-Filled Geiger-Muller Counters," *Phys. Rev., 74*, 902 (1948).

Friedland, S. S., "On the Life of Self-Quenching Counters," *Phys. Rev., 74*, 898 (1948).

Greisen, Kenneth, "The Intensities of the Hard and Soft Components of Cosmic Rays as Functions of Altitude and Zenith Angle," *Phys. Rev., 61*, 212 (1942).

Greisen, Kenneth and Nereson, Norris, "Inefficiency and other Sources of Error in Cosmic-Ray Measurements with Self-Quenching Counters," *Phys. Rev., 62*, 316 (1942).

Keuffel, J. Warren, "Parallel-Plate Counters," *Rev. Sci. Instr., 20*, 202 (1949).

Korff, Serge A., *Electron and Nuclear Counters.* New York: D. Van Nostrand Co., Inc., 1946.

Korff, Serge A. and Present, R. D., "On the Role of Polyatomic Gases in Fast Counters," *Phys. Rev., 65*, 274 (1944).

Liebson, S. H. and Friedman, H., "Self-Quenching Halogen-Filled Counters," *Rev. Sci. Instr., 19*, 303 (1948).

MacKnight, M. L. and Chasson, R. L., "Construction of the External-Cathode Geiger Counter," *Rev. Sci. Instr., 22*, 700 (1951).

Maze, Roland, "Compteurs à paroi de verre et à cathode externe," *J. phys., VII*, 164 (1946).

Montgomery, C. G. and Montgomery, D. D., "The Discharge Mechanism of Geiger-Müller Counters," *Phys. Rev., 57*, 1030 (1940).

Ramsey, W. E., "Measurements of Discharge Characteristics of Geiger-Muller Counters," *Phys. Rev., 57*, 1022 (1940).

Stever, H. G., "The Discharge Mechanism of Fast G-M Counters from the Deadtime Experiment," *Phys. Rev., 61*, 38 (1942).

Street, J. C. and Woodward, R. H., "Counter Calibration and Cosmic-Ray Intensity," *Phys. Rev., 46*, 1029 (1934).

Wilkinson, D. H., *Ionization Chambers and Counters.* London: Cambridge University Press, 1950.

13

SCINTILLATION COUNTERS

Crystal conduction counters have been intensively studied for some time by many workers. A review article of the field is given by Robert Hofstadter.[1]

Harmut Kallman[2] pointed out in 1947 some of the features of scintillation counters, in particular the short pulse time and short dead time, and that gamma and beta rays could be detected by them. Since then the literature has become extensive.

The scintillation counter is a "solid" counter rather than gaseous. This has several advantages:

1. Smaller size detectors are possible, permitting good geometry.
2. Higher stopping powers, useful for gamma ray work and high-energy particles.
3. More efficient conversion of energy into ion pairs than gas, resulting in better signal-to-noise ratios.
4. Low resolving time, allowing high speed counting and coincidence work.
5. Life of the scintillation counter is indefinite.

PHOTOMULTIPLIERS. In present use the phosphor is viewed by a photomultiplier rather than by the eye. Successive dynodes or field plates are each maintained at a higher potential than the preceding one by means of a voltage divider from a single source. The operation of the tube depends upon secondary emission of electrons and is as follows.

A photon striking the photocathode ejects electrons. These electrons because of the electric field are drawn to the first field plate. There, by secondary emission, other electrons are ejected, the multiplication being as high as 5 or more. These in turn strike the second field plate, where the same thing occurs. A typical tube may contain 10 or more such stages, giving rise to very high amplifications.

It is seen that the photomultiplier is a current amplifier. To get a voltage pulse at the output, the current must be integrated. Assume that the light pulse from the phosphor is a rectangular pulse of dura-

[1] Hofstadter, Robert, "Crystal Counters," *Nucleonics*, April, May, 1949.

[2] Kallmann, H., *Natur und Technik*, July, 1947.

tion T (this is not the case), and that the frequency response of the photomultiplier is such that its output current has the same wave shape as the input (this is a pretty good assumption). If this current is integrated, we may consider (Fig. 1)

$$i_0 = 0 \quad 0 > t > T$$

$$i_0 = i_1 + i_2 \quad 0 < t < T \quad i_0 = \text{constant} = i_0 \quad 0 < t < T$$

$$e_0(t) = \frac{1}{C}\int_0^t i_1 dt = i_2 R \quad 0 \leq t \leq T$$

$$i_1 = C\frac{de_0}{dt} = RC\frac{di_2}{dt} = -RC\frac{di_1}{dt}$$

Thus

$$i_1 = i_0 e^{-t/RC} = i_0 e^{-t/\tau} \quad \tau = RC$$

$$e_0 = Ri_0(1 - e^{-t/RC}) = Ri_0(1 - e^{-t/\tau})$$

The behavior of e_0 vs. time is shown by the curve (Fig. 2). To get 95% of the possible voltage output, τ should not be more than $\frac{1}{3}T$.

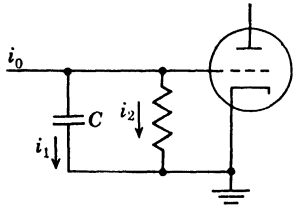


Fig. 1. Amplifier circuit.

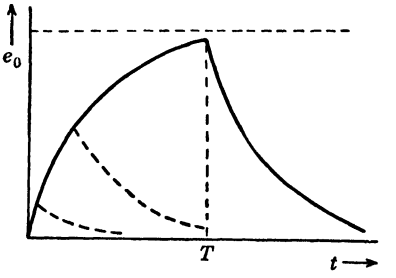


Fig. 2. Pulse shape.

However, good output pulses can still be obtained for larger τ, as shown by the dotted curves drawn for smaller T and fixed τ, which is effectively the same. Ordinarily C is the distributed capacitance of the wiring and the input capacitance of the next tube. A cathode follower will reduce the input capacitance. The lower limit obtainable for C is about 10–15 $\mu\mu$f.

Actually the light pulse from the phosphor has finite rise and decay times[3] but the above is a good qualitative picture.

According to preliminary reports the best tube is the British made VX5032 which gives pulse outputs of about 50 v and has low noise output. Procurement of this tube is difficult however. Photomultiplier tubes readily available in this country are the 931A, 1P21, and RCA 5819. The 1P21 is a selected 931A. The 5819 has the same ele-

[3] Collins, George B., "Decay Times of Scintillations," *Phys. Rev.*, 74, 1543(L) (1948).

ments as the 1P21 but the mounting is designed for scintillation counter work. The relative positions of phosphor and tube for the two types are different (Fig. 3).

The 5819 spectral response is from 3000 A to 6400 A with maximum response at 4800 A. It has a semitransparent cathode $1\frac{1}{2}$ in. in diameter on the face end of the bulb. This large area permits efficient collection of light and allows the crystal to be affixed directly to the tube face. The current amplification is about 400,000 when the tube is operated at 90 v per stage. One drawback is that the tube is extremely sensitive to magnetic fields.

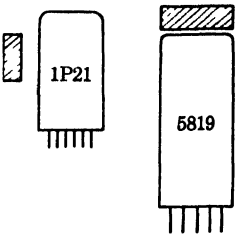


Fig. 3. Arrangement of photomultiplier tubes.

In counting use, the tubes must be protected from light. The crystal also must be shielded to reduce the background counting rate. Even when the crystal background is zero, spurious counts may be recorded due to dark current pulses in the phototube. This dark current background can be reduced very nearly to zero by operating the tube at dry ice or liquid nitrogen temperatures, but this may not be possible for some experiments. British investigators have found noise to be due to electric leakage within the tube, and the VX5032 was designed to minimize this, making refrigeration unnecessary.

Another method to minimize noise is to use two photomultipliers in coincidence, both looking at the same crystal. Noise, being random, will cause coincidences at a much reduced rate. For this the available light is divided between two tubes so that smaller pulse outputs will be obtained from the multipliers.

PHOSPHORS. Many people have investigated materials suitable for counting gamma and beta rays. Opinions differ, however, as to what phosphors are most efficient and as to what the efficiencies are. Anthracene and naphthalene are both widely used. The differences are probably due to slight traces of impurities in the crystals, which may either improve or injure the performance of the crystals. The size of the crystals is important because a conglomerate of larger crystals is more transparent than one of the small crystals. Stilbene and diphenylethyne are more efficient still than the above two materials.

The phosphors used may be divided into two classes, organic and inorganic. Their properties are very different. Inorganic crystals are more efficient than organic, but the pulses from the organic crystals are much shorter. The rise and decay times for some organic crystals may be so fast that there is difficulty in obtaining electronic circuits

fast enough to allow them to be measured. Some special techniques can be used,[3] giving the following decay times.

Naphthalene	5.7×10^{-8} sec
Anthracene	1.3×10^{-8} sec
Phenanthrene	0.9×10^{-8} sec

Such short light pulses make it possible to have coincidence circuits with resolving times at least as short as 5×10^{-8} sec. By using delayed coincidences, half-lives of the order of 2×10^{-8} sec can be measured. Phenanthrene has a slightly higher resolution than naphthalene.

The counting efficiency of some crystals varies with the temperature. In some experiments it might be feasible to improve the counter performance by cooling the crystal.

One more point to be considered in choosing the phosphor is the wavelength of light emitted by it.[4] A crystal should be chosen whose light emission characteristics match the sensitivity characteristics of the photomultiplier used, or vice versa. Anthracene is well suited for use with the 931A or 1P21, for example. Calcium tungstate with peak emission at about 4300 A is well suited for use with 5819.

Since pure crystals are rather expensive, especially in large sizes, liquid phosphors have been developed[5,6] which make it possible to use large collecting volumes. Terphenyl in xylene, for example, while giving pulses only about one-tenth as large as those obtainable from solid crystals, is quite cheap and can be used in large volumes, the problem becoming one of collection of light from a large area.

NEUTRON DETECTION AND PROPORTIONAL COUNTING. This method can be used for detection not only of ionizing radiation but also of neutrons. A counter using a lithium halide has been tested experimentally.[7] Detection efficiency for fast neutrons is 10 to 15% for scintillation counters. Hofstadter proposed that neutrons be absorbed in cadmium, boron, or mercury, and the gamma ray given off be detected by a scintillation counter. This has reasonably high efficiency. According to Wouters no phosphor is needed to detect very high-

[4] Harrison, F. B. and Reynolds, G. T., "Spectral Emission from Scintillation Solutions and Crystals," *Phys. Rev.*, *79*, 732 (1950).

[5] Reynolds, G. T., Harrison, F. B., and Salvini, G., "Liquid Scintillation Counters," *Phys. Rev.*, *78*, 488(L) (1950).

[6] Garwin, R. L., "The Design of Liquid Scintillator Cells, *Rev. Sci. Instr. 23*, 755 (1952).

[7] Farmer, E. C., Moore, H. B., and Goodman, Clark, "Studies on Synthetic LiF as a Scintillation Detector of Heavy Particles and Quanta," *Phys. Rev.*, *76*, 454(A) (1949).

energy neutrons, since upon striking the cathode of a photomultiplier they lose energy, which causes electrons to be boiled off from the cathode surface.

Low energy protons can be detected using zinc sulfide.[8] Calcium tungstate has high counting efficiency and is suitable where high resolution is not needed; it is about 125 times as efficient for radium gamma rays as an ordinary Geiger-Müller tube.[9]

The pulse height is proportional to the energy of the particle detected if the particle is brought to a stop in the crystal. This is easiest with inorganic crystals since their densities are higher. For example, the density of calcium tungstate is 6.06 g/cm^3 and that of sodium iodide is 3.67 g/cm^3, while organic crystals have densities of about 1 g/cm^3. Calcium tungstate, zinc sulfide, and sodium iodide have been tested for energy discrimination, and sodium iodide has been tested for uniformity of response under bombardment by mesons of a single energy.[10] These experiments have established that scintillation counters can be used to discriminate between particles of different energies if the resolution desired is not very great. In this connection it should be remarked that the efficiencies of these counters is to some extent dependent upon the energy of the particles counted.

One problem that this use poses is that the light collection efficiency must be made independent of the position in the crystal where the light flash is formed. The design of the 5819 simplifies this problem.

Light collection. To allow the best use of the light emitted in the flash, the crystal should have a low index of refraction so that the light can easily escape to the photomultiplier. Calcium tungstate has an index of refraction of 1.9, corresponding to a critical angle with respect to air of 31°, which makes it likely that most of the light will be dissipated in multiple reflections within the crystal unless special steps are taken. Canada balsam may be used to cement the crystal to the face of a 5819. This materially reduces the loss of light by reflection since the critical angle is made larger.

The collection can be improved further by surrounding the crystal on all sides except the one in contact with the photomultiplier with an aluminum foil to reflect the light back. Truncated lucite or quartz cones will concentrate the light from a large crystal onto a smaller

[8] Tollestrup, A. V., "Counting Efficiency of Scintillation Counter," *Phys. Rev.*, *74*, 1561(A) (1948).

[9] Moon, Robert J., "Inorganic Crystals for the Detection of High Energy Particles and Quanta," *Phys. Rev.*, *73*, 1210(L) (1948).

[10] Harrison, F. B. and Reynolds, George T., "Measurements of Meson Decay by Scintillation Counters," *Phys. Rev.*, *76*, 169(A) (1949).

photomultiplier cathode surface. Such an arrangement would be necessary if a very high-energy particle were to be stopped, since this requires a large crystal.

If the photomultiplier tube must be some distance from the phosphor (as would be the case if the phosphor had to be placed in a strong magnetic field such as that of the cyclotron), quartz or lucite rods, or metal tubes with polished inner surfaces can be used to conduct the light to the photomultiplier. Quartz is best, a 1 ft rod transmitting 70% of the light from a zinc sulfide screen as compared to 60% for a lucite rod. For the greatest efficiency an optical system of mirrors and lenses can be used.

PREPARATION OF CRYSTALS. A method of preparing thallium-activated sodium iodide is given by Robert Hofstadter.[11] Calcium tungstate crystals can be prepared by the experimenter[12] or obtained commercially from Linde Air Products Co., 30 E. 42 St., New York 17, N.Y. These crystals are available in research quantities with the dimensions $\frac{1}{8}$ up to 2 in. in length. According to the company literature, their density is 7.9 g/cm^3 and peak emission is at 5000 A. The crystals are transparent to their own radiation.

Naphthalene is available in the form of mothballs if a purer source cannot be found. Pure anthracene at $0.10 per gram is available from Reilly Tar and Chemical Corp., Merchants' Bank Bldg., Indianapolis, Ind.

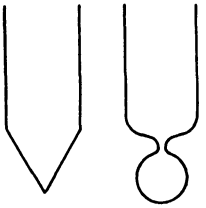


Fig. 4. Shape of containers in which crystals are grown.

The last two materials can be prepared in crystalline form, using what is essentially Bridgman's method.[13] The pure material is placed in a glass or quartz tube of one of the shapes shown in Fig. 4. These shapes allow only one crystal to form in the tube. The material is melted and the melt lowered through a furnace with a high temperature gradient, the highest temperature at the top. As the tube is lowered about an inch a day or slower through the freezing zone, the melt freezes and forms a single crystal under ideal conditions. After the crystal has formed, the temperature is reduced to room temperature over a period of several days. The crystal may be removed from or left in the tube for use. This will depend upon the use to which it is to be put.

[11] Hofstadter, Robert, "Detection of Gamma-Rays with Thallium-activated Sodium Iodide Crystals," *Phys. Rev., 75,* 796 (1949), and *75* 1611(L) (1949).

[12] Zerfoss, S., Johnson, L. R. and Imber, O. "Single Crystal Growth of Scheelite," *Phys. Rev., 75,* 320(L) (1948).

[13] Bridgman, P. W., *Proc. Am. Acad. Arts Sci., 60,* 305 (1925).

AMPLIFIERS AND COINCIDENCE CIRCUITS. No discussion of scintillation counters is complete without reference to the electronic circuits needed to realize the advantages that the scintillation counter offers. However, this is a field in itself. Barring the use of a tube such as the VX5032 as the photomultiplier, or coincidence counting by some such special technique as observing the unamplified photomultiplier pulses displayed directly on a high-sensitivity cathode ray tube, as was done by Dr. Robert Hofstadter in his work at the University of California synchrotron during the summer of 1949, an amplifier will usually be necessary. The distributed amplifier with possible bandwidths of several hundred megacycles is the best type of amplifier for this, since the pulses to be amplified are so short.

For a general review of the entire field of scintillation counters, consult the October 1949 issue of *Nucleonics*.

BIBLIOGRAPHY

Hofstadter, Robert, "Crystal Counters," *Nucleonics*, April, May, 1949.
Jordan, W. H. and Bell, P. R., "Scintillation Counters," *Nucleonics*, Oct. 1949.

14

NUCLEAR EMULSIONS

HISTORICAL SUMMARY. Early work with silver bromide emulsions as recorders of ionizing particles (dating from 1911 when tracks were first observed by Reinganum) was limited mainly to the investigation of alpha tracks in commercial optical emulsions. However, in 1927, moderate energy recoil proton tracks were observed in plates sensitized either by certain dyes or by reduction of grain size. Disadvantages of the first "fine grain" emulsions were: (1) grain spacing was too large to give accurate definition of range; (2) large random development of grains—due to mechanical and thermal activation, pre-exposure cosmic rays, radioactive contamination, and most important, chemical effects occurring in the preparation of the emulsion—gave rise to high backgrounds. These disadvantages have largely been eliminated in modern plates, which have emulsions containing high concentrations of silver halides approximately 10 times ordinary optical plates, in the form of fine grains from 0.1 to 0.6 microns in diameter. Background from preparation and development is generally low.

A further important development from these early plates is the increased thickness of emulsion, as much as 2000 microns (but generally 100 to 600 microns) as compared with 10 to 15 microns in an ordinary optical plate.

Plates in general use in the United States are manufactured by Ilford Co. in England and by Eastman Kodak Co. in America. Both companies manufacture plates that will record, and in part discriminate between tracks of fission fragments, alpha particles, protons, mesons, and more recently, electrons.

COMPARISON WITH CLOUD CHAMBER. It is of interest to compare the nuclear emulsion with the cloud chamber as a means of recording paths of ionizing particles, since the instruments have many common properties. The basic similarity is, of course, the formation of visible (microscopically visible) traces made up of drops (grains) whose density is essentially proportional to the rate of energy loss due to ionization of the incident particle. It is just this last point, the limited proportionality of the response of the emulsion to the ionization loss, that serves as both advantage and disadvantage of the emulsion over the cloud

chamber. To clarify the effects and origin of this limited response, it is desirable that certain aspects of the formation of the latent image by ionizing particles be discussed. This will be taken up briefly below. It may be mentioned now that, whereas in the cloud chamber a drop may form on each primary ion, in the emulsion it probably requires (depending on the sensitivity of the emulsion) as many as 150 ion pairs to produce a single developable grain.[1] Thus the lowest recording rate of energy loss is the rate that will produce a minimum of 150 ion pairs in the distance of a single grain diameter. This threshold energy loss is about 0.013 mev/cm, and corresponding to it for each mass and charge, there is an upper limit for the energy of the particle that will record. It is this differential sensitivity that allows the use of the emulsion to record events that take place in a high background of electrons mainly, which do not record on particular types of plates. Further advantages over the cloud chamber are:

1. *High stopping power.* Present-day emulsions have about 1800 times the stopping power of standard air and not only will many of the low energy tracks stop in the thick emulsions, but also events will be initiated in the emulsion, e.g., nuclear stars and recoils from nonionizing particles.

2. *Operative time.* As contrasted with the cloud chamber, the sensitive time of which is perhaps a thousandth of the working time, the emulsion is, of course, operative continuously, and "integrates" over the time of a cyclotron run or over the time of exposure to cosmic rays at balloon or mountain altitudes.[2]

3. *Mechanical characteristics.* Plates are simple, light, and relatively rugged. As an example of the necessary mechanical strength, it may be noted that on some plates the bonding of the thick emulsion to the glass is such that differential thermal expansion may result in such stresses that the glass chips and breaks away along the edges.

The above facts do not, in every case, comprise unmitigated advantage for the emulsion. In fact, along with the often useful differential sensitivity was until very recently the important limitation of emulsions to consistent recording of only those tracks of particles that have energies up to values corresponding to approximately twice minimum ionization loss. Emulsions of greater sensitivity have now been developed, and continued improvement in this direction can be expected.

[1] Webb, J. H., "Photographic Plates for Use in Nuclear Physics," *Phys. Rev.*, *74*, 511 (1948).

[2] Fading of the latent image provides a limit for acceptable exposure times. Thus if the exposure is as long as several weeks, a new pi meson track may look like an early and thus faded proton. In accelerating machine exposures the time is based on a compromise between the desired frequency and the allowable background for a particular event.

On the other hand, the high stopping power sometimes augments the background tracks in addition to adding to the frequency of the desired event. As to the integrating "operative time," it is just this that makes it difficult to connect events in time, although simultaneity can be deduced in some cases from spatial relations of the tracks. Thus in a pi-mu meson decay and in disintegration of unstable Li^8 into two alphas (so-called "hammer tracks"), as well as in nuclear events (stars), one can reasonably assume that the tracks are in general associated in time. Other disadvantages lie perhaps in the length of time and the strain involved in microscopic examination of large areas of emulsion, the field of view being only about 200 microns and the depth of field about 2 microns with magnification adequate for discerning individual grains (better than 1000 times). Another respect in which the emulsion perhaps suffers in comparison with the cloud chamber is that because of the scattering in the emulsion it cannot record directly the curvature of particles moving in a magnetic field.

PLATES AVAILABLE—THEIR CHARACTERISTICS. It is desirable initially to describe briefly the character of the rate of ionization energy loss versus velocity curve, which has been calculated theoretically and which predominantly governs the rate of grain formation along the track through the emulsion. The curve has the shape indicated in Fig. 1, and is the same for all particles with a single charge independent

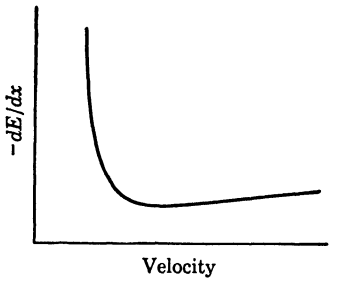


Fig. 1. Energy loss of fast charged particles.

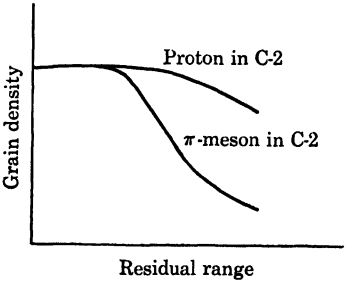


Fig. 2. Characteristics of emulsions.

of the mass value of the particle. The initial rate of energy loss goes inversely as v^2 and approaches a minimum at a velocity corresponding to about twice the rest energy. (Note that if the energy loss is plotted in a somewhat more convenient way against the energy, the curves become mass dependent, dE/dx, increasing with increasing mass.) Above the minimum the energy loss increases very slowly (as the log of v^2) and in this region the particle is spoken of as being "minimum ionizing."

The plates available from Ilford will be described first, (in the order of decreasing sensitivity).

G-5. Sensitive to electrons at high energies and records close to minimum ionizing particles exhibiting, on average plates (on a few plates one may get perhaps double the average sensitivity) a grain density of approximately 20 grains per 100 microns. Random background amounts to approximately one-quarter of this.

C-3. (old B-2) Sensitive to electrons up to about 50–100 kev. Difficulty in observing electron tracks arises from large amounts of scattering which makes the tracks very circuitous. (50 kev electrons correspond to approximately 100 mev protons in ionization.)

C-2. Sensitive to protons up to 40–60 mev and barely sensitive to mu mesons of 4 mev. Rarely, tracks attributable to low-energy electrons may be seen. As mentioned above, this differential sensitivity allows use of this emulsion in high backgrounds of electrons and gammas (thus around synchrotrons) as well as around cyclotrons. This is the most widely used Ilford type.

F-3. Upper limit is approximately 30 mev protons; no tracks from 4 mev mu mesons.

E-1. Very slight difference from F-3; except more fading. Slightly finer grain gives better alpha tracks, distinguishable clearly from protons.

D-1. Finer grain than E-1. Alpha tracks poor but good fission tracks obtained.

The plates available from Eastman are as follows.[1]

NTB-3. Comparable with Ilford G-5, but possibly less consistent. Grain somewhat smaller.

NTB. Comparable to Ilford C-3, but possibly more "clumping" and staining difficulties. Suitable for recording alpha particles to 400 mev, protons to 50 mev, mesons to 5 mev.

NTA. Sensitive to alphas to 200 mev, protons to 20 mev. May be used for fission.

NTC. Sensitive to nuclear fission fragments of high ionizing power only.

Certain qualitative differences between the plates manufactured by the two companies are of interest. The Ilford plates generally show more random grain development and more cosmic ray background. The latter is due to the greater elapsed time between manufacture and use because of the increased transportation time. Shipment by air is not desirable because of the increase in cosmic ray intensity with altitude. On the other hand, the Eastman plates have exhibited somewhat greater distortion (possibly due to uneven layering of the emulsion); occasionally the thicker emulsions start to peel off the backing plate in vacuum and frequently sensitivity has been found to vary by fair amounts over a single plate and by excessive amounts between batches. Perhaps more important than these difficulties is the fact that the Eastman plates (specifically type NTB) do not exhibit the characteristic illustrated in Fig. 2 which is a plot of grain density versus

residual range for protons and pi mesons in Ilford C-2 emulsion. The curves of the figure are normalized to the same value at 0 residual range. It is evident that the relatively large difference in shapes of these characteristic curves affords a clear-cut method of distinguishing the two particles. The difference holds true despite the variations in absolute grain density that may result from differential fading or processing. In the NTB emulsion, however, both have approximately the same variation of grain density with residual range, the variation being like that of the proton in the C-2 emulsion. Improvement in this regard may possibly be achieved by alteration of development technique but this has not been accomplished yet.

Both Eastman and Ilford plates can be obtained with "loadings" of the standard emulsions with the elements lithium, beryllium, and boron. Ilford plates can also be obtained with bismuth as well as, on special order, silver, lead, and other heavy elements.

PROCESSING. In general, development and fixing of thick nuclear emulsions entail difficulties not present in ordinary photographic film work. Despite these, a relatively simple technique is apparently adequate. This procedure is as follows for 100–200 micron Ilford plates in optimum sensitivity uses.

Development. 30–40 min in 6:1 Eastman D-19 at 68°F. Wash in clear water for 10 min.

Fixing. Approximately 1 hour per 50 microns emulsion thickness or twice the time to clear in Kodak Acid Fixer. Wash for 3–4 hours and dry.

Mechanical agitation is used throughout the processing. Variants of the above procedure may be used, and adjustments are made for special purposes. Thus pre-swelling in water for 10 min before development is sometimes used in the hope that it will give superior uniformity of development, and is recommended by Eastman in conjunction with a 10 min full-strength D-19 development. However, better track differentiation is obtained by slow development with weaker solutions of D-19, and in general, with thicker emulsions it is desirable to use longer development times and appropriately weaker solutions. Apparently, if the development time is made sufficiently long compared with the time for the diffusion of the developer through the emulsion, a more uniform development results. For very uniform development, a technique of "cold development" has been originated by the Bristol Group.[3,4] This consists of saturating the emulsion with a 1:1 solution

[3] Dilworth, C. C. *et al.*, "Processing Thick Emulsions for Nuclear Research," *Nature, 162,* 102 (1949).

[4] Dainton, A. D., Gattiker, A. R., and Lock, W. O., "On the Processing of Thick Photographic Emulsions," *Phil. Mag., 42,* 396 (1951).

of D-19 at 5°C for about 30 min, then raising the temperature by diluting with 2 parts water at 15°C and finally warming to 20°C and maintaining this temperature for the remainder of the developing time (about 25 minutes). The processing is then continued as described above. It is sometimes desirable, however, to ensure a complete cessation of the development by using a short stop, a 2% acetic acid bath for about 25 min.

Processing of thick emulsions when distortions must be minimized calls for rather special techniques. The following procedure has been used successfully with 400 micron G5 emulsions. The entire process is done in a refrigerator at 9°C.

1. The plates are placed in a tray containing distilled water at room temperature, and the trays are placed in the refrigerator. Thus the temperature of the plates changes slowly. It is important to avoid thermal shock. The cooling takes about $2\frac{1}{2}$ hours.

2. The plates are transferred to the developer in the refrigerator, and left in the developer for 4 hours. The plates must always be maintained in a horizontal position. The developer is

3 grams amidol
6.7 grams anhydrous sodium sulfite
1.4 ml sodium bisulfite liquor (this is made of a *saturated* solution of $\frac{1}{2}$ sodium sulfite and $\frac{1}{2}$ sodium bisulfite)
930 ml distilled water

Mix in order given. The *p*H should be 6.7.

3. The plates are transferred to a stop bath consisting of 0.5% acetic acid for 1 hour (5% sodium bisulfite may also be used).

4. Plates are moved into the hypo and left for 40 hours or $1\frac{1}{2}$ the time required to clear. They should be kept off the bottom of the hypo tray because silver-laden salts fall to the bottom. Do not overfix. The plates must be agitated in the hypo. The hypo solution is

400 grams hypo
30 grams sodium bisulfite
100 ml distilled water

5. Plates are transferred to the wash water and left for at least 60 hours, with the water changed periodically. The wash water should be tested for hypo before the plates are removed.

6. The plates are soaked in 1% glycerin solution for 1 hour.

7. On removal from the glycerin solution, the edge of the plate is coated with immersion oil, then the temperature is raised slowly (two to three hours) to room temperature. Coating of the edge prevents rapid drying along the edge of the emulsion.

8. The drying is done in an enclosure with hydrated salts, which keep the humidity fairly high and prevent too rapid drying. Drying takes about 3 days.

Surface pressure development as well as unwanted marks and any surface layer of amorphous silver that may form may be removed by rubbing lightly with xylol or alcohol, but care must be taken to avoid removing more than a few microns of the emulsion. Because of the low sensitivity of nuclear emulsions to light, the processing may be carried out under safelight at low intensity using Eastman yellow series 00 or red series 1 filters.

A technique for "sweeping" the undeveloped plates of unwanted tracks without reducing the sensitivity is sometimes very useful. The process apparently depends on oxidation of the emulsion, achieved by exposing it to a high-humidity atmosphere or one containing in addition moderate concentrations of hydrogen peroxide. This may be carried out by placing the plates inside of a fairly airtight and light-tight container on a rack suspended a few inches over the water or hydrogen peroxide surface. The plates should remain over the water, which is kept at 90–95°F for 16 to 24 hours. In the case of the hydrogen peroxide, which is a 3% solution and is also kept at 90–95°F, the plates should remain no more than 1 hour to avoid desensitization. These methods apparently are capable of removing as much as 90% of the background.

Further details of development procedures and sweeping may be obtained from Yagoda, *Radioactive Measurements with Nuclear Emulsions.*

LOADING AND COMPOSITION OF EMULSIONS. To supply target elements for various nuclear reactions, loading of emulsions is frequently resorted to. As mentioned above, loading with specific elements may be obtained from the manufacturer, and these are incorporated into the emulsion in known amounts during manufacture. For elements not so available it is necessary to utilize loading by adsorption (i.e., soaking the emulsion in a solution of the desired ions) or by an evaporation technique which consists of "delivering a measured volume of known concentration to the surface and allowing the solution to evaporate to dryness."[5] Only relatively small amounts of foreign nuclei can be added without danger of losing sensitivity. In the adsorption or soaking technique, the plates are placed in the solution for 15–30 min depending on the desired concentration and thickness of the emulsion. Since the uptake is a complicated function of many variables, determination of the final concentrations of the loaded atoms must be obtained by

[5] Yagoda, H., *Radioactive Measurements with Nuclear Emulsions.* New York: John Wiley & Sons, Inc., 1949.

chemical analysis or by a method of difference involving the initial and final concentration and volume of the solution. For alpha-emitting atoms a calibration may be carried out in terms of the rate at which ions from a known concentration solution are adsorbed in a given emulsion. This rate can be determined by subsequently examining the density of alpha tracks from the atoms adsorbed. It has been shown that the uptake from uranium acetate into C-2 emulsions may be such as to give as much as 1 atom per 300 grains but the plates will be desensitized in a few hours after this uptake.[5] In addition to acetates, nitrates and lactates may be used with satisfactory rate of uptake. Care must be exercised in the choice of salts to avoid those that will desensitize the emulsion.[5]

For calculation of the stopping power of nuclear emulsions as well as to determine what and how many target nuclei are present, it is important to know the composition of the emulsion used.

Table 1 gives the composition in terms of atoms per molecule of silver bromide for Ilford plates (G-5, B-2, C-3, C-2, F-3, but not E-1) and Eastman plates (NTA and NTB, but not NTC).

TABLE 1
COMPOSITION OF EMULSIONS

	Atoms per molecule of silver bromide	
Element	*Ilford**	*Eastman*
Ag	1.0	1 0
I	..	0 027
Br	1.0	0.97
C	1 6	1.62
H	3 2	2.68†
N	0 34	0 50
O	0 87	0.69†

* Small amounts of sulfur and traces of calcium, phosphorus, chromium, silicon, and sodium are also present in Ilford plates.

† Not including moisture content, which at 50% relative humidity is 2.2% of the dry weight.

In terms of weight it is interesting to note that 82.5% and 85.1% of the emulsion is silver bromide on Eastman and Ilford plates, respectively, the remainder being gelatin.

EXAMINATION OF PLATES. Nuclear plates are explored by observation under microscopic magnification, the magnitude of which is determined by several factors. For rapid scanning for large events or counting heavily ionizing tracks moderate magnifications can be used, such as 500×. To measure track lengths, grain spacing, and dip, oil immersion apochromatic objectives giving magnifications better than 90×

are required, together with 15× or greater oculars. Microscopes preferably should be binocular with inclined eyepieces and easily adjustable illumination, and should have micrometer screw feed precision stages as well as vertical micrometer depth adjustment. It should be noted that high-power short focal length objectives have limited working distance. For instance, a 97 power 1.8 mm focal length oil immersion objective has a maximum working distance of 300 microns.

Photographic reproduction of the observed tracks is frequently necessary for demonstration and reexamination purposes. Since tracks rarely stay in a single plane, and in many cases travel over large numbers of fields of view, it is necessary to take many separate photographs (adjusting for proper illumination for each) and piece them together to form a mosaic showing the entire track or event. A convenient arrangement for taking these pictures is described by Morrison and Pickup.[6] A motion picture technique for recording events in the nuclear plates, which has been tried at the University of California Radiation Laboratory, has certain advantages over the mosaics. In the first place, it is much faster than piecing mosaics; secondly, it allows the viewer of the movie to follow the track three-dimensionally, thus bringing out the details of track structure necessarily absent from mosaics. The technique requires a demonstration eyepiece on the microscope, allowing the simultaneous observation of the track by the manipulator of the stage and focus and the recording of the image by the movie camera.

TECHNIQUES AVAILABLE FOR EVALUATION OF PARTICLES. *Grain density.* As mentioned previously, the developed grain density of a nuclear track in an emulsion depends greatly on the space rate of ionization energy loss of the nuclear particle in the emulsion, but it should be realized that the mechanism of formation of the latent image imposes limitations on the exact relation of grain density to ionization loss. This mechanism as described by Webb[1] is briefly as follows. A high-velocity charged particle passing through a particular grain of silver bromide will free by ionization very rapidly many electrons, which will be able to move around freely in the crystal until they come in contact with a so-called sensitivity speck (minute impurity centers, possibly silver sulfide) where they become trapped in a low energy level and exert a localized electrostatic attraction for the mobile silver ions present in the crystal. These ions are neutralized by the electrons and eventually form silver specks large enough to make the entire grain developable. The process is affected by the rate at which the electrons are freed along the straight, narrow path through the grain and which

[6] Morrison, A. and Pickup, E., "Primary and Secondary Meson Event in Photographic Emulsions," *Phys. Rev.*, *74*, 706 (1948).

therefore depends on velocity and ionization of the particle. Very rapid electron liberation may lead to inefficiency in the process because of increased probability for recombination and greater competition of the sensitivity specks for the available latent image silver. Thus, even if there is sufficient ionization in a particular grain, it is sometimes possible that the grain will not develop. In contrast to this, when velocities are low and ionization high, the formation of latently developable grains will take place for all grains along the path, and the grain density will approach the maximum value given by the grain density of the unexposed emulsion, about 2 grains per micron for commercial emulsions.

Within limitations of this type, it is possible to make use of the theoretical average energy loss per unit thickness of stopping material to give some indication of variation in grain density. This energy loss has the form[7]

$$-\frac{dE}{dx} = \frac{4\pi e^4 z^2 NZ}{mv^2}\left[\ln\frac{2mv^2}{I} - \ln\left(1-\frac{v^2}{c^2}\right) - \frac{v^2}{c^2}\right] \tag{1}$$

where ez is the particle charge, v is the particle velocity, m is the mass of the electron, N is the number of atoms per cubic centimeter of stopping material, Z is the atomic number of the stopping material, and I is the average ionization potential of the stopping material.

Thus $dE/dx = z^2 f_1(v)$, and if one writes the grain density as $dN/dx = f(dE/dx)$, it is apparent that particles with the same rate of ionization will have the same grain density. This means that particles having the same velocity and charge will have the same grain density.

Since the range $R = \int_0^E dE/(dE/dx)$ and dE/dx may be written as $z^2\phi(E/M)$, we have

$$R = \frac{M}{z^2}\int_0^{E/m}\frac{d(E/M)}{\phi(E/M)}$$

and thus

$$R = \frac{M}{z^2} f_2(v) \tag{2}$$

From this relation of range and velocity it is evident that for particles having a given v, R is merely proportional to M and $dN/dx = f(v) = F(R/M)$. Therefore, integrating over the range, one obtains for the total number of grains within that range

$$N = M\int_0^{R/M} F\left(\frac{R}{M}\right) d\left(\frac{R}{M}\right)$$

[7] Wheeler, J. A. and Ladenburg, Rudolf, "Mass of the Meson by the Method of Momentum Loss," *Phys. Rev.*, *60*, 754 (1941).

so that for a particle having a given initial velocity

$$N = M\theta\left(\frac{R}{M}\right) \tag{3}$$

and Eq. (2) similarly determines the range, so that both the range and number of grains in that range are directly proportional to the mass. This allows a convenient method of comparing masses of two particles having equal charges from grain counting by finding points along the trajectories of the two particles which have equal densities; then the masses are in the ratio of the residual ranges or the ratio of the total number of residual grains. This method can be applied only for tracks lying in the same region of a single emulsion and which, moreover, end in the emulsion and have been formed at about the same time.

Range-energy relations. From a determination of the range-energy relation experimentally for one type of particle one may use Eq. (2) written in terms of E to determine the relation for a second particle of different z and M. This relation is

$$R_1(E) = \frac{M_1}{M_2}\left(\frac{z_2}{z_1}\right)^2 R_2\left(\frac{M_2}{M_1}E\right)$$

In particular it is convenient to remember that the range of a 35 mev proton is the same as that of a 4 mev meson.

An experimentally valid range-energy relation obtained for protons on Ilford emulsion between the energies 10 to 39 mev is[8]

$$E = 0.251R^{0.58}$$

where E is in mev and R is in microns.
This relation is accurate to about 1% on a single plate and to about 2% for different plates.

Eastman emulsions have approximately $8 \pm 2\%$ larger range for protons at 36 mev. See Webb[1] for details on stopping power of Eastman emulsions.

If the empirical relation above is written in more general form as $E = KR^n$ it may be made to include the charge and mass explicitly by writing $E = K'M^{1-n}R^nz^{2n}$ from Eq. (2).

A sometimes important point which may affect the experimental determination of the range in an emulsion is the large shrinkage (from 10 to 50%) in thickness resulting from the removal during fixation of grains that were not activated. The original track length may be reconstructed by measurement of the shrinkage factor, which can be accomplished by allowing tracks making a known angle with the emulsion

[8] Bradner, H., Smith, F., Barkas, W., and Bishop, A., "Range-Energy Relation for Protons in Nuclear Emulsions," *Phys. Rev.*, 77, 462 (1950).

surface to enter the plate and subsequently determining the path in the developed emulsion.

Scattering. The amount of bending of the tracks due to multiple small angle Coulomb scattering by the nuclei of the emulsion affords a convenient qualitative measure of the masses of similarly charged particles. Thus it is quite easy to distinguish in this way pi mesons of relatively low energy from fast protons: the meson tracks will show a large number of small "wobbles" whereas the protons will exhibit very straight tracks.

Attempts at quantitative utilization of scattering to determine meson masses have been made by the Bristol laboratory,[9] utilizing the multiple scattering theory of Williams and others.

BIBLIOGRAPHY

Webb, J. H., "Photographic Plates for Use in Nuclear Physics," *Phys. Rev.*, *74*, 511 (1948).

Yagoda, H., *Radioactive Measurements with Nuclear Emulsions*. New York: John Wiley & Sons, Inc., 1949.

[9] Menon, M. G. K., Gottstein, K., Mulvey, J. H., O'Ceallaigh, C., Rochat, O., "Observations on the Multiple Scattering of Ionizing Particles in Photographic Emulsions," *Phil. Mag.*, *42*, 708, 932, 1050, 1089, 1232 (1951).

15

CLOUD CHAMBERS

THE PHYSICS OF CLOUD CHAMBER OPERATION. C. T. R. Wilson published his first paper on cloud chambers in 1897, and this was followed by several others on all types of cloud chambers through 1935. The operation of a Wilson cloud chamber depends upon the expansion of a condensable vapor to supersaturation. Supersaturation is defined as the ratio of the density of vapor present to the saturation density at the same temperature. It was shown by Lord Kelvin that, because of the effects of surface tension, the saturation vapor pressure increases if the curvature of the liquid surface is increased. Consequently, drops of very small size will immediately evaporate. Under supersaturated conditions, droplets form on ions, dust, or any other small nuclei that provide a radius of curvature great enough for droplets to grow. Nuclei of greater than a certain critical size for any given supersaturation must be present to enable the formation of drops. For a discussion of the physics of drop formation, see Das Gupta and Ghosh[1] or Wilson.[2] With appropriate operation, the expansion can be arranged so that droplets form only on ions and not on other nuclei (if the dust has been removed sufficiently well from the chamber first). This can be brought about because the effect of a charge on the drop is opposite to the above described effect so that a charged drop will always have lower vapor pressure than an uncharged one. Consequently it is possible to supersaturate to an extent sufficient to form drops on charged nuclei only.

The expansion ratio required depends upon the value of γ, the ratio of the specific heat at constant pressure to that at constant volume, for the gas to be used. The higher γ, the lower is the required expansion ratio. This ratio is usually defined as the ratio of the final volume of the chamber to the initial volume, although sometimes it is expressed in terms of the pressures. The required expansion ratio for a given supersaturation differs for various gases and mixtures of gases. Typical expansion ratios are

[1] Das Gupta, N. N. and Ghosh, S. K., "A Report on the Wilson Cloud Chamber and its Applications in Physics," *Revs. Mod. Phys.*, *18*, 225 (1946).

[2] Wilson, J. G., *The Principles of Cloud Chamber Technique*. London: Cambridge University Press, 1951.

Argon and water	1.2
Argon and ethyl alcohol	1.12
Argon and alcohol 3 parts, water 1 part	1.08
Air and water	1 25

Figure 1[3] shows how the variation of the composition of a vapor may change the expansion ratio for best tracks.

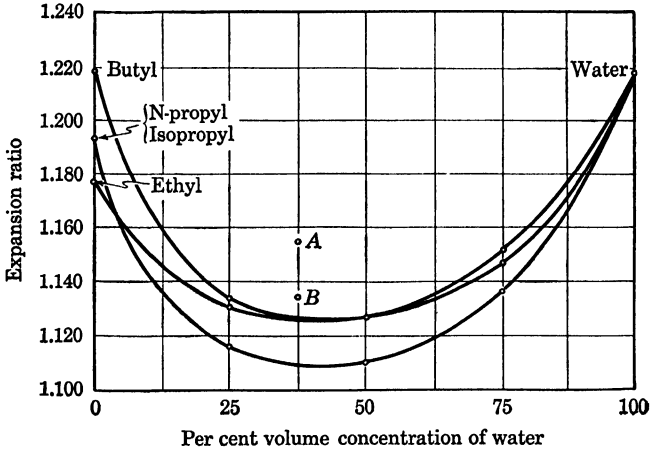


Fig. 1. Expansion ratio as a function of concentration for various liquid mixtures.

A fast expansion is required in order that it be adiabatic. There is always some heat gain by the gas, so the supersaturation S is always somewhat less than

$$S = \frac{D_2}{SD_2} = \frac{D_1}{SD_2(1+r)} = \frac{P_1}{SP_2(1+r)}$$

where D is vapor density, SD is saturated vapor density, r is relative expansion, i.e., $(1 + r)$ is the expansion ratio, and 1 and 2 refer to before and after the expansion.

Operation of cloud chambers. There are two general types of chamber, volume controlled and pressure controlled. C. T. R. Wilson carried out extensive experiments with all types. The volume controlled type expands from a definite initial volume to a definite final volume. This can be done by a moving piston, a Sylphon bellows, or by a rubber diaphragm moving between fixed stops. The pressure controlled chamber expands to a definite final pressure, and this may be accomplished by expanding the chamber with a rubber diaphragm backed by the

[3] Beck, Clifford, "Optimum Liquid Combinations for Cloud Chambers," *Rev. Sci. Instr., 12,* 602 (1941).

desired final pressure. In most chambers the driving force for expansion is derived from a pressure difference applied by suddenly opening a pop valve. It is inadvisable to allow the rubber diaphragm to separate from a metal backing plate as this is likely to cause a separation of charge at the surfaces.

Having decided on the proper range of expansion ratios, one can design the chamber. The procedure in starting to operate a new chamber is the following.

1. Fill the front with gas, and add enough liquid to get a saturated vapor.

2. Set the proper expansion ratio.

3. Expand the chamber. This results in a dense fog which precipitates on dust particles in the chamber. To get rid of these, one makes a series of slow expansions or expansions at a lower expansion ratio. Each time a fog will appear and the weight of the drops will carry the dust particles out of suspension in the air. Enough slow expansions should clear the gas. To shorten this process, one should take particular care to have the chamber as clean as possible before making the initial expansion. *After* no cloud appears on slow expansion, *then* one can make a rapid expansion. A bad overexpansion makes clearing difficult. An abundance of small drops appears, which evaporate into invisibility only to serve as nuclei for another dense fog on the next expansion. A large number of slow expansions is necessary to precipitate this fog. The first rapid expansion will also produce a fog on the background ions that have accumulated in the chamber. These are removed by the clearing field, which should be about 20 v/cm. The back of the chamber is usually grounded, and a grid may be mounted within the chamber to carry the voltage. With the clearing field, one can look for tracks, using an auxiliary source of radioactivity. The clearing field is left on until just before the expansion. A bright light is needed for viewing the tracks; a 100 w bulb with a condensing lens is satisfactory. One should be able to see individual drops, if the chamber is performing satisfactorily.

Helium is sometimes difficult to use because of oil drops in it from the loading mechanism. One should fill a chamber with helium through a liquid air trap or aerosol filter. This is also advisable when filling with air.

An air-water mixture produces tracks over a range of expansion ratios of 1.2 to 1.4. The latter produces a large background. One should adjust for a low background. This requires careful cleaning beforehand.

OPERATION OF CHAMBER WITH PARTICLE ACCELERATOR. The pulse from a machine lasts about 0.001 sec. The procedure used is as follows.

1. Turn on magnet.
2. Expand chamber.
3. Pulse machine.
4. Flash lights.

The only widening of the track is that produced by diffusion during the delay between the passage of the particle and the formation of the drops. This is a very short time compared with the delay present in a counter-controlled chamber; so the tracks can be photographed when very sharp.

Operation of counter-controlled chamber. Coincidence counters are used to trigger off an expansion only when a particle has passed through a chamber to avoid the waste of exposures resulting when random operation is used. A coincidence in counters above and below the chamber releases the pop valve, causing the expansion. A thyratron operating a relay to open the solenoid circuit takes about 0.010 sec since there is a spark upon opening the circuit. An alternate procedure is to discharge a capacitor through a bucking coil to reduce the field to zero. This reduces the delay to about 0.002 sec. Speed is needed to overcome the diffusion of the ions. The width of the track increases with the square root of the time between the formation of the ions and the condensation of the droplets. It is necessary to avoid turbulence in this type of operation since the tracks move during the expansion. Argon is a good filler since it is heavy and is therefore a slow diffuser and less subject to turbulence. A disadvantage is its low thermal conductivity.

Photography of tracks. Illumination may be provided by flash of krypton and xenon discharge tubes, which provide a very intense and short flash. The flash must be short since the drops are falling in the gravitational field and must be stopped, and must be intense because the cameras must be operated at a small aperture to give a good depth of field. The flash tubes are manufactured by General Electric Co. and by Amglo Corp. They are flashed at 2000 v by discharging capacitors through them. The discharge is initiated by a high-voltage pulse on a tickler coil outside the tube. The lights must be flashed after a short delay to allow the drops to grow. The tracks break up in a few tenths of a second. Following the flash, one or two slow expansions (4 or 5 sec) clear the fog, and another exposure may be made. The flash tubes are mounted in front of cylindrical parabolic reflectors or with cylindrical lenses. An increase in intensity may be achieved by mounting the tubes in a block of transparent plastic and using this in conjunction with a second lens. The best lighting uses a transparent chamber with lighting from the back. If the light source is directly behind the chamber, a strip

may be put on the face of the chamber to block the direct ray. The intensity goes up rapidly as the angle between the camera axis and the direction of illumination decreases.[4]

Flash tubes must be mounted securely if they are to be flashed in a magnetic field or the force from the large current pulse may tear them loose or break them. Cameras are usually specially constructed and stereoscopic. The film may be linagraph ortho, or pan. There is some advantage in ortho film with flash tubes because of its high sensitivity in the blue. Super-XX and Superpan Press are also used.

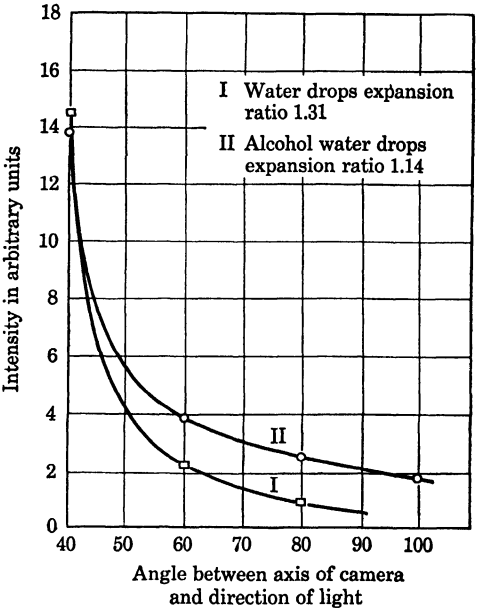


Fig. 2. Brightness of drop in a cloud chamber at different angles with the direction of the illuminating beam.

A convenient arrangement is two cameras permanently mounted in a stereoscopic casting. When the films are developed, they are projected with the same lens system. If the depth of field is good, a movable translucent screen may be used to locate the plane of the track system in space.

SENSITIVE TIME. When a chamber is expanded, the pressure drops and the supersaturation increases through a critical point at which nuclei can form drops that will grow. Following the expansion, heat

[4] Webb, C. G., "The Scattering of Light by Drops in a Wilson Chamber," *Phil. Mag., 19,* 927 (1935).

flows into the gas of the chamber, raising the temperature and decreasing the supersaturation. After a period of time, it decreases through the critical point, and nuclei will not cause droplet formation. The length of time during which nuclei will cause condensation of drops is called the sensitive time; it varies from about two milliseconds in low-pressure chambers to two or three seconds in high-pressure chambers.

Hazen has studied the sensitive time.[5] He has found that the sensitive time can be increased by

1. Increasing the pressure.
2. Slowing the expansion.

However, the length of time taken for the expansion must be of the same order of magnitude as the sensitive time before this helps much, and expansions taking this long (about a half second or so) tend to cause turbulence. The use of Sylphon bellows decreases turbulence and is preferable to use of piston and diaphragm machines in this respect. Bearden,[6] who was the first to realize that an increase in sensitive time would increase the cosmic ray track yield, increased the sensitive time to 2 sec by this method.

3. Increasing the expansion ratio. This is limited by the ratio that will cause general fog.

4. Increasing the ratio of volume to surface area of the chamber. This may result in a chamber that is hard to manipulate.

A long sensitive time is an important consideration if the chamber is to be operated at random in cosmic ray work. The probability of getting a track is directly proportional to the sensitive time.

Langsdorf[7] constructed a chamber of continuous sensitivity. It consisted of a column in a thermal gradient. Saturated vapor is introduced at the top (warm) end and settles through the tube, cooling and becoming supersaturated near the middle, where tracks may be formed continuously. The vapor finally condenses and "rains out" at the cooler, lower end. It was difficult to get sharp tracks with this device. Recently[8] simple continuous cloud chambers have been constructed that use a thermal gradient produced by a block of solid carbon dioxide placed under a saturated volume of gas. With vapor fed in from a pad at the top of the chamber, tracks can be seen to form in the rain that generally appears in the chamber.

[5] Hazen, W. E., "Some Operating Characteristics of the Wilson Cloud Chamber," *Rev. Sci. Instr.*, *13*, 247 (1942).

[6] Bearden, J. A., "Wilson Cloud Chambers with an Increased Time of Sensitivity," *Rev. Sci. Instr.*, *6*, 256 (1935).

[7] Langsdorf, Alexander, "A Continuously Sensitive Diffusion Cloud Chamber," *Rev. Sci. Inst.*, 10, 91 (1939).

[8] Nielsen, C. E., Needels, T. S. and Weddle, O. H., "Diffusion Cloud Chambers," *Rev. Sci. Instr.*, *22*, 673 (1951).

SPECIAL DIFFICULTIES. The curvature of the track in a known magnetic field is often required, and there are various sources of spurious curvature, as for example, thermal circulation of air and expansion turbulence. Thermal air currents may be reduced by placing a 1 w heating resistor on the top of the chamber. Turbulence problems are reduced by putting the chamber on its back rather than on its side. Turbulence can easily cause an apparent radius of curvature of 5–10 m. With considerable trouble this can be increased to about 30 m, but improving it beyond this is very difficult. Other sources of distortion of the tracks are lens distortion and the distortion caused by photographing through the transparent cover of the chamber. These effects can be measured and corrected for by photographing straight wires mounted in the chamber and measuring the resulting distortion in the photographs.

Sometimes it is desirable to place several absorbing plates in a chamber to shorten the range, detect direction, or produce showers. These should be arranged so that they all point toward the camera lens to appear edge-on in the photographs. This results in a set of smaller chambers. A rectangular chamber is most suitable for such an arrangement.

The liquid in a chamber tends to climb up on the front face or deposit on it, obscuring tracks near the edge. Two possible cures for this are (1) blowing a small hair dryer on the face to warm it, (2) running the chamber unsaturated. A difficulty with the latter is that sometimes there is not sufficient vapor present to enable complete clearing. Also, an undersaturated chamber is much more temperature sensitive. A temperature difference can easily cause fogging on the front when the temperature of the gases rises slightly from the numerous compressions.

POISONING OF CLOUD CHAMBERS. Certain substances (e.g., aromatic compounds) chemically poison the chamber. The result is to provide nuclei for droplets which will continuously form fog at very low expansions. This phenomenon needs further investigation. Neoprene and bakelite may cause this although some experimenters report no difficulty with them.

LOW- AND HIGH-PRESSURE CHAMBERS. Ordinarily a cloud chamber is operated at or near atmospheric pressure. For certain applications, chambers operated at considerably different pressures are very useful. The study of low-energy processes of only a few millimeters or less in gas at atmospheric pressure can be carried out in reduced-pressure chambers. The length of track varies inversely with the pressure. On the other hand, higher energy processes involving very penetrating particles that have long ranges cannot be studied effectively in the

ordinary chamber because it will only cover a part of the track. A high-pressure chamber will enable a study of a longer portion of it. Another advantage of a high-pressure chamber is its long sensitive time. This reaches two or three seconds at 100 atm.

LOW-PRESSURE CHAMBER. Joliot describes a chamber designed for operation at various low pressures.[9] Since low pressure cannot produce a large enough force to produce a rapid expansion, an auxiliary piston is used to provide the driving force. This is operated with a separate source of pressure. Joliot found that at low pressures, a very rapid expansion is necessary for satisfactory track formation. Since rubber gasket pressure seals caused too great a frictional force for rapid expansion, Sylphon bellows were used. These gave very satisfactory operation and lasted for years without repair. A disadvantage of low-pressure operation is that the required expansion ratio increases with a decrease in the pressure. It increases from 1.3 at atmosphere to 2 at about 14 mm (saturated water vapor at 17°C). The increased expansion ratio augments the turbulence problem. Since the object of using low pressure is to increase the track length, a gas of very low stopping power should be used. Pure water vapor is the logical choice. The sensitive time is a function of the pressure, decreasing as the pressure decreases. Joliot operated his chamber down to a pressure of about 10 mm. The persistence of tracks at low pressure is very great, around five seconds. Below 20 cm, low-energy electron paths disappear.

HIGH-PRESSURE CHAMBER. The chamber described by Johnson[10] was designed to operate at 200 atm, and possibly extendable to 300 atm. When filled with argon, it has the stopping power of 80 m of atmosphere, 10 cm of water, or 1 cm of lead. Consequently a relatively long portion of the track of a high-energy particle can be observed. The high pressure produces a long sensitive time, and this coupled with the increased track length made visible, makes this chamber about 1000 times as effective as a chamber operated at normal pressure. A disadvantage of the high-pressure chamber is its long clearing time; pictures can be made at a rate of only 1 in 15 min. A high clearing field is needed.

One must be careful in filling a high-pressure chamber, or excessive fogging will result. Careful filtering over a period of several days may

[9] Joliot, F., "Etude des rayons de recul radioactifs par la methode des détentes de Wilson," *J. phys. radium, 5,* 216 (1934).

[10] Johnson, T. H., Benedetti, S., and Shutt, R. P., "A Hydrostatically Supported Cloud Chamber of New Design for Operation at High Pressures," *Rev. Sci. Instr., 14,* 265 (1943).

be necessary. The required expansion ratio decreases with increased pressure and where it is 1.08 at a pressure of 1 atm, it is 1.04 at 110 atm.

Counter control is achieved by mounting 18-in. counters 1 mm in diameter inside the chamber. The wall thickness is 2 mm, and since the ends are rounded, no difficulty is experienced in crushing under the high pressure.

Note on references for cloud chamber technique: the long article by Das Gupta and Ghosh,[1] although now somewhat out of date, contains a great deal of useful information and an exhaustive bibliography. The book by J. G. Wilson is excellent and contains much material not published elsewhere.

BIBLIOGRAPHY

Blumer, H., *Z. Physik*, *39*, 195 (1926).

Das Gupta, N. N. and Ghosh, S. K., "A Report on the Wilson Cloud Chamber and its Applications in Physics," *Revs. Mod. Phys.*, *18*, 225 (1946).

Hazen, W. E., "Some Operating Characteristics of the Wilson Cloud Chamber," *Rev. Sci. Instr.*, *13*, 247 (1942).

Montgomery, D. J. X., *Cosmic Ray Physics*. Princeton: Princeton University Press, 1949.

Rossi, Bruno, *High-Energy Particles*. New York: Prentice-Hall, Inc., 1952.

Wilson, J. G., *The Principles of Cloud Chamber Technique*. London: Cambridge University Press, 1951.

16

PARTICLE ACCELERATORS: GENERAL CONSIDERATIONS

The use of fast particles as probes in atomic and nuclear structure has been largely responsible for the great advances in our knowledge of these fields during the last fifty years. Rutherford's experiments on the scattering of alpha particles and the first transmutations of elements by the use of alpha particles indicated the need for controlled sources of particles. For investigation of the nucleus itself, the probe particles must be energetic enough to penetrate the Coulomb barrier. Another field in which high-energy particles are required is that of x-ray research, where fast electrons, accelerated in electric fields, are shot into targets to produce x rays.

There have been three different methods of approach to the problem of accelerating charged particles: by electric field machines, induction accelerators, and resonance accelerators.

ELECTRIC FIELD. In principle such a machine consists simply of two metal plates with an electric field between them. Particles gain an energy Ve in traversing the field and are said to have acquired V electron volts of energy. The upper limit of voltage is determined by the spark breakdown of such a gap, which limits the voltage to about 400 kv. To overcome this limitation, gaps may be put in series on a voltage divider, as is done in the Van de Graaff machine.

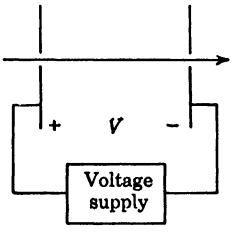


Fig. 1. Elementary acceleration circuit.

A more serious limitation comes in the voltage supply. The most obvious supply is high-voltage transformers, used on 60 cycles. Voltages up to 50 kv are rather easily obtained and rectified, and voltages up to 10^6 v have been obtained in high-voltage laboratories, using cascaded transformers. It is also possible to use radio frequency inductance coils for producing high voltages: Sloan has built a million volt x-ray tube using this principle.

The impulse generator (Fig. 2) consists of a number of high-voltage capacitors that are charged in parallel, then connected in series, producing momentarily a very high voltage. Spark gaps are used to make

the series connection. Transient voltages up to 10^6 v can thus be obtained.

A somewhat similar device is the Cockcroft-Walton generator (Fig. 3). The alternating current applied to one end causes the capacitors to charge in series, in much the same way that a voltage multiplier works at low voltages. Use of radio frequency on a Cockcroft-Walton type

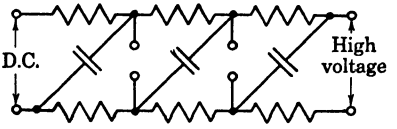


Fig. 2. Impulse generator.

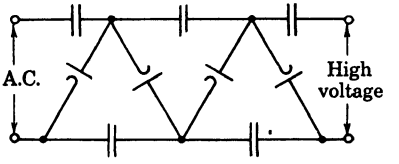


Fig. 3. Cockcroft-Walton generator.

generator has also been proposed, and will perhaps allow attainment of a few million volts.

Electrostatic machines, in which charge is transferred from low potential to high potential by mechanical means, have been useful in the range of a few million volts. A machine designed to give 10–12 million volts is being developed at Los Alamos Scientific Laboratory, but generally experimenters have found that electrostatic machines are quite difficult to operate over 4×10^6 v.

Let us now consider the advantages and disadvantages of these "low-voltage" machines. The only machine presently in use for nuclear studies is the electrostatic machine because it has such excellent voltage regulation, and extremely small ripple. Cockcroft-Walton generators have considerable ripple, although use of higher frequencies may improve the situation. Impulse generators are useless in nuclear studies because of poor voltage control. The difficulties with electrostatic machines will be considered in detail later.

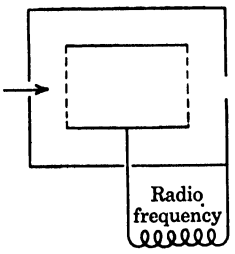


Fig. 4. Linear accelerator.

INDUCTION ACCELERATORS. If a charged particle is moving in a circular path in a magnetic field, and the flux enclosed by the orbit is increasing, the particle will experience a continuous force in such a way as to increase its energy. Machines using this principle were first built by Kerst to accelerate electrons, and have been successfully used in the voltage range of 20 to 300 mev. The upper limit of voltage depends on the radiation loss that the electrons experience in going around the circular paths. When the radiation loss per turn equals the energy gain per turn due to the changing flux, the particle is no longer accelerated.

RESONANCE ACCELERATORS. A third approach to the problem of obtaining high-energy particles has been to give them a series of small accelerations, rather than a single large acceleration. The first such device to be proposed was the linear accelerator. Positive ions admitted through the hole in the cavity (Fig. 4) are accelerated toward the radio frequency electrode when it is negative. They then enter the tube, travel through it free from the influence of electric fields, and emerge from the other end when the electrode is positive. They thus get a second kick and more energy. They must, of course, travel with the proper speed to arrive in the correct phase. Extension of linear accelerators to more gaps came very quickly when Lawrence and Sloan made a linear accelerator, first for mercury ions, and later for lithium. The early linear accelerators were not suited to proton, deuteron, or alpha particle acceleration because the radio frequency techniques at the time were inadequate. The lighter the ion, the faster it goes, and either the drift tubes must be made very long or the radio frequency power must be of very high frequency. It was not until ultrahigh frequency techniques developed during World War II became available that it was possible to build a proton or electron linear accelerator. Such machines, later discussed in detail, have been made for 32 mev protons and one bev electrons. The upper limit of energy on these machines depends primarily on economic considerations.

Lawrence was also responsible for the first circular resonance accelerator, the cyclotron. In the cyclotron the particles bend in a magnetic field so that the same gaps can give successive impulses. The angular frequency of rotation in a magnetic field is independent of the energy of the particle until relativistic effects on the mass become important. The conventional cyclotron is practical for energies up to about 50 mev.

For higher energies the synchrocyclotron may be used. Here the frequency is changed to keep step with the changing mass of the particle, but otherwise the machine resembles a conventional cyclotron. Synchrocyclotrons are used in the range 50 to 500 mev, the upper limit resulting from economic considerations.

A synchrotron can be used to accelerate electrons to high energy. In the synchrotron, described in some detail in Chapter 20, the frequency of the accelerating electric field can be kept constant, since the high-energy electrons, traveling at nearly the velocity of light, always take the same length of time to traverse the circular orbit at constant radius. The magnetic field is changed, and the electrons absorb energy from the electric field to remain in equilibrium orbits as the magnetic field increases. Attainment of energies as high as 1 bev is practical with the synchrotron.

To accelerate protons to billion volt energies, a machine variously

called a bevatron, cosmotron, or proton-synchrotron must be used. In these machines, the magnetic field and the frequency both vary, the radius of the orbit remaining approximately constant. The magnet field structure is made in the form of a large ($\approx$ 50 ft) doughnut to keep the cost of the iron down.

Focusing. In accelerators where the particles must travel great distances during the acceleration process, it is of prime importance to provide focusing for the particles. If the particle, as is inevitable, strays from the prescribed path in position or time, forces must be provided that bring it back where it should be. Let us first consider focusing due to magnetic fields. In a magnetic field such as shown in Fig. 5 the particles tend to circulate in the median plant; if they go off center they are forced back to the median plane by the action of the curving magnetic field. This focusing action is important in the cyclotron and other magnetic field devices.

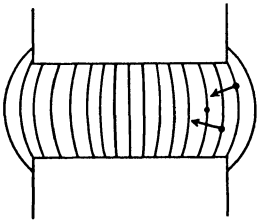


Fig. 5. Focusing in diverging magnetic field.

Electric field focusing must be provided when the particle crosses the accelerating gap. First-order focusing can be obtained by the use of foils across the accelerating tubes and a configuration similar to Fig. 6. Here the particle is always forced toward the center. If the energy of the particles is low, however, foils may introduce scattering and destroy the focusing effect; then second-order focusing (Fig. 7) may be used. The particle, on crossing the gap and being accelerated, spends less time in the diverging field than in the converging field, and thus tends to move toward the center. This type of focusing is used in

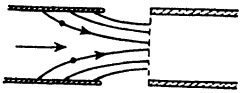


Fig. 6. Electrostatic focusing.

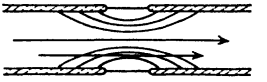


Fig. 7. Second order focusing.

cyclotrons but cannot be used in the proton linear accelerator because of phase focusing requirements.

In addition to spatial focusing, phase focusing must exist in resonance machines. Provision must be made so that if a particle gets slightly out of step with the accelerating radio frequency voltage, the phase of the voltage must be such as to bring the particle back into step. This requirement is discussed in more detail in the chapters on specific accelerators.

Economically, two interdependent factors predominate in the design of accelerators to produce particles of extremely high energy. They are (1) the cost of materials and construction, and (2) spatial focusing. If the beam can be focused into a very small space, the amount of material required to produce a magnetic field in such a space can be correspondingly small, and the cost kept to a minimum. The "strong-focusing principle" recently worked out by Courant, Livingston, and Snyder,[1,2] may provide a method for focusing a beam of high-energy particles into a small area, and thereby make possible an accelerator of 100 bev particles. Such a machine would have a radius of 1030 ft.

BIBLIOGRAPHY

Fremlin, J. H. and Gooden, J. S., "Cyclic Accelerators," *Reports on Progress in Physics*. London: The Physical Society, 1950.

Livingston, M. S., "Particle Accelerators," *Advances in Electronics*, New York: Academic Press, Inc., 1948, Vol. 1.

McMillan, E. M., "High Energy Accelerators," *Helv. Phys. Acta*, *23*, Supp. III, 1950.

[1] Courant, E. D., Livingston, M. S., and Snyder, H. S., *Phys. Rev*, *88*, 1190, 1952.

[2] Livingston, M. S., "The Strong Focussing Synchrotron," *Nucleonics*, Jan. 1953, p. 12.

17

VAN de GRAAFF ELECTROSTATIC ACCELERATORS

Belt-driven electrostatic generators operate (and sometimes do not operate) because of the existence of corona phenomena. To obtain a picture of why corona takes place, recall that the field E at the surface of a conducting sphere of radius r immersed in an infinite insulating medium of dielectric constant ϵ is given by $E = e/\epsilon r^2$, where e is the charge on the sphere. Its potential ϕ is given by $\phi = e/\epsilon r$. Hence $E = \phi/r$. The field E may be increased by increasing ϕ and/or decreasing r. Furthermore, most of the potential difference ϕ is concentrated in the vicinity of the sphere. These relationships are approximately true at the end of a conducting rod considered to have a hemispherical point. When the field E at the point of the rod becomes so high that the surrounding gas is ionized, a corona discharge occurs which is essentially independent of the distance of the point of the rod from the oppositely charged terminal. The use of these corona points will be discussed later in connection with charging belts and with their use as variable "resistors."

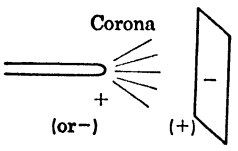


Fig. 1. Corona point.

An electrostatic generator fundamentally is composed of a high-voltage terminal, its insulating support, and a means of conveying electric charge to the terminal. The potential gradient at the surface of the conducting terminal must be minimum with respect to its potential above ground, i.e., it must be smoothly rounded with no sharp points or discontinuities. In the early unpressurized machines this requirement was most simply and compactly met by a hollow metal sphere in space supported by a vertical insulating column (e.g., textolite tube or glass rod). The University of California Radiation Laboratory machine uses a cylindrical electrode with rounded ends to conform to the shape of the enclosing pressurized tank.

Electric charge is transported by a long endless motor driven belt of some suitable insulating material (silk, canvas, paper, rubber, etc.). Charge (negative or positive) is "sprayed" onto the belt at ground potential and is transported into the field-free region within the high-voltage terminal. Here it is removed and conducted to the terminal's

surface. As the high-voltage terminal accumulates charge, its potential increases to a limit determined by the insulating properties of the medium (atmospheric air, for unpressurized generators) surrounding it, by the field gradient over the surface of the high-voltage terminal as determined by its geometry and surface smoothness, and by the leakage of charge along the insulating support.

Charge is sprayed on and removed from the belt by the corona from a row of sharp points (phonograph or sewing needles, fine wires, etc., not unlike a comb) directed toward the belt surface.[1,2] This charging process may run self-excited, in which case the friction of belt and pulley generates static electricity and initiates the action, or an external potential may be impressed upon the "combs." The latter is much to be preferred as it permits more precise control of the charging rate and thus of the potential of the generator.

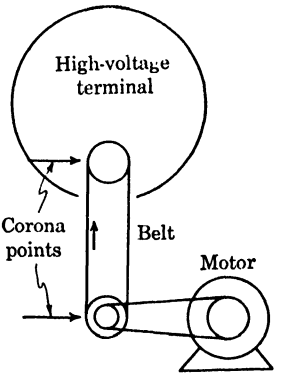


Fig. 2. Elements of Van de Graaff generator.

The amount of charge that can be transported by a belt is limited by three factors; the speed of the belt, its width, and the charge density that it can support on its surface. The first two are mechanical considerations. The third depends on the potential gradient existing at the belt surface and on the dielectric breakdown point of the insulating medium in which the belt is immersed. Several belts, of course, may be (and have been)[3] operated in parallel. If one assumes the belt to be a portion of an infinite plane in a medium of dielectric constant ϵ, the surface charge density is $\sigma = \epsilon E/4\pi$, where E is the potential gradient at the surface of the belt. For air at atmospheric pressure the limiting (breakdown) potential gradient is about 30 kv/cm. The theoretical maximum charge that a belt will carry is then 2.65×10^{-9} coulomb/cm^2 per side, or double this figure for both sides. Van Atta and his associates[1] found they could transport about 60% of this figure.

The charging system of the Univ. of Calif. Radiation Lab machine consists of a 4-ply cotton belt 20 in. wide and 16 ft between pulley cen-

[1] Van Atta, L. C., Northrup, D. L., Van Atta, C. M., and Van de Graaff, R. J., "The Design, Operation, and Performance of the Round Hill Electrostatic Generator," *Phys. Rev.*, *49*, 761 (1936).

[2] Tuve, M. A., Hafstad, L. R., and Dahl, O., "High Voltage Technique for Nuclear Physics Studies," *Phys. Rev.*, *48*, 315 (1935).

[3] Getting, I. A., Fisk, J. B., and Vogt, H. G., "Some Features of an Electrostatic Generator and Ion Source for High Voltage Research," *Phys. Rev.*, *56*, 1098 (1939).

ters, traveling horizontally at 90 fps. The high-voltage terminal is 6 ft long and 56 in. in diameter. Sewing needles placed about $\frac{1}{2}$ in. from the belt and spaced $\frac{1}{2}$ in. apart are used as corona points to spray positive charge onto the grounded end of the belt. The charging voltage is 35 to 50 kv. A similar set of needles inside the high-voltage terminal removes the charge. The corona gap inside the terminal shell acts as a variable resistance having a constant current characteristic. The voltage drop is primarily a function of the gap spacing. Because of this voltage drop the belt pulley within the shell is positive with respect to the shell and draws negative charge (electrons) from the second set of corona points onto the belt. This doubles the belt's charge-carrying capacity. The belt will carry about 0.4 ma in one direction and double this amount in

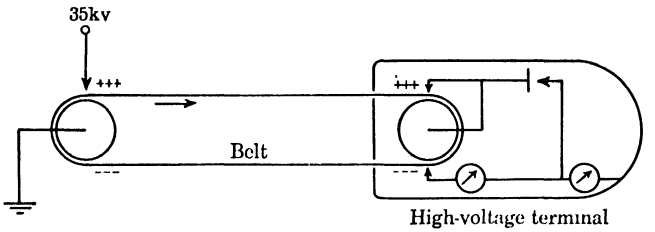


Fig. 3. Modified Van de Graaff generator.

both directions. The energy stored in the electric field surrounding the high-voltage terminal is a result of the work done by the driving motor in transporting positive charge up to the shell and negative charge back down again. About 2 hp is required to deliver 0.4 ma at 4 mev.

Surface charge on the belt may be very high, and will accumulate at points of good insulation if the belt is nonuniform in this respect. Such an uneven charge distribution will give rise to places of high field gradient and result in sparks as effective as a razor in splitting the belt. The life of a belt on the Univ. of Calif. Radiation Lab generator is around 1000 hours.

INSULATION. Unpressurized generators depend upon atmospheric air for insulation. This relatively poor insulation requires huge (hence, costly) machines if high potentials are to be attained. For example, the largest unpressurized electrostatic generator built was the one constructed at Round Hill[1,4] by M.I.T. It consisted of two units, each having an aluminum sphere 15 ft in diameter, supported by a Textolite insulating column 6 ft in diameter and over 22 ft high. A potential of 5.1 mev was claimed for this machine. In addition to their large size,

[4] Van de Graaff, R. J., Compton, K. T., and Van Atta, L. C., "The Electrostatic Production of High Voltage for Nuclear Investigations," *Phys. Rev.*, *43*, 149 (1933).

unpressurized machines are subject to unsteady operation occasioned by vagaries in atmospheric conditions, particularly changes in humidity.

The first attempt to circumvent these objections was to place a generator in an evacuated steel tank.[5] This is not a satisfactory solution. Outgassing the tank requires a very large vacuum pumping capacity. Performance is likely to be erratic and unsteady because of the sudden release of trapped gas during operation.

High-pressure air used as an insulating medium was first reported in 1932.[6] The greatest contributions to the development of pressure insulated generators were made at Wisconsin.[5,6,7,8] Pressurization is accomplished by enclosing the whole machine in a steel tank (usually long and cylindrical) oriented in either a horizontal or a vertical position. The horizontal position complicates the problem of supporting the high potential electrode (it necessitates a cantilever support construction), but requires less room ceiling space and provides a horizontal nuclear beam, which is usually more convenient to use.

The gas used in the Univ. of Calif. Radiation Lab machine is nitrogen at 100 to 110 psi. Air, with its 20% oxygen content, is not used because of the great fire hazard. Five % by weight of Freon 12 (CCl_2F_2, a refrigerant) is added to the nitrogen to increase its electric breakdown strength. It does this by virtue of its ability to pick up free electrons, thus decreasing the ion mobility. Besides Freon 12, carbon tetrachloride was also tried at Wisconsin. It was effective in increasing the insulation of the air but impaired the insulation properties of the Textolite supports used in the tank.

Insulation breakdown is caused by the maximum existing potential gradient and not by the average. A uniform gradient is a necessity in order that all parts of the system be uniformly and efficiently stressed to the maximum permissible value. Ways of equalizing gradients will be brought out in what follows.

The support insulation for the main electrode consists of 6 Textolite tubes 5 in. in diameter of $\frac{3}{8}$ in. wall in a "hexapod" or six-sided pyramidal structure. It carries a load of 2 tons on a 16 ft cantilever from a base plate fastened to the low voltage end of the tank. Its deflection is 1 in. per 1200 lb weight. The apex of the hexapod extends into the main electrode shell. Textolite (General Electric Co. compound 1824) is built up

[5] Herb, R. G., Parkinson, D. B., and Kerst, D. W., "A Van de Graaff Electrostatic Generator Operating Under High Air Pressure," *Rev. Sci. Instr.*, *6*, 261 (1935).

[6] Barton, H. A., Mueller, D. W., and Van Atta, L. C., "A Compact High Potential Electrostatic Generator," *Phys. Rev.*, *42*, 901(A) (1932).

[7] Parkinson, D. B., Herb, R. G., Gernst, E. J., and McKibben, J. L., "Electrostatic Generator Operating Under High Air Pressure," *Phys. Rev.*, *53*, 642 (1938).

[8] Herb, R. G., Turner, C. M., Hudson, C. M., and Warren, R. E., "Electrostatic Generator with Concentric Electrodes," *Phys. Rev.*, *58*, 579 (1940).

of layers of kraft paper bonded under pressure with molten shellac. It withstands a mechanical stress of 4000–5000 psi and is an excellent electric insulator. Because 10 ft lengths only were obtainable it was necessary to make a splice halfway up each leg. Each spliced section is electrically short-circuited over a distance of 1 ft to prevent sparks from being initiated there. The space between base plate and high-voltage terminal is divided into 4 quarters by 3 equally spaced sheet metal diaphragms serving as ground planes. The charging belt passes through rounded slots in the diaphragms so that it is effectively divided electrically into 4 equal short lengths. This reduces the likelihood of failure by insuring a more uniform potential distribution.

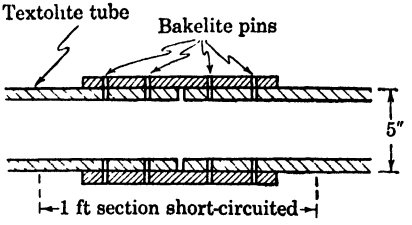


Fig. 4. Splice in support leg.

The field within the hexapod is isolated from the disturbing proximity of the tank walls by a graduated series of large hoops. They are of $\frac{3}{4}$ in. dural tubing and are spaced $1\frac{1}{2}$ in. apart to form a conelike envelope. A uniform gradient is obtained by negative point-plate corona gaps between successive hoops. The voltage characteristic of such corona gaps is shown in Fig. 6. These corona gaps perform the same function as resistive voltage dividers and are much cheaper.

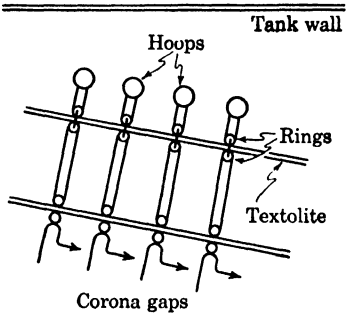


Fig. 5. Arrangement of corona gap.

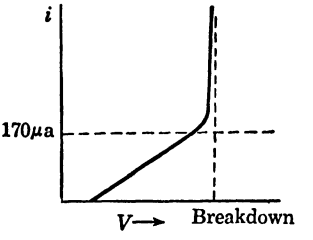


Fig. 6. Voltage characteristic of corona gap.

There is nothing sacred about their geometry and they are much better than resistors in that flashovers of as much as 1000 amperes do not harm them. They are set for 30 kv at about 50 μamp, equivalent to 1 to 5×10^6 ohms. One needle can handle about 170 μamp before breakdown. This figure is relatively independent of gap spacing and pressure. Actually, two sets of corona gaps are used. Along one support they are adjusted to $\frac{1}{4}$ in. spacing and are placed between every hoop.

On a diametrically opposite support $\frac{3}{4}$ in. gaps are used on alternate hoops. The larger gaps have a voltage drop that is less critical to small variations in gap spacing. If a needle were missing from one gap, the charge would accumulate at that gap until the whole voltage of the machine was concentrated there, or at least a large enough voltage would develop to cause a spark to jump across and initiate a much larger destructive spark elsewhere.

Flashovers are prevented along the hexapod legs by rings of $\frac{1}{2}$ in. dural rod snugly fitted both inside and outside the Textolite tubing and lying in the plane of each hoop to which they are electrically connected.

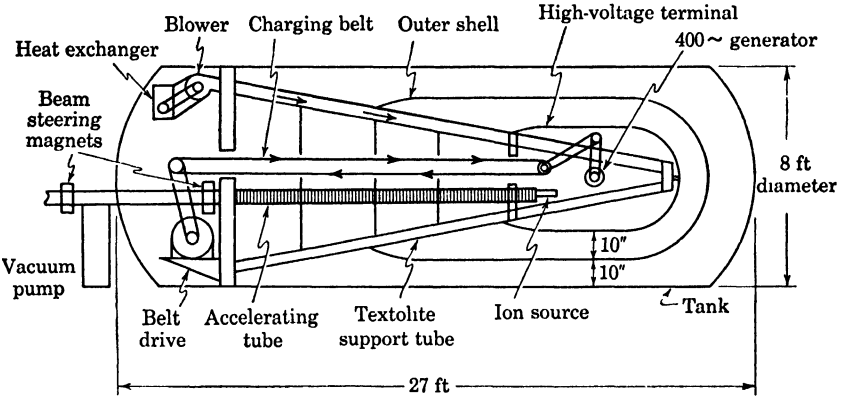


Fig. 7. General schematic of Univ. of Calif. Radiation Lab machine.

Finally, there is the matter of the potential gradient between the high-voltage electrode and tank. As with any concentric geometry, the gradient is greatest at the inner conductor. Following the method first used at Wisconsin, an intermediate conducting shell surrounds the high-voltage electrode and divides the 20 in. electrode-to-tank-wall gap into two smaller 10 in. intervals. This shell is entirely supported from the one end, where it is fastened at a point halfway down the hexapod. Here it is also electrically connected to the mid- or 2 mev-point of the hoop and corona gap voltage divider system. Provision was originally made for two more shells at 1 mev and 3 mev, respectively. They were not found necessary.

The stored electric energy in the Univ. of Calif. Radiation Lab generator is tremendous. The capacitance between outer shell and tank is 1100 $\mu\mu$f, and between outer and inner shell is 300 $\mu\mu$f. At 2 mev each, about 3000 joules of energy is stored—enough for quite a bang!

ACCELERATING TUBE. The accelerating tubes used with electrostatic generators are descendants of Coolidge's hot cathode high-

vacuum x-ray tube.[9] Cold emission at the cathode, electron bombardment of the glass tube walls which tended to puncture them, and runaway arc discharges brought about by bombardment of the anode placed a practical upper limit of 250,000 to 300,000 v on his early tubes. By subdividing a tube into three sections and dividing the total potential difference equally between pairs of tubular electrodes in cascade, he was able to triple this figure. The main features of modern accelerating tubes are the cascading of sections, the use of a voltage divider for equalizing potentials between sections, and the use of electrostatic shielding to minimize potential gradients and corona.

The Univ. of Calif. Radiation Lab accelerating tube is practically identical with that used at Wisconsin.[8] Several sections are shown in

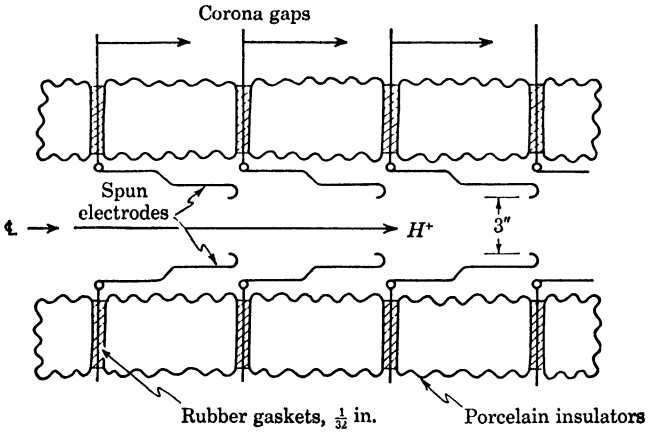


Fig. 8. Accelerating tube (section).

Fig. 8. Four sections, each containing 17 porcelain insulators and spun metal electrodes, form a tube that is approximately 15 ft long. The tube is held together by springs that exert 5000 lb axial end thrust to maintain the vacuum seals. Temporary tie rods are used during assembly only. Corona gaps between electrodes equalize the potential gradient in the same manner as those along the hexapod legs.

Holding a high vacuum inside the tube is quite a problem because there is 110 psi pressure outside. The vacuum pump is at the low-voltage end and maintains 5×10^{-6} mm pressure with a speed of 10 liters/sec. Pressure at the high-voltage end is probably 10 times higher because of the long distance and 3 in. effective inside tube diameter. Considerable outgassing may take place from the seal gasket and insulator surfaces. For an as yet unknown reason, sandblasting the inside

[9] Tuve, M. A., Breit, G., and Hafstad, L. R., "The Application of High Potentials to Vacuum Tubes," *Phys. Rev.*, *35*, 66 (1930).

insulator surfaces improves the tube voltage characteristics and apparently the vacuum characteristics also.

Focusing of the proton beam is accomplished by the field between electrodes. At the end of the tube the beam is reduced to $\frac{1}{32}$ to $\frac{1}{16}$ in. diameter. The pulses are of 1.75 ma for 450 μsec at a repetition rate of 15/sec. This high beam intensity is sufficient to cut through $\frac{1}{4}$ in. thick quartz in $\frac{1}{2}$ sec. Pulsed operation, though required by the mode of operation of the linear accelerator, is also useful in that the continuous current capabilities are much lower; 10 only μamp steady direct current can be obtained. The following explanation has been suggested for the large pulsed current capabilities. Secondary ions and electrons produced in the residual gas of the tube may collect at the accelerating electrodes, and if very large compared with the normal corona gap bleeder current, will distort the potential gradient along the tube. However, in pulsed operation, the capacitance between electrodes (about 30 $\mu\mu$f) is sufficiently great that momentary accumulations of charge do not materially alter the potential gradient.

The ionization of the residual gas is serious for other reasons. The secondary ions and electrons, by striking the insulator walls and electrodes, may cause further outgassing. The electrons freed by the ionizing processes can be accelerated up to 4 mev along the tube in the backward direction to produce an x-ray background dangerous to personnel. One would expect the secondary ionization to be exponential, by analogy with the multiplication process of an electron avalanche in a gas. Assuming one electron freed per secondary ion produced, the number of electrons N_e is given by $N_e = N_{H^+} \exp(-L/\lambda)$, where N_{H^+} is the number of protons in the beam (assuming none are lost), L is the length of the tube, and λ is the mean free path. Naturally, λ is proportional to the residual gas pressure. The x-ray background should then be proportional to the beam current (N_{H^+}) and exponential with respect to the residual gas pressure $\exp(-L/\lambda)$. This has been found to be so. With a good vacuum the x-ray intensity along side the tank may be 1 roentgen per hour. Shielding for these x rays is impractical. Adequate protection would require 12 in. of lead. The only solution is to keep personnel out of the x-ray region during operation of the accelerator.

Ion source. Ion sources that have usually been used with Van de Graaff accelerators contain heated cathodes or filaments to supply electrons which ionize a gas through which they travel. By a suitable arrangement of electric fields some of these ions are caused to pass through a small hole or canal and into the accelerating tube.

The Univ. of Calif. Radiation Lab machine now employs a source using a PIG (Philips Ion Gage) type discharge which has proved much better than the Zinn arc type previously employed. It requires no

filament, the hydrogen pressure is about one-tenth that needed for the Zinn arc, and it delivers greater currents into a more sharply focused beam. To understand its operation, consider first a simplified Philips ion gage (Chapter 8) having two parallel aluminum plates connected to ground through a milliameter. Between the plates is a metal ring at, say, 2 kv positive potential. A magnetic field of 500 to 5000 gausses is maintained between and perpendicular to the plates. Electrons between the plates are constrained to move along the magnetic lines of force (spirally) and will not strike the ring unless they lie in its projected volume. The electric field establishes a potential trough in which the electrons will oscillate about the plane of the ring. The positive ions produced by collisions of the electrons with gas molecules will be collected at the plates and appear as a current through the milliameter. This current is roughly proportional to the gas pressure. If the device is calibrated it may be used to measure pressures down to 10^{-6} mm Hg, using a regulated direct voltage source. An alternating voltage supply

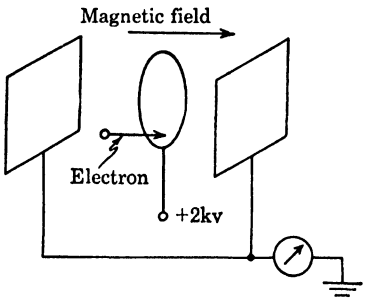


Fig. 9. Philips ion gage.

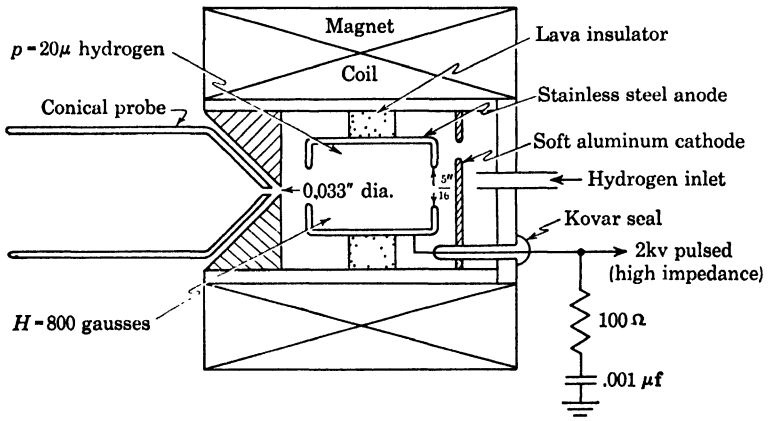


Fig. 10. Modified PIG ion source.

may be used, but the gage is good down to only 10^{-5} mm. The electrons are eventually slowed down and collected by the ring anode. A fresh supply is continually being released by positive ion bombardment of the aluminum plates.

Gow and Foster of the Univ. of Calif. Radiation Lab have developed

this device into a source of hydrogen ions. Their arrangement is indicated in Fig. 10. The cylindrical stainless steel anode replaces the ring. One soft aluminum cathode is pierced by a No. 70 (0.033 in.) drill to allow ions to escape. A conical probe draws these escaped ions into an electrostatic lens system which focuses them into a beam entering the accelerating tube. It was found that a 90° cone and probe geometry yields the largest beam currents.

Hydrogen is admitted at about 20 microns pressure, and an axial magnetic field of 500 to 800 gausses is used. A lower pressure could be utilized if the anode were made longer than the present $1\frac{1}{4}$ in., but the resulting greater magnetic field in homogeneity would permit more electrons to be drained off by hitting the anode prematurely.

The 2 kv high-impedance source allows the PIG to choose its own voltage operating level—about 200 to 300 v. The anode is connected to ground for radio frequency only by a resistor and capacitor in series. This suppresses any plasma oscillations that might occur.

Other metals are suitable as a cathode material; notably beryllium, zirconium, and magnesium. Zirconium has not been tried at the Laboratory. Magnesium sputters but is otherwise all right. Aluminum is used simply because it is cheapest and most easily worked. Metals other than those named could also be used, but would require higher voltage in order to produce sufficient electrons by ion bombardment. A trick used with the aluminum cathode is to admit a little oxygen to form an oxide layer better suited to the production of electrons. Other metals have a less stable oxide coating.

The ratio of electron current to positive ion current (mostly protons) is extremely good for the PIG. In most ion sources the current is inversely proportional to the square root of the electron and ion masses, e.g., $\sqrt{1800} \approx 35$. For the PIG, about one electron is produced for every ion striking the cathode. Since the production of each ion releases 1 electron, there are 6 (original plus 5 released) electrons for every 5 ions, or a ratio of $\frac{6}{11}$ electron current to $\frac{5}{11}$ positive ion current. Three types of ions are produced; H_1^+, H_2^+, and H_3^+ in the ratio 5:4:1. This is very good since the usual ratio is 1:2:1.

The PIG is designed for pulsed operation to synchronize with the pulsed operation of the associated linear accelerator. This necessitates a 450 μsec pulse and a duty cycle of less than 1%. Adequate cooling would allow steady operation if desired. Two amperes peak (50 ma average) current from the 2 kv input yields a maximum resolved proton current of 1.75 ma from those which escape through the 0.033 in. diameter hole. A newer model yields about 6 ma pulsed proton current. The actual current out of the hole is the product of the hole area and the positive ion flux density falling on the cathode at that point. The

current density falls off for larger holes because the discharge runs with a hollow center.

The electrostatic lens focusing arrangement is diagrammed in Fig. 11. The beam focused by this system is less than $\frac{1}{4}$ in. in diameter. There is, of course, 20 microns of hydrogen inside the ion source, which leaks out into the high-vacuum accelerating tube system. The problem is to remove this unwanted hydrogen. It could be allowed to leak down the length of the accelerating tube. However, this would give rise to troubles mentioned previously in connection with the residual gas in the accelerating tube. Ionized gas will also reduce the voltage that the tube will withstand.

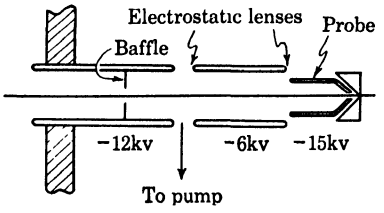


Fig. 11. Electrostatic lens system.

Differential pumping. One way of reducing the residual gas pressure in the tube is to have a fast vacuum pumping speed. Another way is to use a differential pumping arrangement. This consists of placing a baffle in the path of the escaping hydrogen to impede its leakage into the accelerating tube itself. A small opening in the baffle permits the proton beam to pass through. An auxiliary vacuum pump draws off the hydrogen from in front of the baffle so that only a small percentage leaks into the accelerating system.

In the Univ. of Calif. Radiation Lab machine hydrogen is continuously circulated through a closed system and reused in the ion source. An oil diffusion pump (100 liters/sec) backed by a Cenco hypervac mechanical pump removes hydrogen from the electrostatic lens system in front of the baffle. Cooling gas for the diffusion pump is blown from a water cooled heat exchanger at the ground end (see Fig. 7) up through the Textolite tube into the high-voltage terminal. The mechanical pump and its driving motor are enclosed inside a small steel pressure vessel for protection from the high pressure inside the main tank. They are lubricated by diffusion pump oil. Hydrogen is stored at 1 atm pressure inside this pressure vessel. From there it passes through an oil filter and then through a palladium leak into the PIG arc chamber. The leak is a 10 mil heated palladium tube. The hydrogen forms a palladium hydride soluble in the metal, which appears on the other side of the tube. This ensures that only pure hydrogen (no pumping oil) gets into the ion source. Sixteen liters/hr are passed through by using 450 psi pressure and 450 w power. In case of power failure an electric vacuum valve in the intake of the diffusion pump prevents the

system from "going down to air" and dumping hydrogen into the accelerating tube.

POWER AND CONTROLS. The high voltage terminal of the Univ. of Calif. Radiation Lab machine was made large (56 in. diameter by 6 ft long) to accommodate the huge amount of control and auxiliary equipment required for the ion source. Power for this equipment is supplied mechanically through the charging belt which drives a 5 kv 110 v 400 c generator with a permanent magnet field. The generator uses no slip rings or brushes because of the extremely rapid wear these experience as a consequence of the insulation between brush and ring from the high-pressure atmosphere. Brush life on earlier equipment was only 2 to 30 hr.

The pulse timing signal for the PIG ion source is transmitted over a beam of light through one of the Textolite supports up to a photomultiplier tube in the terminal. It is taken from the timing signal which starts the oscillators on the linear accelerator resonant cavity. The buildup time for the linear accelerator cavity is much greater than the combined time delay in starting the arc in the ion source (about 10 μsec) plus the proton transit time through the accelerating tube.

The high-pressure atmosphere creates peculiar equipment problems. The high heat capacity is excellent for cooling but increases the windage work tremendously to the point where most of the belt driving motor power is used to overcome just this windage. The high heat capacity also makes time delay overload relays difficult to use. They do not operate until practically short-circuited, and then there is no time delay. The Freon in the gas mixture decomposes to produce fluorides, which corrode everything but stainless steel. Radio tubes and light bulbs collapse under the high pressure. Electric batteries soak up nitrogen under pressure and then blow up when it is released. Electrolytic capacitors cannot be used.

High voltage for the electrostatic lens system is supplied by straightforward rectifier supplies using Eimac 35R's, which will withstand 200 psi. Selenium rectifiers are very practical for the PIG 2 kv supply. Selenium rectifiers are also used for the 24 v d-c motors, but their poor regulation requires separate supplies for fields and armature currents.

Vibration of the high-voltage terminal is a real problem because its high moment of inertia leads to low resonances. Another difficulty is in the dissipation of the heat from the 30 hp that is fed into the accelerator tank. A heat exchanger must be used at the ground end of the machine, and a compromise must be made between a large efficient one

and a small one which is economical of the expensive space inside the tank but not efficient.

GENERAL. Most of the tank of the Univ. of Calif. Radiation Lab Van de Graaff is detachable near its ground end. It rides on a small railway carriage so that it may be rolled out of the way. The procedure in putting the machine back into operation after the tank has been removed is as follows. Clear dust very thoroughly off all shelves. Bolt the tank together and evacuate for about 8 hr to remove moisture. Put in nitrogen and Freon. The machine then must be operated for from 2 hr to a day's time in order that the high-potential fields may transfer any remaining dust from the shelves to regions of high gradient. The maximum voltage of the machine cannot be attained until this is done.

In regard to the maximum voltage attainable; the machine was designed for up to 200 psi pressure, but the 4 mev required for the linear accelerator is readily realized with 100 psi, and the ultimate voltage limit has not been explored with any thoroughness. Higher pressures would cause some difficulty in increased windage loading of the belt and with instability of the charge spray system, but these possibly could be remedied.

BIBLIOGRAPHY

Hanson, A. O., "Voltage-Measuring and Control Equipment for Electrostatic Generators," *Rev. Sci. Instr.*, *15*, 57 (1944).

Herb, R. G., Parkinson, D. B., and Kerst, D. W., "The Development and Performance of an Electrostatic Generator Operating Under High Air Pressure," *Phys. Rev.*, *51*, 75 (1937).

Parkinson, D. B., Herb, R. G., Gernst, E. J., and McKibben, J. L., "Electrostatic Generator Operating Under High Air Pressure," *Phys. Rev.*, *53*, 642 (1938).

Turner, C. M., Hudgins, A. J., Fillmore, F. L., Jeppson, M. R., and Gow, J. D., "Design of the UCRL Electrostatic Generator," *Univ. of Calif. Radiation Lab, UCRL-352* (1949).

Van de Graaff, R. J., Compton, K. T., and Van Atta, L. C., "The Electrostatic Production of High Voltage for Nuclear Investigations," *Phys. Rev.*, *43*, 149 (1933).

Wells, W. H., Haxby, R. O., Stephens, W. E., and Shoupp, W. E., "Design and Preliminary Performance Tests of the Westinghouse Electrostatic Generator," *Phys. Rev.*, *58*, 162 (1940).

Williams, J. H., Rumbaugh, L. M., and Tate, J. T., "Design of the Minnesota Electrostatic Generator," *Rev. Sci. Instr.*, *13*, 202 (1942).

18

LINEAR ACCELERATORS

GENERAL CONSIDERATIONS FOR LINEAR ACCELERATORS. The linear accelerator of the Sloan-Lawrence[1] type was used to demonstrate that linear accelerators would work. No new nuclear discoveries were made with this type of accelerator, but several known reactions were observed. Figure 1 shows a schematic of this accelerator. The particle enters the first tube at a given velocity (β = ratio of velocity v of the particle to the velocity c of light); then for a time $T/2$ it is in the drift tube. As the particle leaves the first drift tube, the phase of the radio frequency voltage is such that it is accelerated toward tube 2. The

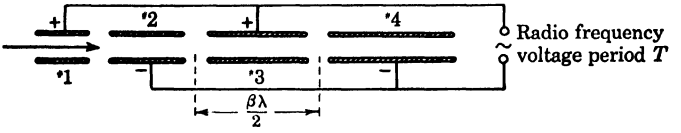


Fig. 1. Basic design of linear accelerator.

length of the second tube is such that in time $T/2$ the particle is in the space between 2 and 3. The radio frequency phase has now changed so that the particle can again be accelerated. The particle takes a time $T/2$ to traverse tube 3 and thus the phase is proper to accelerate the particle again. It travels a time $T/2$ through tube 4, and so forth. At the low frequencies available with power tubes before the war, the size of the Sloan-Lawrence accelerator becomes prohibitive if high velocities are desired. High-power, high-frequency tubes designed and mass produced during the war opened up the field again for further development of linear accelerators.

A linear accelerator is an accelerator that ordinarily does not have both longitudinal and transverse stability. Thus a circular accelerator is not a linear one wrapped around into a circle. Longitudinal (phase) stability is a mode of operation wherein slow particles are speeded up and fast ones slowed down. Transverse or radial stability is a mode of operation wherein particles displaced radially from the axis of the beam tend to be returned to the axis.

[1] Sloan, D. H. and Lawrence, E. O., "The Production of Heavy High-Speed Ions Without the Use of High Voltages," *Phys. Rev., 38,* 2021 (1931).

In a linear accelerator the particles just naturally come out the end of the machine, while in a circular machine, ejection of the beam is one of the problems. The ratio of beam current outside (useful) to beam current inside is much higher for a linear device than for the circular one.

No radiation damping is to be expected in a linear accelerator since there is no radial acceleration as in a circular machine. Radiation damping may set a limit to circular electron accelerators, leaving the linear accelerator the practical device for extremely high-energy electrons.

The beam from a linear accelerator is much more nearly a thin pencil of particles than that from a circular accelerator. This is one of the main advantages of a linear accelerator as a research tool.

The Berkeley proton accelerator has a very homogeneous beam—an energy spread of only about 0.3%. The electron accelerator, on the other hand, has quite a wide energy spread because several electrons traveling at about the speed of light can have energies differing even by several orders of magnitude.

The protons may be injected into their linear accelerator at any convenient energy consistent with other requirements of the whole system (at least 400 kev for the Berkeley accelerator). In the M.I.T. electron accelerator the electrons will be injected at 2 mev, or a $\beta = 0.97$, which is practically the velocity of light. The Stanford machine uses an injection energy of only 78 kev, or a β of 0.5. In an electron linear accelerator the injection velocity must be near the speed of light since the electron accelerators actually operate at a nearly constant velocity and increase the mass of the electron.

The cost of a linear vs. a circular machine is such that for a given ion energy it is usually cheaper to build a circular machine. For electrons in the billion volt range the linear machine, if workable, may be cheaper than the circular one.

PROTON LINEAR ACCELERATOR. The proton accelerator described here is located at University of California Radiation Laboratory, Berkeley. It accelerates 4 mev protons from a Van de Graaff generator to 32.0 ± 0.1 mev. It is being used as a research tool in experiments where a precisely collimated, small diameter ($\frac{1}{8}$ in.), low-background, low-energy spread proton beam is desired.

A resonant cavity oscillator, with drift tubes to isolate the protons except when the field is correct, is used for the accelerator. The big problem is choice of wavelength. Mechanical factors usually point toward a long wavelength. The longer the wavelength, the less electrically long is (L/λ) the accelerator, and the easier is mode separation. The minimum beam aperture size prohibits use of a wavelength less

than 75 cm. Electrical considerations, on the other hand, usually point to a shorter wavelength. The attainable voltage V in an accelerator of length L is

$$V = kP^{\frac{1}{2}}L^{\frac{1}{2}}\lambda^{-\frac{1}{4}} \tag{1}$$

where P is the peak power, λ is the free space wavelength, and k is a numerical factor depending on the geometry details. The Q of the cavity is proportional to $\lambda^{\frac{1}{2}}$ through the skin depth relation, and has been eliminated from Eq. (1). High Q of a cavity helps keep the oscillation energy constant just as a high moment of inertia helps keep the energy of a rotating device constant. As a result the pulse must be at least Q periods long in order for the oscillations to build up to their full energy before the pulse goes off. Similarly, the pulse takes about Q periods to go to zero. Thus a relatively low Q is desired. If U is the energy per pulse,

$$U \propto PQ\lambda \propto P\lambda^{\frac{1}{2}} \tag{2}$$

Substituting Eq. (2) in Eq. (1),

$$V = kU^{\frac{1}{2}}L^{\frac{1}{2}}\lambda^{-1} \tag{3}$$

Thus for a given V, k, L, we see that $U \propto \lambda^2$. Since the equipment cost is proportional to U, the value of λ should be small. These factors were resolved mostly on the basis that equipment at 150 cm wavelength was available with little modification. The pulse length was chosen to be several times QT as in Fig. 2; the repetition rate was then determined by the available power.

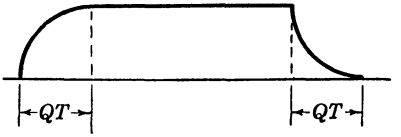


Fig. 2. Radio frequency pulse shape.

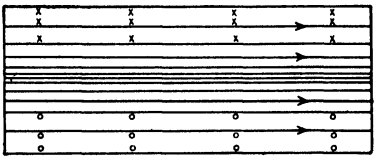


Fig 3. Configuration of electric and magnetic fields in cavity.

The injection voltage was set at a rather high value, 4 mev, since the intention was to accomplish transverse focusing with the aid of thin beryllium foils in which, unfortunately, there would be multiple scattering. It is not difficult to obtain 4 mev protons with a Van de Graaff generator, and one was desired for a general research tool anyway. Subsequently, grids were used instead of the beryllium foils, and a considerably lower injection voltage (400 kev) could be used. The limitations are cavity geometry and possible electron multiplication effects.

There are two choices in the cavity design. (1) The wave guide can

be loaded by insertion of irises to reduce the phase velocity so that it equals the particle velocity (and will constantly increase toward the output end of the tube). The phase velocity is the velocity v, in the expression $e^{i\omega(t-z/v)}$. In a wave guide, this will be one of the Fourier components. There will be but one useful component; the others will do nothing but waste power. If the guide were loaded enough to get the phase velocity down to $\beta = 0.1$ at the beginning (4 mev protons), the other Fourier components would be extremely large and would result in an undue waste of power. (2) The protons can be shielded from the field when it is in the wrong direction by passage through properly designed drift tubes. This is the method actually used.

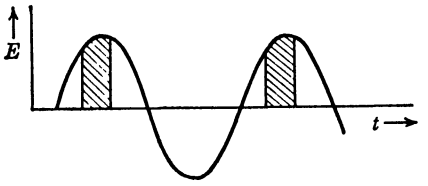


Fig. 4. Accelerating voltage cycle.

The cavity is excited in the axial electric mode (0,1,0), that is, the mode in which an axial electric field with no nodes exists. The magnetic fields are circumferential in nature. Efficient acceleration is possible only when the proton is accelerated with the field at its maximum value, and in no event during a period when the field changes sign. Thus the drift tubes are introduced so that the protons meet only a field in the proper direction. The shaded sections in Fig. 4 are the times when the proton is being accelerated. In Fig. 5 the drift distances get longer because, for the same time, the particles (now traveling faster) go a longer distance. The protons are being accelerated for a time equivalent to $\frac{1}{4}$ to $\frac{1}{3}$ wavelengths. The diameters of these drift tubes become smaller in order to maintain the resonant frequency of the entire cavity at the proper value. The theory for this basic design unit (Fig. 5) is similar to klystron, or other reentrant cavity, theory. The actual voltage gain per stage of a proton is

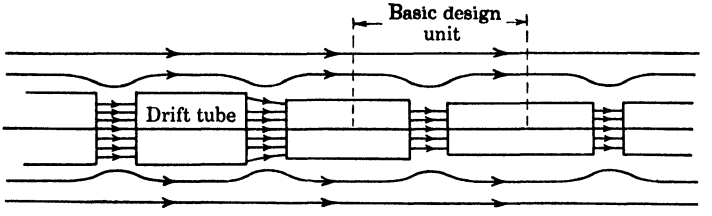


Fig. 5. Linear accelerator cavity.

$$V_1^* = \omega H \cos\phi \frac{\sin \pi f}{\pi f} \qquad (4)$$

where $\omega = 2\pi c/\lambda$, H is the maximum value of the magnetic flux, ϕ is the transit phase angle relative to the phase of a proton crossing the center of the gap at the time of maximum field, and f is the fraction of a wavelength a proton is in the gap. The value of f should be small to cut down transit time losses, but the spacing cannot be too short, because of arcing problems. The maximum voltage V_1 across the gap is related to the shunt impedance per unit cell z_1 and the peak power P by

$$V_1 = \omega H = 2PZ_1 \tag{5}$$

The magnetic fields were evaluated with the aid of exploring loops and numerical integration. From this, Z_1 can be calculated (which is related to the Q), and thus a value of P is obtained for a given V_1. There will be other losses in the actual cavity accelerator that will result in reduced Z_1 and Q, and consequently require a higher P. The new Q can be measured by the relation $\nu_0/\Delta\nu$ where ν_0 is the resonant frequency of the completed cavity and $\Delta\nu$ is the frequency change necessary to reduce the power excitation of the cavity by one-half, the input power staying constant. More peak power will then be required to give the same V_1; this increase will be proportional to the decrease in Q. The value of power for $V = 32$ mev ($f = \frac{1}{4}$, $\phi = 30°$) is 2 megawatts.

Since this design is only approximate, minor corrections must be applied to the cavity geometry in order that the final field distribution shall be the desired one. This involves inserting shims of the proper thickness into the drift spaces. Since large currents flow along the surface of the tubes, care must be taken that the joints are solid. The drift tubes may be positioned with either insulating or conducting rods; it makes no difference to the field distribution.

Longitudinal or phase stability occurs when a slow particle, arriving late, is speeded up at a gap, and a fast particle, arriving early, is slowed down. A particle arriving before A (Fig. 6) is going too fast and receives a smaller impulse than one arriving at A; one coming shortly after A is going too slowly and gets a larger impulse. There is then longitudinal stability at A. A particle arriving shortly before B is going too fast and receives a larger impulse; one arriving after B is going too slowly and receives a smaller impulse. There is longitudinal instability at B.

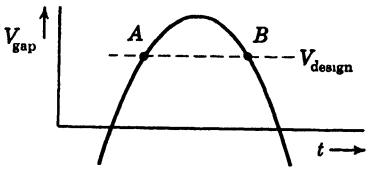


Fig. 6. Voltage across accelerating gap.

The change in radial momentum Δp_r is, at a small constant distance r from the axis,

$$\Delta p_r = e \int_a^b \frac{1}{v} E_r \, dz \tag{6}$$

If v is essentially constant across the gap,

$$\Delta p_r = \frac{e}{v} \int_a^b E_r \, dz = \frac{e\phi_r}{2\pi r v} \tag{7}$$

where θ_r is the total radial flux. If v is not constant there will be second-order focusing, which is not important here (although it is in some d-c accelerators) since the velocity change per stage is small. If the beam crosses no charge, θ_r is positive when the field is increasing and negative when the field is decreasing. For transverse stability Δp_r and thus ϕ_r should be negative, or the field should be decreasing, which corresponds to B in Fig. 6. But this is just the criterion for longitudinal instability. Therefore we are confronted with a fundamental characteristic of linear accelerators: longitudinal and transverse stability are conflicting requirements. If some of the field lines can end in the hole,

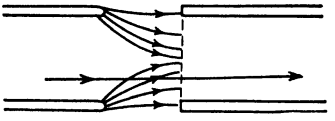


Fig. 7. Focusing of beam.

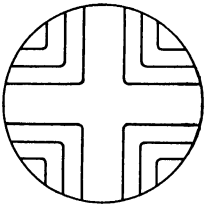


Fig. 8. Grid structure.

the beam goes through (there will then be charge there), and ϕ_r can be negative even if the field is increasing somewhat. It can be seen from Fig. 7 that a proton, if the field stayed constant in time, would cross more inward lines than outward ones, thus giving the desired transverse or radial stability. The grids are used in the entrance hole for flux terminal points. Proper design of these grids, shown in Fig. 8, was developed after some experimentation. The grid elements are slatlike so that flux lines can end on the flat surfaces without reducing the aperture area. Thus the problem of both longitudinal and transverse focusing in the proton accelerator is solved. The phase angle over which both stabilities are achieved is necessarily limited, and only a fraction of the entire incoming beam gets through.

With the recent development of the strong focusing principle,[2]

[2] Courant, E. D., Livingston, M. S., and Snyder, H. S., "The Strong Focusing Synchrotron," *Phys. Rev.*, *88*, 1190 (1952).

however, it is possible to obtain focusing in the linear accelerator[3,4] without use of grids. Electrodes are introduced into the drift tubes and electric fields are applied to them in such a way that focusing is obtained, and appreciable increases in beam current have been observed. The focusing action is quite comparable with that obtained with optical lenses, where a net focusing action occurs when light passes successively through converging and diverging lenses of equal strength, slightly separated.

The resonant cavity is enclosed in a vacuum-tight steel housing about 40 ft long and 4 ft in diameter. The top is hinged 4 in. above the center line and is opened by a pair of hydraulic cylinders. The tank is mounted with a minimum of constraints so that it can expand freely. The cavity is mounted inside the housing so that it is free to expand independently of the housing.

The liner or resonant cavity itself is a 10-sided copper cylinder made like an airplane fuselage, only inside out; for this reason it was made by the Douglas Aircraft Company. A prototype, about 4 ft long, was first made both to test the constructional method and to check resonant frequencies. Extreme care was used in the fabrication and handling of the liner to preserve the intended shape. The tolerances on the liner dimensions were held to 0.005 in.

At each end of the tank is an adjustable tuner made of a movable drift tube that slides through the end of the liner.

Elaborate provisions were made for the lining up of the drift tubes and their final adjustments. Cooling had to be provided for all the parts. Later experience showed that the allowable tolerances on alignment and positioning were much less severe than had originally been considered.

X rays from electron bombardment of drift tube ends constitute the predominant stray radiation. Appropriate shielding has reduced this to a safe level.

The vacuum system is a 30 in. 3-stage oil diffusion pump backed by an 8 in. 2-stage booster working into two paralleled 43 cfm Kinney mechanical pumps. A refrigerated baffle system keeps the oil vapor pressure in the system to a negligible value. The vacuum outlet from the liner is a series of long narrow slots. Since large longitudinal radio frequency currents flow in the liner, care must be taken to remove as little of the circumference as possible. The pumping speed measured inside the liner is 2500 liters/sec or about one-third the rated capacity

[3] Blewett, J. P., "Radial Focussing in the Linear Accelerator," *Phys. Rev. 88*, 1197 (1952).

[4] Bell, J. S., "New Focussing Principle Applied to the Proton Linear Accelerator," *Nature 171*, 167 (1953).

of the 30 in. pump (this is about the fraction to be expected after cooling baffles, etc., are inserted). A series of automatic valves, connected to an ion gage supply, close whenever the pressure rises to an unsafe value.

The proton accelerator is designed to operate with a total voltage of 36 megavolts (this is higher than the energy gain, 32 mev, due to operating phase angle and transit time losses), and the shunt impedance of the cavity on the proper mode is 316 megohms; therefore the peak power required is 2.5 megawatts. The repetition rate is 15 c with a pulse length of 300 μsec (duty cycle = 225), giving an average power of about 10 kw.

Modified BC-677 radar transmitters were used at first, but these proved unsatisfactory. The tubes, GL-434A, were salvaged, and a new oscillator, very tightly coupled to the resonant cavity, was developed. The oscillator itself was quite efficient, giving an over-all d-c to radio frequency efficiency of 50%. The oscillators are coupled to the cavity by a resonant transmission line which utilizes ingenious vacuum seals. It makes little difference where in the cavity wall the coupling loop from the oscillator is placed because of the magnetic field distribution (see Fig. 3). Thus the 25 oscillators are evenly spaced along the cavity, for convenience. In order that the oscillators shall all work at the proper frequency and phase, a loosely coupled set of oscillators running at a frequency determined by their own tunable tank circuits are turned on 100 μsec ahead of the main group. The oscillations build up to about 10% of the final voltage, thus adequately locking in the power oscillators. Several tubes go bad each day of operation; a bad tube can be located and replaced in a few minutes. Other servicing, by replacement of units, is equally easy (and necessary).

"Multipactor" action occurs when the voltage across a gap is just right for an electron to travel across a gap in $\frac{1}{2}$ cycle, knock out several secondaries, which go back to the other side in the next $\frac{1}{2}$ cycle, and knock out several secondaries, which repeat the performance. This occurs near the injection end where the gap spacing is favorable. The action is avoided by feeding in the pre-exciter power at the ejection end; the low group velocity of energy propagation permits the voltage to build up fast enough to avoid this effect.

A pulse network providing a 5 megawatt 300 μsec pulse was constructed in a standard fashion. A smaller pulse for the pre-exciters is triggered 100 μsec before the main pulse. The triggers for the two pulsing operations above and for various monitoring devices are provided by a separate trigger supply. The power supplies have various interlocks for the safety of personnel and equipment.

A deflecting magnet is provided at the output to deflect the proton

beam to any one of the 3 experimental setups. The current output is about 4×10^{-9} amp, accounted for as follows: (1) injection current 50 μa; (2) phase angle loss factor, 15; (3) focusing grid loss factor, 2.5; (4) effective duty cycle, 300. It is expected that the current can be increased by: (1) ion source improvements; (2) focusing grid redesign; (3) bunching of the ion beam prior to entering the linear accelerator, which will speed up the particles due to arrive too late for the cycle and slow down the ones due to arrive too early for the acceleration cycle; (4) improved duty cycle. About 85% of the beam will pass through a hole $\frac{1}{8}$ in. in diameter; the beam has a divergence

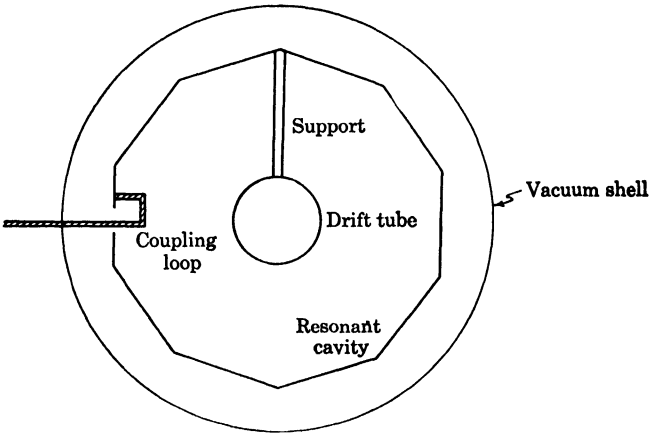


Fig. 9. Cross section view of linear accelerator.

of about 10^{-4} radians. Figure 9 is a schematic cross section diagram of the cavity and drift tubes.

ELECTRON LINEAR ACCELERATOR AT STANFORD. The accelerator is a wave guide loaded so that one of the waves has a phase velocity less than the velocity of light. It uses a traveling rather than a standing wave. The focusing problems are not too great, but the power feed presents difficulties.

It has been shown that longitudinal and transverse stability are conflicting requirements. However, the solution of the problem in the case of the electron accelerator is quite different. The electrons are traveling within a few per cent of the speed of light, and this makes a marked difference in their response to the fields.

Only one wave will be treated in this discussion; other waves are essential in order to satisfy the boundary conditions, but their effects on an electron ordinarily cancel out, since their velocities are different and the forces they produce will oscillate.

If the proper Lorentz transformation is made, the new equations of motion show that transverse stability is not a problem; this is because of the short "new" time that an electron is in the tube, and because most of the particles entering the tube on the correct half cycle get down the tube. This latter phenomenon causes a serious energy spread unless bunching is used. While instability can be ignored, none of the advantages of stability are present; this causes very close mechanical tolerances to be necessary.

Space charge effects turn out to be negligible, also, for the same reason that transverse instability is small.

The simplest electron accelerator would be a single cavity. The shunt resistance of a cavity is $kc\lambda/\delta$, where δ is skin depth, λ is the

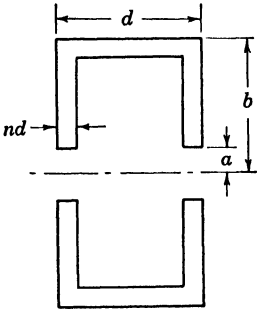


Fig 10. Electron linear accelerator elementary unit.

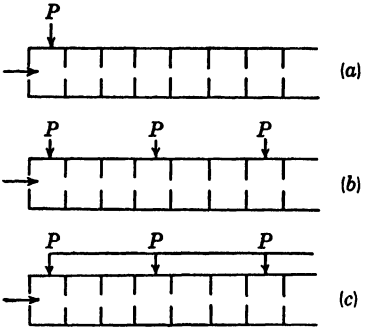


Fig. 11. Different power feed arrangements.

free space wavelength, and k is a geometrical factor. Several fancy shapes have a good k, but when mechanical considerations are taken into account, the shape in Fig. 10, which has a lower k, is decidedly preferable. For high energy gains, several such sections should be put together, since the power losses are proportional to the square of the field and the acceleration is linear in the field. This assumes that the power can be applied in phase to all cavities. If a is made small, little power will be transported down the tube; if a is made larger, the tube will act as a wave guide. There are 5 design factors, λ, a/λ, b/λ, d/λ, and n. The power considerations point to a small λ. In this case, 10.5 cm was chosen. The value of n is not too important and is chosen on the basis of mechanical and space considerations. Three factors are concerned in the choice of d. (1) The longer d is, the higher is the shunt impedance and the less is the power required for a given energy increase; (2) the longer d, the greater are transit time losses; (3) as d is made longer, there will be room for fewer cavities in a given length.

There is a rather broad maximum around $d = \lambda/3.5$. The valve $d = \lambda/4$ was chosen as being the most convenient near this maximum. (The quantity maximized is the effective shunt impedance per unit length.) The valves of a and b are chosen to give the correct phase velocity and an amount of power propagation down the guide consistent with the method of power feed.

A traveling wave rather than a standing wave is used. The tube is terminated with a device such that there are no reflections, that is, the power is all absorbed. No power is actually wasted since the reflected wave could not have accelerated any electrons. A relation between a and b to give a phase velocity of the speed of light was determined empirically. This leaves the ratio a/b to be set by power feed considerations.

The method of Fig. 11a is perhaps the simplest. However, too much power may be wasted out the open end if the field is constant throughout the tube. For very long tubes, practically no power would get to the far end. Figure 11b removes these disadvantages, although the compromise between a good wave guide (power propagator) and a good accelerator still exists to a small extent. In Fig. 11c the power is distributed by an auxiliary transmission line. This in practice would be a complicated method. The following table considers the 3 types of power feed and various accelerator lengths. The body of the table gives the power in megawatts to give an output energy of 10^9 v.

Length (ft)	n	25	50	100	200
Single feed (Fig. 11a)		3000	1700	850	550
Multiple feed (Fig. 11b)	2	2600	1400	770	440
	4	2400	1300	700	380
	8	2300	1200	650	350
	16	2100	1100	610	320
Multiple feed (Fig. 11c)	2	2300	1200	580	290
Separate feed line	4	2200	1100	550	280
	8	2100	1100	530	260
	16	2000	1000	510	260

The following conclusions are drawn by the Stanford group: (1) if the distance between feeds is less than 25 ft, there is no point in introducing the auxiliary feed line, or at least its use must be justified on other grounds; (2) for lengths of 50 ft or less there is not too much to choose between Figs. 11a and 11c, again except on other grounds. However, if tubes were available which could be phased easily, like power amplifiers, the multiple feed method of Fig. 11b might be preferable.

The final arrangement consists of a number of klystron amplifiers phased by wave guides and fed by magnetrons. Approximately 25 mev per 10 ft section is obtained, each section being fed by one high-power klystron. It is hoped that higher voltages per section can be obtained, perhaps up to 5 mev/ft with more refined techniques. Pulses of length 1 μsec are applied, and well-focused electron beams are obtained. Energies in excess of 100 mev have been attained and much higher energies depend on the addition of further sections.

ELECTRON LINEAR ACCELERATOR AT M.I.T. Most considerations for this accelerator are the same as for the one at Stanford. The principal difference is that this is a standing-wave, rather than a traveling-wave accelerator.

The accelerator operates in a π-mode, that is, there is a phase difference of π between adjoining cavities. Figure 12 shows the direction of

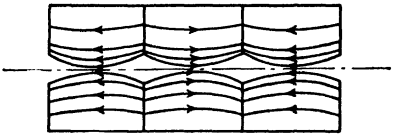


Fig. 12. Electric field configuration.

the field lines. For a greater number of cavities in cascade, the mode separation problems and phasing considerations for the distributed power oscillators are similar to the proton accelerator at Berkeley. For further information, see the paper by Demos.[5]

BIBLIOGRAPHY

Alvarez, L. W. and others, "Berkeley Proton Linear Accelerator," *Univ. of Calif. Radiation Laboratory Report* 236, Nov. 1948.

Glinzton, E. L., Hansen, W. W., and Kennedy, W. R., "A Linear Electron Accelerator," *Rev. Sci. Instr., 19,* 89 (1948).

Livingston, M. Stanley, "Particle Accelerators," *Advances in Electronics.* New York: Academic Press, 1948, Vol. 1.

Slater, J. C., "The Design of Linear Accelerators," *Revs. Mod. Phys., 20,* 473 (1948).

[5] Demos, P. T., *M.I.T. Technical Report 50,* May 1, 1951.

19

CYCLOTRONS

PHYSICAL AND OPERATIONAL CONSIDERATIONS OF NONRELATIVISTIC CYCLOTRON. The conventional cyclotron consists of a large circular vacuum chamber that contains two symmetrically located hollow semicircular electrodes shaped like the capital letter D. These two "dees" form an approximate circle and have their straight edges slightly separated. A source of radio frequency power is coupled to the

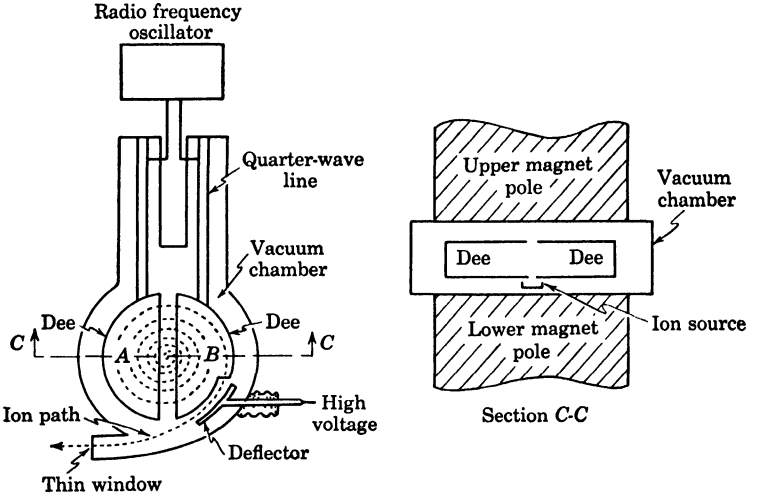


Fig. 1. Plan view of nonrelativistic cyclotron.

dees so that the dees form the high-voltage end of a quarter-wave line. A source of positive ions is provided below the center of dee circle. The entire vacuum tank containing the dees and ion source is placed between the pole faces of a powerful electromagnet so that the plane of the dees is perpendicular to the uniform d-c magnetic field. Usually the magnet has a strength of approximately 15,000 gausses. Figure 1 shows the arrangement of the accelerating electrode system of the cyclotron.

Ions starting out from the ion source at the center of the dee circle are acted on by both the radio frequency field produced between the dees and the magnetic field. The potential difference between the dees

accelerates the positive ion towards the negative dee (A). Under the influence of the magnetic field the ion moves in a circle inside of the electric field free region of this dee. When it reaches the dee gap after making a semicircular path in (A), the electric field reverses and the ion is accelerated across the gap to the new negative (B) dee. Having received an increment of energy the ion travels with greater velocity in (B) and so describes a semicircle of larger radius. The principle of operation depends on the physical fact that a charged free particle moving in a magnetic field describes a circular path with a rotational frequency that depends only on the strength of the magnetic field. The magnetic field is so adjusted that the angular velocity v/r of the ion in the field is just equal to the angular velocity of the radio frequency power supplied to the dees. This condition of proper phasing is called resonance between the frequency of the applied voltage and the strength of the magnetic field. Thus with proper resonance condition, the magnetic field causes the positive ions to revolve outward from the center in a sort of spiral path and thus to be subjected to the accelerating effects of properly phased electric fields between the dees. When the ions approach the edge of the magnetic field they pass through an opening in the side of the dee into a region where they are acted on by a negatively charged deflector plate. The action of the plate pulls them out of the magnetic field and they pass out of the vacuum chamber through a window to a suitable target.

Basic theory of operation. The principle of operation of the constant-frequency cyclotron depends on the following.

1. The kinetic energy gained by a particle of charge e in falling through a potential difference V is equal to the product Ve. Increments of energy are gained by the circulating ions each time they cross the dee gap.

2. Within a hollow conductor containing no charge there is no electric field. Only the magnetic field acts on the ions circulating within the hollow metal dees. The ions within are not affected by the charge on the dee, which reverses sign while the ion is inside in order to keep the electric field between the dees in proper direction for continued acceleration.

3. On a charge e moving with velocity v at an angle θ with a magnetic field of strength H, the force is $Hev \sin \theta$, and its direction is at right angles to v and H. This force causes the ions to travel in circles within the dees such that the centrifugal and magnetic forces just balance. From this, one can derive angular frequency of rotation, which depends only on e, H, and the mass m of the ion. To illustrate:

Let H = magnetic field strength (gausses), e = effective charge on ion (emu), v = linear velocity of ion (cm/sec), r = radius of curvature of ion path (cm), m = mass of ion (grams). Then magnetic force is $Hev \sin \theta = Hev$, since for cyclotron $\sin \theta = 1$. Centrifugal force $= mv^2/r$. For circular motion, magnetic force and centrifugal force must balance; therefore $Hev = mv^2/r$, or $v = Her/m$, or $r = v(m/He)$, where m/He is essentially constant for the constant frequency cyclotron operation. The angular velocity ω of the oscillator driving the system ($2\pi \times$ frequency of applied electric field) must equal the angular velocity v/r of the ions that circulate under the effect of the magnetic field. Hence $\omega = v/r = H(e/m)$, where e/m is essentially constant for the constant frequency cyclotron. For protons whose e/m is twice that for deuterons, the oscillator frequency would be twice as great for the same magnetic field. Now by using principle (1) above, the energy of an ion can be expressed in terms of the radius r of its path.

$$\text{kinetic energy} = \frac{1}{2} mv^2 = Ve \quad \text{or} \quad V = \frac{1}{2} v^2 \frac{m}{e}$$

or

$$V = \frac{1}{2} H^2 r^2 \frac{e}{m}$$

If the ion is subjected to an accelerating voltage E between the dees for n accelerations, $V = nE$ and

$$r = \left(2 \frac{mE}{eH^2} \right)^{\frac{1}{2}} \sqrt{n}$$

Thus the radii of successive paths are seen to vary with the square roots of integral numbers.

4. The kinetic energy of the accelerated particle is small compared with its rest energy, which is m_0c^2 ergs, where m_0 is the rest energy of the particle in grams and c is the velocity of light in centimeters per second.

5. Loss of most of the ions in their long flight is prevented by magnetic focusing obtained by decreasing the magnetic field radially a few per cent. Focusing exists when the lines of magnetic flux are curved outward in such a way that an ion above or below the median plane will encounter a force like that in a current in a magnetic field, tending to bring the positive ion back to the median plane.

CONVENIENT WORKING FORMULAS. A few convenient formulas that simplify calculations and discussion of the engineering features of the cyclotron are based on the following rather odd but practical units.

f = oscillator frequency (megacycles per second)
ϵ = electric charge on accelerated particle (in multiples of the charge on an electron). Example: for $_1H^1$, $\epsilon = 1$; $_1H^2$, $\epsilon = 1$; $_2He^4$, $\epsilon = 2$.
H = magnetic field strength (in kilo-oersteds).
M = mass of accelerated particle (in multiple of mass of proton = 1). Example: for $_1H^1$, $M = 1$; $_1H^2$, $M = 2$; $_2He^4$, $M = 4$.
R = maximum accelerating radius (in.).
W = maximum energy of beam (mev).
V = accelerating voltage (peak voltage between accelerating electrodes).

NONRELATIVISTIC FORMULAS. Resonance relation between the rotating ions in the magnetic field and the frequency of the oscillator in megacycles is

$$f = 1.52 \frac{\epsilon H}{M}$$

The energy (mev) as a function of the radius of the cyclotron is

$$W = \frac{1}{3240} \cdot \frac{\epsilon^2 H^2 R^2}{M}$$

RELATIVISTIC FORMULAS. The total energy E (mev), expressed as a function of the radius, consists of a kinetic energy term and a rest energy term E_0.

$$E = 0.580\epsilon^2 H^2 R^2 - E_0$$

where E_0 has values for ion as follows: H^1, 938 mev; H^2, 1874 mev; He^4, 3729 mev. The frequency (megacycles) expressed as a function of the total energy E involved is

$$f = 1430 \frac{\epsilon H}{E}$$

Should the velocity be required, it may be found from the relation of β as a function of total energy E as follows.

$$\beta^2 = \frac{E^2 - E_0^2}{E^2}$$

where $\beta = v/c$, v = velocity of ion, c = velocity of light.

PRACTICAL LIMITATIONS ON BASIC THEORY OF CONVENTIONAL CYCLOTRON. There is a minimum voltage to which the electrodes of the cyclotron can be reduced and still permit its operation. Two conditions determine this limit:

1. The necessity of having a radially decreasing magnetic field to provide focusing forces so as to prevent loss of ions from the beam.

2. The relativistic effect of increasing mass of the ions as they gain energy.

These conditions decrease the angular velocity of the ions, as can be seen from the formula for ω. This formula is good relativistically if m is the relativistic mass of the particle accelerated. The expression is $\omega = He/m$—where ω = angular velocity (radians per second), H = field (gausses), e = charge on electron (emu), m = relativistic mass of particle (grams). The resonance relation therefore cannot be maintained over the entire path of the particle, and the ions must accordingly gain or lose with respect to the angular velocity of the oscillator voltage. Since this phase error is a function of the number of revolutions and cannot exceed 90°, the number of revolutions is definitely fixed. There is then a minimum voltage across the dees below which the cyclotron will not operate.

The limiting or minimum voltage across the dees expressed as a function of the final energy of the beam and in terms of the practical units outlined above is

$$V = 236 \frac{W^2}{M\epsilon}$$

For high particle energies the dee voltage must be large. It is seen then that the essential feature of the conventional constant frequency cyclotron is not applicable at very high energies. Since the voltage required increases as the square of the energy, the radio frequency power supply becomes a difficult problem in high-energy cyclotrons. For deuterons, a practical limit for high voltages would be somewhere in the region of 50 to 100 mev. To build a constant frequency cyclotron for 100 mev deuterons would be a major engineering project both from the antisparking point of view and from the point of view of the radio frequency power engineering involved. A voltage of approximately 1,200,000 v would be required across the electrodes, and 2000 or 3000 kw would be required at this voltage.

Relativistic ion cyclotron considerations. In considering the problem of the high-energy cyclotron it is important to note that 300 mev protons have a relativistic mass $m = 1.3m_0$ and a velocity of $0.64c$, where c is the velocity of light. These protons are not fast enough for radiation energy loss to be significant. Now from the resonance condition derived above it is evident that to obtain high energies either the magnetic field must be increased or angular velocity ω of the oscillator decreased as the proton gains energy.

$$\omega = \frac{He}{mc} = \frac{Hec}{mc^2} = \frac{Hec}{E}$$

where E = total energy, m = relativistic mass, c = velocity of light, e = charge on particle (esu). Note: $e_{emu} = e_{esu}/c$. The problem is to maintain resonance and focusing at the same time and still have a machine that will be practical to construct. The magnetic focusing

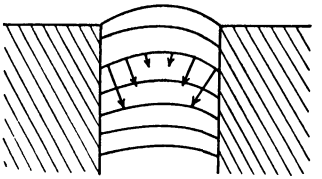


Fig. 2. Focusing in magnetic

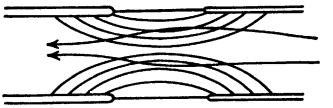


Fig. 3. Velocity focusing.

principle is illustrated in Fig. 2 above, and Fig. 3 shows the mechanism of electric focusing.

The basic solution makes use of a phase stability property possessed by certain orbits in a cyclotron. Call the equilibrium energy the energy that a particle has when its angular velocity v/r just matches the impressed electric field angular velocity. Let the particle cross the gap as the electric field goes to zero so that at an early arrival at the gap the ion will gain energy from the electric field and at a late arrival the ion will lose energy. To show that the orbit is stable, one notes from the above formula that a slight increase in energy reduces the angular velocity and causes the particle to arrive later at the gap the next time. The time of crossing will become later and later until deceleration takes place. This deceleration continues until the particle reaches its original equilibrium energy value. Thus the time of gap crossing and the energy can undergo stable oscillations and the particle will try to stay in or near the stable orbit of a certain equilibrium value for small perturbations. Hence, since it is desired to accelerate the particles, it is necessary only to alter the value of the equilibrium energy slightly. This can be accomplished, as the above equation shows, by variation of either the magnetic field or the frequency of the oscillator.

Variations of these quantities can be accomplished by three general methods, retaining both focusing and resonance conditions:

1. Space variation of the magnetic field.
2. Time variation of the magnetic field.
3. Time variation of the frequency of the electric field.
4. Variation of both frequency and magnetic field.

L. H. Thomas has pointed out that the defocusing effect of the space variation of the magnetic field can be counteracted by introducing a circumferential variation of the magnetic field having an amplitude of order of c/v or $1/\beta$. Although such method is considered possible, the

circumferential field variation would be undesirably large for focusing very high-energy particles.

To carry out method (2) the magnetic field could be tapered slightly with radius to provide focusing and then be made to increase in time in order to maintain resonance as energy of accelerated particle increases. The synchrotron employs method (2), while the bevatron uses (4).

Method (3) is employed in the University of California's 184 in. frequency modulated cyclotron. Variation of the oscillator frequency has the advantage over the other two methods of making use of a simple magnetic field—a most important design consideration. The variation does not have to be extremely rapid and is accomplished by means of motor driven mechanical turning devices such as a rotating capacitor. Since the rate of variation is slow enough, phase stability is maintained throughout the acceleration. No precise control of field or frequency or their rates of variation is necessary even though the charged particle makes many revolutions during the course of the acceleration.

Synchrocyclotron. The synchrocyclotron, or frequency modulated cyclotron, is like the conventional cyclotron with the added principle of phase stability obtained by use of frequency modulation. The higher energies are obtained at the expense of a decrease in the ion current compared to the constant frequency machine.

In the present design of the 184 in. cyclotron it appears to be practical to start the ions at the center of the dee circles as in the case of the constant frequency machine. However, there is only a limited range of times in which the ions so started are locked into stable orbits and never return to the origin. Because of this intermittent operation, for a given dee voltage and vacuum, the average output of the frequency modulated cyclotron is less than that of conventional type under corresponding operating conditions. The fractional part of the frequency modulated cycle in which starting particles may be captured into phase stable orbits is defined as the capture efficiency η.

The *capture efficiency* has a maximum at an equilibrium phase angle ϕ_a of 30° for a constant dee voltage and a varying rate of frequency modulation. Consider $\rho = \sin \phi_a$ proportional to rate of rotation of the capacitor. Then the capture efficiency η is proportional to a function $L(\rho)$. A plot of the theoretical efficiency for constant dee voltage and varying capacitor speed is given in Fig. 4. Efficiency decreases for ϕ_a greater than 30° as a result of the smaller range of phase stability. At phase angles smaller than 30° greater stability is obtained, but efficiency decreases as a result of ions returning to the center. Under typical operation conditions of the 184 in. synchrocyclotron, with the rate of rotation of the capacitor adjusted for a maximum yield, an

experimental value of 0.31% was obtained from an acceptance time of 35 μsec a period $\tau = 0.0114$ sec. The theoretical value from typical operating conditions was 0.33% and in good agreement. Capture efficiencies for synchrocyclotrons are generally of the order of 0.1 to 2%.

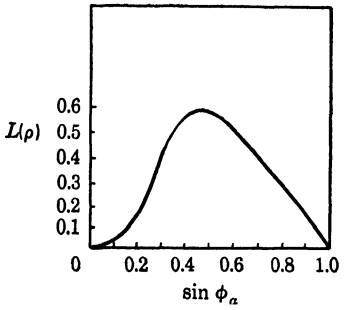


Fig. 4. Capture efficiency vs. phase.

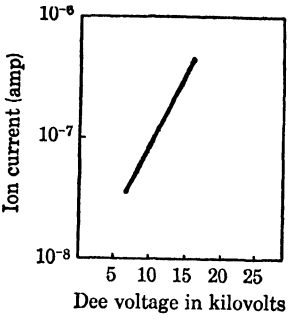


Fig. 5. Ion current vs. dee voltage.

VARIATION OF ION CURRENT WITH VOLTAGE. Figure 5 shows the maximum obtainable ion current as a function of dee voltage when all other adjustments of the 184 in. machine were set to give maximum current. Factors that influence the variation of current with voltage include source current, loss of ions due to collisions, and the capture efficiency variation with dee voltage. The ion current that can be pulled out of the ion source is approximately proportional to $V^{\frac{3}{2}}$. For higher dee voltages the collision loss is less. This is due to the path length, which is shorter before a given energy is gained, and most important to the fact that the shorter path at low energies, where collision loss is greatest, is decreased. At a fixed phase angle, the capture efficiency will vary as the square root of the dee voltage.

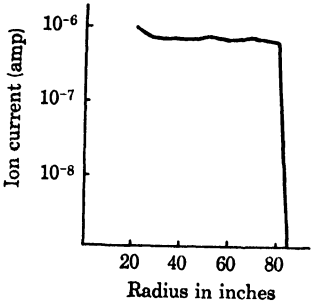


Fig. 6. Ion current vs. radius of orbit.

LOSS OF BEAM AT LARGE RADIUS. Figure 6 is a plot of the deuteron beam current of the 184 in. cyclotron as a function of the radius. These measurements were made in the tank at different radii, starting at 25 in. radius by means of copper probes in shape of a C. The probes were radioautographed to determine why the beam was lost at large radius. Design basis was for the maximum energy to be obtained at a radius of between 84 and 85 in. corresponding to $n = 1$. At $n = 1$, the

motion was calculated to be radially unstable. The variable n at a given radius r and a given field strength H is defined as

$$n = -\frac{R}{H_z}\left(\frac{dH_z}{dr}\right)_{r=R}$$

The coupling between the vertical oscillation and azimuthal inhomogeneities in the magnetic field accounts for the loss of the beam at radius corresponding to $n = 0.2$. Very high accuracy in machine design would be required to exceed this practical value.

DESIGN CONSIDERATIONS. Many cyclotron installations, both in the United States and abroad, are examples of poor engineering design, and consequently are not in full time operation, because of breakdowns. A good philosophy for the design of such a high-priced, high-powered, complicated apparatus is to give careful detail attention to its design. Such a machine involves huge quantities of costly material and auxiliary equipment, and is not readily changed once it is constructed. Alterations are quite expensive and time consuming.

MAGNET DESIGN. Use is made of the H-type magnet, which is quite efficient. It uses a minimum of high-quality magnetic iron consistent with simple mechanical structure that provides adequate space for coils.

The ampere turns required depend on (1) the radius of the ion orbit, which is definitely fixed by the energy of the accelerated particle; and (2) the gap between the poles. This gap, when made small, increases the magnet efficiency but complicates the mechanical construction of the heart of the machine, which must be located in the limited space. Size of the gap is then a matter of design skill and is related to parasitic discharges that will occur if the size of the vacuum chamber is made too large.

The 184 in. magnet is very inefficient. It has a leakage coefficient of about 1.6. The original design called for a gap of 42 in. for constant frequency operation, but with the development of the frequency modulated idea a gap of only 24 in. was needed.

It should be remarked that proper connection of coils is most important because of the large magnetic forces on the coils.

TAPERING OF POLES. Experiments have shown that tapering of the poles does not have an appreciable effect in cyclotron design. Although a greater flux is available in the gap for a specified number of ampere turns, the radius of the magnet is decreased. Model experiments have to decide which is best for the particular purpose for which the magnet

is to be used. Gain is only a few per cent in energy one way or the other.

PROFILE OF MAGNETIC FIELD. Experimental results show that a linearly decreasing magnetic field is preferable to a quadratic shaped field. Linear gradient appears to be better by a factor of 3.

Adjustment of the magnetic field for best operation can be accomplished by employing thin plates of soft iron in the magnetic circuit. Normally a shimming gap is provided in the circuit for this purpose. The amount required will depend, of course, on the imperfections in magnetic uniformity and on mechanical design.

VACUUM SYSTEM. A pillbox with steel covers and nonmagnetic side walls is adequate for a fairly small machine. As nonmagnetic materials, brass castings are sometimes used, but porous casting makes them generally unsatisfactory. Fabricated structure of brass or other nonmagnetic material such as stainless steel is best for this purpose. In large machines, because of the long span involved, deflection of the plate cover and distortion of the magnetic field results. Use is then made of the pole faces as cover for the vacuum system, and the sides of the magnet are gasketed. The dee shape can be varied if desired, but this alters the voltage requirements. They must be rigid, adequately cooled, suitable for the radio frequency system, and economical in construction. A circular piece of copper is usually employed in constructing the dee. A grid is often used to replace one dee.

ION SOURCE. Most of the sources in use are of the low-voltage type but depend on the requirements of a particular problem. It is usually preferable to use extensions on the dee faces called feelers to intensify the high-frequency electric field in the gap between the cone tips, and thus improve ion emission. A hooded arc type source is sometimes used, in which ions are drawn out from one side. The arc source operates at about 3 amp, of which 2 amp cross the $\frac{1}{4}$ in. gap between cone tips. The potential across the arc is about 125 v. The erosion and heating problem is severe, so tips of carbon or Elkonite (an alloy of tungsten and copper or tungsten and silver) are used. These tips do not require local water cooling.

TYPICAL OPERATING CONDITIONS FOR THE 184 IN. CYCLOTRON.[1]

1. The magnetic field is 15,000 gausses at the center of the gap.

2. The dee voltage (average over the cycle) is 20 kv on the dee, or a total maximum possible energy gain per turn of 40 kev.

[1] Henrich, L. R., Sewell, D. C., and Vale, J., "Operation of 184 Inch Cyclotron," *Rev. Sci. Instr.*, *20*, 887 (1949).

3. A direct voltage bias of 1 kv negative with respect to ground is applied to the dee.

4. A dummy dee is used, which is held at ground.

5. The radio frequency is modulated 100 times per second. This is accomplished by a variable capacitor that is rotated at 250 rpm. The capacitor has 24 teeth, so that it then gives $250 \cdot \frac{24}{60} = 100$. The total frequency range is from 12.6 to 9.0 megacycles.

6. The pressure in the main vacuum tank and in the rotating capacitor tank is about 10^{-5} mm Hg.

7. The arc voltage of 150 v. is applied for 100 μsec during each radio frequency modulation cycle. The arc current during this 100 μsec pulse is about 20 amp.

8. The time average beam current is about 1 μa.

BIBLIOGRAPHY

Bohm, D. and Foldy, L., "The Theory of the Synchrotron," *Phys. Rev., 70,* 249 (1946).

Bohm, D. and Foldy, L., "Theory of the Synchro-Cyclotron," *Phys. Rev., 72,* 649 (1947).

Brobeck, Lawrence, MacKenzie, McMillan, Serber, Sewell, Simpson, and Thornton, "Initial Performance of the 184 Inch Cyclotron of the University of California," *Phys. Rev., 71,* 449 (1947).

Brobeck, William, "The Cyclotron," *Elec. Eng., 61,* 1 (1942).

Lawrence, Alvarez, Brobeck, Cooksey, Corson, McMillan, Salisbury, and Thornton, "Initial Performance of the 60 Inch Cyclotron of the William H. Crocker Radiation Laboratory, University of California," *Phys. Rev., 56,* 124 (1939).

Livingston, M. Stanley, "The Cyclotron" I, *J. Appl. Phys., 15,* 2 (1944).

Livingston, M. Stanley, "The Cyclotron" II, *J. Appl. Phys., 15,* 128 (1944).

MacKenzie, K. R., Schmidt, F. H., Woodyard, J. R., and Wouters, L. F., "Design of the Radiofrequency System for the 184-inch Cyclotron," *Rev. Sci. Instr., 20,* 126 (1949).

MacKenzie, K. R. and Waitham, V. B., "R-F System for Frequency Modulated Cyclotron," *Rev. Sci. Instr., 18,* 900 (1947).

McMillan, E. M., "The Synchrotron—A Proposed High Energy Particle Accelerator," *Phys. Rev., 68,* 143 (1945).

McMillan, E. M. and Salisbury, W. W., "A Modified Arc Source for the Cyclotron," *Phys. Rev., 56,* 836 (1939).

Richardson, J. R., MacKenzie, K., Lofgren, E. L., and Wright, B., "Frequency Modulated Cyclotron," *Phys. Rev., 69,* 669 (1946).

Schmidt, F. H., "Mechanical Frequency Modulated System as Applied to the Cyclotron," *Rev. Sci. Instr., 17,* 301 (1946).

Wilson, R. R., "Magnetic and Electrostatic Focusing in the Cyclotron," *Phys. Rev., 53,* 408 (1938).

Wilson, R. R., "Theory of the Cyclotron," *J. Appl. Phys., 11,* 781 (1940).

20

CIRCULAR ELECTRON ACCELERATORS

This chapter is devoted to a discussion of circular electron accelerators. The two types are betatron and synchrotron. In both machines the small mass of the electron produces effects that are both good and bad. First of all, the small mass means that the electron is easily accelerated to nearly the velocity of light and remains at practically constant speed throughout its acceleration period. Many design problems are simplified because of this effect. In the synchrotron, for example, the radio frequency accelerating voltage can remain at constant frequency because the particles always take the same time to traverse the circular orbit at constant radius. The other effect of the small mass of the electron is that the energy loss by radiation due to the centripetal acceleration becomes important at high energies.

The theory of these machines will not be given in detail, since it is very complicated and can be found in the literature. Instead, the physical principles will be given and most of the chapter will be devoted to the techniques incidental to operating the machines and performing experiments with them.

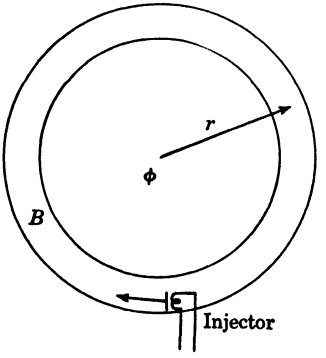


Fig. 1. Schematic of betatron.

BETATRON. In the betatron, developed by D. W. Kerst,[1] electrons are shot from an electron gun into a circular orbit inside an evacuated doughnut-shaped tube. The magnetic field changes sinusoidally with time, and injection occurs when the field is at a low value. As the field B increases the flux contained inside the orbit also increases, the changing flux producing a force on the electron in such a way as to increase its momentum. The value of B at the orbit increases simultaneously, keeping the radius of the electron orbit constant. The rate of change of flux within the orbit and the rate of change of field at the orbit are related by the "flux condition,"

[1] Kerst, D. W., "Acceleration of Electrons by Magnetic Induction," *Phys. Rev.*, *58*, 841 (1940).

$$\frac{d\phi}{dt} = 2\pi r^2 \frac{dB}{dt}$$

The many factors involved in the design of large betatrons are discussed in detail by Kerst and his collaborators in their description of an 80 mev betatron built as a model for a 300 mev betatron. Some operating characteristics of the final design are as follows.[2]

The energy attained by the electrons is 315 mev, the orbit radius is 122 cm. Orbital radiation loss of 9% is compensated by a special separate flux pulse, which is also used to expand the electron beam so that it will strike the x-ray target. Great care was taken in the design and construction to minimize magnetic nonuniformity, and the result was that little difficulty was experienced in getting the machine operating.

The magnet does not operate on alternating current but is pulsed by means of mechanical switching devices. The form of the pulses of flux and field is shown in Fig. 2.

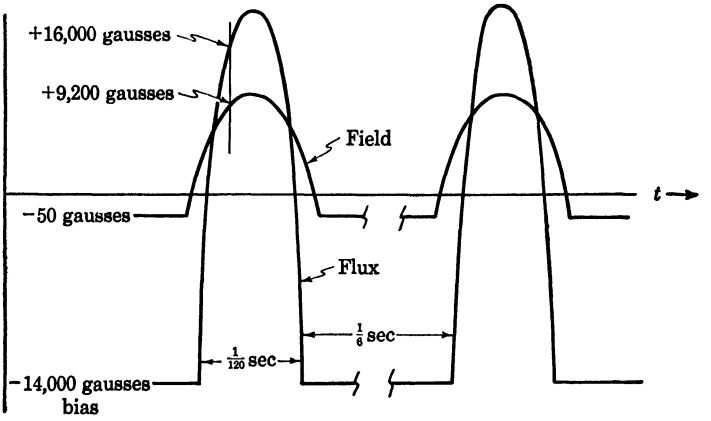


Fig. 2. Time sequence in biased betatron.

Injection is at 80 kv, and the peak current injected is about 1 amp. It was found that injection with a rising voltage waveform improved the yield by a factor of 5 over that found with a flat-topped pulse. The injection pulse is about 5 μsec long.

After the electrons reach their final energy they strike a thin target and produce x rays, which then emerge from the machine. The problems involved in the use of the x-ray beam will be discussed after the other circular electron accelerating machine, the synchrotron, is described.

[2] Kerst, D. W., Adams, G. D., Koch, H. W., and Robinson, C. S., "An 80-Mev Model of a 300-Mev Betatron," *Rev. Sci. Instr., 21,* 462 (1950).

SYNCHROTRON. The idea of *phase stability*, discussed first by V. Veksler[3] and E. M. McMillan[4,5] and already mentioned in Chapters 18 and 19, is used also in the synchrotron. In the synchrotron a magnetic field is used to keep the electrons in circular paths, and a radio frequency electric field accelerates them. The field is applied across a gap which the electrons must traverse once during each revolution in the magnetic field. The frequency of revolution and the frequency of the voltage applied to the gap are the same, and the force experienced by the electron on crossing the gap depends on the phase of the radio frequency voltage. The angular velocity ω of the electron is given by $\omega = Be/mc$. Consider an electron crossing the gap at the time when the radio frequency voltage is zero, and falling. If for some reason the electron's energy is a little too high, its angular velocity will be too low, because the extra energy is in the extra mass which appears in the

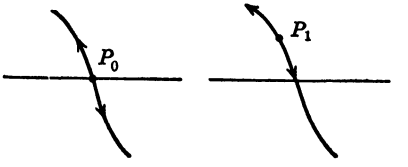


Fig. 3. Phase stability in synchrotron.

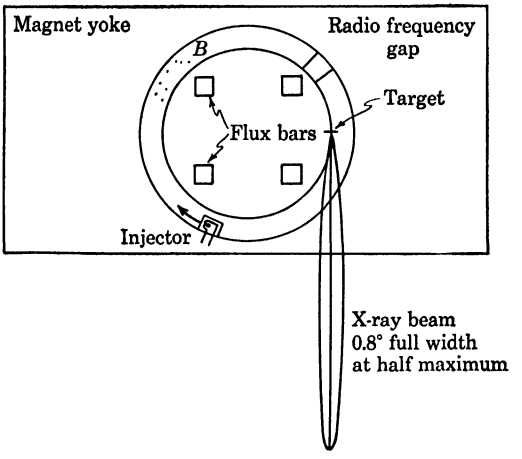


Fig. 4. Schematic of synchrotron.

denominator of the expression for ω. The next time around, the electron will arrive at the gap a little late, and the voltage will have had a chance to go somewhat negative, causing the particle to lose some energy. The particle will be trapped in the phase region about zero voltage, making

[3] Veksler, V., *J. Phys. USSR 9*, 3 (1945).

[4] McMillan, E. M., "The Synchrotron—A Proposed High Energy Particle Accelerator," *Phys. Rev., 68*, 143 (1945).

[5] McMillan, E. M., "High Energy Accelerators," *Helv. Phys. Acta, 23*, Supp. (1950).

oscillations back and forth, and theoretically remaining in a stable orbit indefinitely. If the magnetic field is raised slowly while it is in one of these orbits, the angular velocity of the particle tends to increase because the angular velocity is proportional to the field. The early arrival at the gap results in an increase in energy each turn, however, instead of an increase in angular velocity, and the phase oscillates about P_1 (Fig. 3) instead of P_0 in the equilibrium orbit. The increase in energy per turn demanded by the increasing magnetic field must not be larger than the available voltage at the gap, as otherwise the electrons are accelerated continuously as long as the field rises.

As in the betatron, the radiation loss by the electrons limits the energy attainable; when the energy loss per turn by radiation exceeds the net energy gain per turn due to the radio frequency voltage, the particles are no longer accelerated.

Let us now consider in some detail the operation of the 322 mev synchrotron at the University of California. Schematically the machine is shown (top view) in Fig. 4.

INJECTION. Since the synchrotron is a constant frequency device it is essential to have the electron speed very close to the velocity of light. Thus the injector must supply electrons at several mev. In the Berkeley machine, to eliminate the problem of a high-voltage electrostatic injector, betatron injection is used. The electrons are injected at 100 kev and accelerated for 120 μsec by betatron action. Since the synchrotron does not require any flux *inside* the orbit, it is necessary to provide the flux during the betatron action, and this is done by means of the flux bars through the center of the magnet. The flux bars saturate rather quickly and are not effective during the synchrotron portion of the cycle.

The machine operates at 6 pulses per second. The time sequence during a given pulse is shown in Fig. 5.

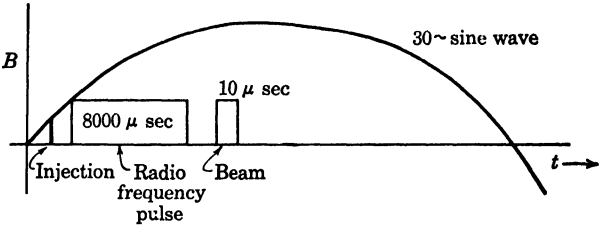


Fig. 5. Time sequence in synchrotron.

When the radio frequency voltage is turned off, the electrons continue to lose energy by radiation, so they spiral in and hit a 0.020 in. platinum target. The length of the output pulse can be greatly in-

creased, if so desired, by shaping the radio frequency pulse. Up to 5000 μsec output pulses can be obtained in this way.

TABLE 1

t	B	*Operation*
20 μsec	12 gausses	Injection at 100 kv, betatron starts
140 μsec	130 gausses	Radio frequency turned on
8000 μsec	10,000 gausses	Radio frequency turned off, 322 mev

USE OF ELECTRON ACCELERATORS. Betatrons and synchrotrons have rather different design problems, but once the x-ray beam is produced, the problems of use are much the same. In a 300 mev machine the angular width of the beam is very narrow. If there were no scattering of electrons in the target, the beam width would be of the order of mc^2/E or $(5 \times 10^5)/(3 \times 10^8)$ radians (0.009°). The scattering broadens the beam to 0.8°.

The distribution in energy of the x rays in the beam follows the bremstrahlung spectrum as given by Heitler (Fig. 6), with the proba-

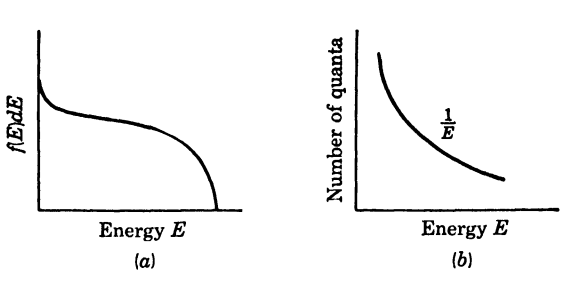


Fig. 6. Energy distribution of quanta.

bility of radiation as a function of energy nearly flat and the energy spectrum of the quanta following a $1/E$ law. The continuous distribution in energy of the x rays must always be considered in using electron accelerators.

The most important difficulty in the experimental work undertaken with electron accelerators is that of background, and the worst source of background is the beam itself. A high-energy x-ray beam striking nearly any material is likely to produce electron pairs. If the experiment has to do with electron pair production, the background problem is not serious, but in other experiments, where the cross sections are of orders of magnitude lower than that of the pair production cross section, the background of electron pairs may be overwhelmingly large, and strenuous measures must be taken to reduce the background.

Shielding is frequently used, but the shields must be carefully designed lest they increase, rather than decrease the background. Figure 7

shows a tapered lead shield, or collimator, the apex of the cone being at the target. The background coming from the walls of the collimator is then blocked by a second lead shield whose walls clear the beam by a good margin. Generally a great deal of lead shielding is used, to collimate the beam and protect the detection apparatus.

A second method of reducing background is to use coincidence and discrimination circuits. The rate of accidental coincidences is diminished for high orders of coincidence and for very short resolving times. Thus scintillation counters with very fast coincidence circuits are frequently used with these machines.

Another experimental problem that is often troublesome is that of pile-up of counts during a resolving time. If more than one event occurs

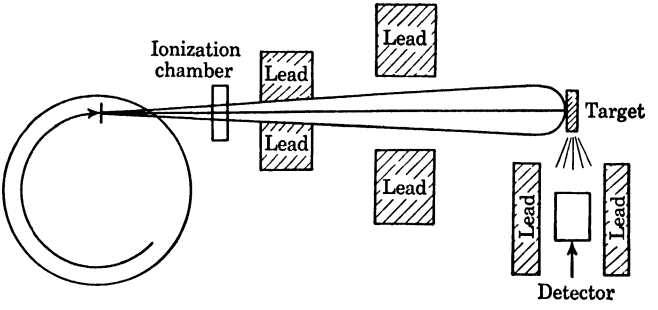


Fig. 7. Typical synchrotron experiment.

during a given resolving time, only one event is recorded. This effect can be checked by varying the beam intensity; if the net counting rate is not proportional to the beam intensity, counts are probably piling up within a given resolving time. A way to get around this difficulty is to lengthen the output pulse by shaping the radio frequency voltage, thus reducing the intensity but keeping the total counts during a given pulse the same.

The beam is monitored by an ionization chamber (Fig. 7), and for experiments where absolute cross sections are measured, the monitoring must be carefully done.

Cloud chambers are often used in conjunction with electron accelerators, but usually the beam must be reduced to a very low value to make the background in a cloud chamber small enough. The cloud chamber is often placed at quite a long distance from the accelerator.

Photographic emulsions have been used to detect production of π-mesons by x rays. A typical setup is shown in Fig. 8. The plates are imbedded in a copper block and serve to sample the meson beam coming from the target.

To illustrate the use of scintillation crystal counters, a typical experiment is shown in Fig. 9. A fast coincidence circuit establishes a coincidence between the two crystals. The stopping π's produce large pulses, which go to a pulse height discriminator. If the pulse is large enough in both crystals, the coincidence circuit operates and the pulses

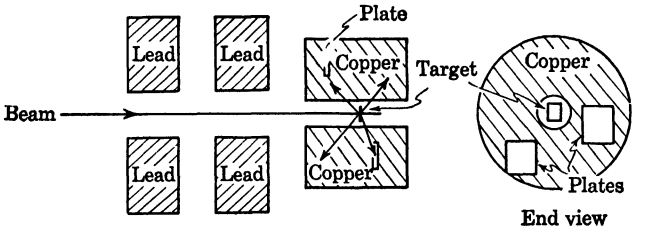


Fig. 8. Photographic emulsion experiment.

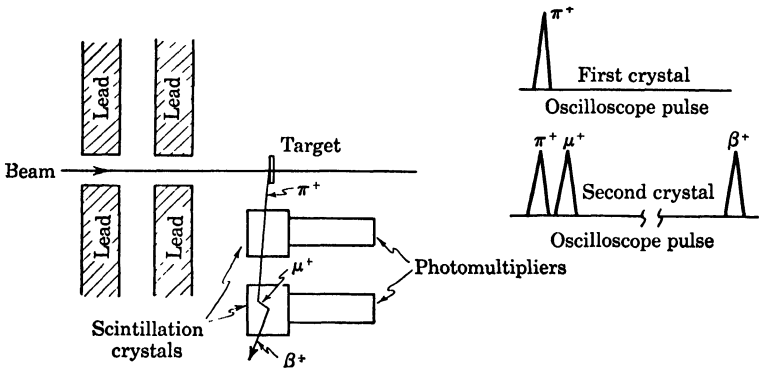


Fig. 9. Scintillation counter experiment.

are displayed on an oscilloscope. The lifetimes for π-μ and μ-β decay can then be found.

In any such experiment the accidental rate is likely to be appreciable. The counting rates were of the following order of magnitude.

Event	*Rate*
$\mu - \beta$	4 per min (30% background)
π coincidence	64 per min
π (single)	200 per min
β (single)	640 per min
$\pi - \mu$	1 per min (20% background)
$\pi - \mu - \beta$	$\frac{1}{3}$ per min (no background)

It is always important to check the background with the target removed. A meson beam is produced at the Cornell synchrotron as in Fig. 10. The beam can be of either positive or negative mesons simply by revers-

ing the direction of the magnetic field. About 10 mesons per minute arrive at the cloud chamber, necessitating counter control.

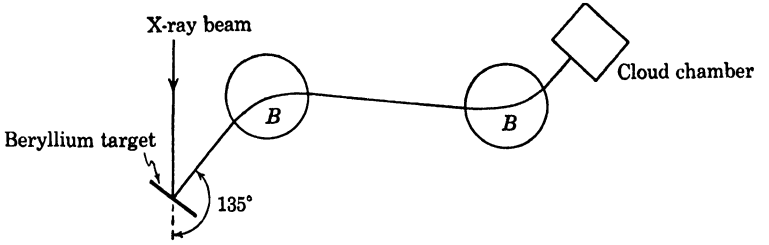


Fig. 10. Meson beam from synchrotron.

Many experiments can be done with the x-ray beam. Following is a partial list.

1. Pair production: $\gamma + Z \rightarrow Z + \beta^+ - \beta^-$, $\sigma \approx 10^{-23}$ to 10^{-24}
2. Compton scattering: $\gamma + \beta \rightarrow \gamma' + \beta'$, $\sigma \approx 10^{-25}$ cm^2
3. Meson production: $\gamma + p \rightarrow n + \pi^+$, $\sigma \approx 10^{-28}$ cm^2
   $\gamma + n \rightarrow p + \pi^-$, $\sigma \approx 10^{-28}$ cm^2
   $\gamma + n,p \rightarrow n,p + \pi^\circ$, $\sigma \approx 10^{-28}$ cm^2
4. Meson experiments
5. (γ,p), (γ,n) nuclear reactions, $\sigma \approx 10^{-28}$ cm^2
6. Compton scattering on protons, $\sigma = ?$

BIBLIOGRAPHY

Amaldi, E. and Feretti, B., "On Two Possible Modifications of the Induction Accelerator," *Rev. Sci. Instr.*, *17*, 389 (1946).

Bohm, D. and Foldy, L., "The Theory of the Synchrotron," *Phys. Rev.*, *70*, 249 (1946).

Kerst, D. W., "The Acceleration of Electrons by Magnetic Induction," *Phys. Rev.*, *60*, 47 (1941).

Kerst, D. W., Adams, G. D. Koch, H. W., and Robinson, C. S., "An 80-Mev Model of a 300-Mev Beatron," *Rev. Sci. Instr.*, *21*, 462 (1950).

McMillan, E. M., "High Energy Accelerators," *Helv. Phys. Acta*, *23*, Supp. (1950).

21

COSMIC-RAY TECHNIQUES

INTRODUCTION. Modern research in the field of cosmic rays is generally directed along one of two lines: to learn more about cosmic rays themselves, or to use the cosmic rays as high-energy probes in nuclear physics. These are not totally separate purposes; they overlap in many areas, and many cosmic ray experiments provide information in both fields. It is the purpose of this chapter to outline briefly the present state of our knowledge of cosmic rays and to indicate the techniques used in the experiments in nuclear physics employing cosmic rays as the bombarding particles.

The high-energy accelerators provide particles with energies up to several billion electron volts, with the prospect of extending the energy range to 100 bev. Cosmic rays provide particles with very much higher energies that produce events that may not be observable with lower energy particles. The inconvenience of working with an uncontrollable source of particles, random in time and very low in flux, is often overbalanced by the fact that certain events can be observed only in cosmic rays.

PRIMARY COSMIC RAYS. Experiments with detecting apparatus carried high into the atmosphere or even above the atmosphere have provided a fairly consistent picture of the primary cosmic rays. They consist largely of protons, but an appreciable fraction of the total energy is brought in by nuclei of greater complexity; α particles, carbon nuclei, and many others up to iron nuclei have been detected. The evidence seems fairly conclusive that few of the primary cosmic rays are electronic in character; γ rays, electrons, and positrons cannot be present in very large numbers.

The magnetic field of the earth has a large effect on the motion of the incoming cosmic ray particles. Although the primary cosmic rays are probably distributed at random in space and time far from the earth, the bending of the particles in the earth's magnetic field creates measurable changes in the distribution. Low-energy particles are rejected entirely, and the distribution of the higher energy particles is noticeably affected. The low-energy cutoff depends on the latitude, and varies from about 10 bev at the equator to about 3 bev at 45°. Changes in the distribution of the higher energy particles with latitude can be

interpreted to give the energy distribution and sign of the primary particles.

The energy distribution seems to be a power law in energy

$$N(E) = AE^{-2.8}$$

The equation applies above the magnetic cutoff energy, and the exact value of the exponent is not known.

The total flux of primaries is of the order of 0.3 particles/cm^2 sec at 50°N latitude.[1]

The upper limit of the energies of cosmic ray particles is not known, but particles of energy 10^{14} ev have been directly detected[2] and there is evidence for energies up to 10^{17} ev. The total energy flux of cosmic rays is small, however, in macroscopic terms, being of the same order of magnitude as that received from starlight.

SECONDARY COSMIC RAYS. When the primary cosmic rays enter the atmosphere of the earth they collide with the atoms in the atmosphere. We can distinguish two general types of interactions. The collisions between the primary charged particles and the atomic electrons produce ionization and excitation of the atoms and a gradual loss of energy of the primary particles. The ionization loss of a fast singly charged particle in traversing the atmosphere is about 2 bev.

The second type of interaction is one that occurs when a primary particle collides with the nucleus of an air atom. In such collisions the interaction is very violent, mesons and secondary nucleons are ejected, and the primary particle may lose a large fraction of its energy. The cross section for nuclear interactions is approximately the geometrical cross section of the oxygen or nitrogen nucleus. Thus the primary particles are exponentially absorbed with a mean free path of about 100 g/cm^2. Since the atmosphere is 10 free paths in thickness, few primary particles reach the surface of the earth.

The secondaries or their progeny, however, are easily detected throughout the atmosphere. In the collisions of primary cosmic rays with the nuclei in the atmosphere, various secondary particles may be produced. Positive and negative π-mesons, neutral mesons, fast nucleons, and slow nucleons are known to be produced. The charged π-mesons decay into charged μ-mesons high in the atmosphere and many of the resultant μ-mesons traverse the atmosphere and reach

[1] Montgomery, D. J. X., *Cosmic Ray Physics*. Princeton: Princeton University Press, 1949.

[2] Lal, D., Pal, Yash, Peters, B., and Swami, M. S., *Proc. Indian Acad. Sci., 36A*, 75 (1952).

the surface of the earth. The neutral mesons decay with an extremely short lifetime (less than 10^{-13} sec) into two γ rays, which then create cascade showers of electrons by pair production and bremstrahlung. The μ-mesons may also decay into an electron and neutrinos and contribute to the electronic component.

The part of the cosmic radiation that produces nuclear interactions is called the N-component and includes the fast nucleons and the π-mesons. Since the μ-mesons do not interact strongly with nuclei they are very penetrating and may traverse large thicknesses of absorber.

At sea level the charged cosmic rays are about 70% μ-mesons and 30% electrons. Some γ rays are present in equilibrium with the electrons, but there is very little of the N component remaining. A convenient figure to remember is that one cosmic ray passes through 1 sq cm of surface at sea level per minute. The absolute intensity of the N component at sea level is not known, but there is still enough present so that some experiments on the interaction of the N component can be done at sea level.

Such experiments will have higher counting rates at higher altitude and often it is necessary to perform cosmic ray experiments at mountain, airplane, balloon, or rocket altitudes. The variation of the different components of cosmic rays with altitude is fairly well known.[1,3] A few examples can be cited and full information can be found in the references at the end of the chapter. For example, from sea level to 10,000 ft, the flux of μ-mesons increases by a factor of 2, the flux of the N component by a factor of 12, and the rate of extensive showers by a factor of 10. Further increases are found at airplane altitudes, but the limited time available for observations often counterbalances the increased flux of particles. Some phenomena occur only at extremely high altitudes. The fast heavy nuclei, for example, are observed only in balloons or rockets near the very top of the atmosphere.

PROPERTIES OF COSMIC RAY PARTICLES. Since the physicist cannot control the flow of cosmic rays or change from one type of particle to another at will, he is often faced with the problem of identification of the particles with which he is working. For this purpose various techniques are available, depending on the measurement or estimate of the physical properties of the particles, such as mass, charge, and spin.

Electrons, with low mass and unit charge, can usually be identified relatively easily. High-energy electrons make large cascade showers, and low-energy electrons can be identified by their large scattering. The measurement of the total flux of electrons is complicated at the

[3] Janossy, L., *Cosmic Rays*. New York: Oxford University Press, 1948.

low-energy end by the effect of the walls of the counters. In any kind of measurements on cosmic ray electrons a thorough knowledge of cascade shower theory is very important, but this complicated subject is beyond the scope of this book. For an introduction to cascade shower theory see Montgomery[1] and Rossi and Greisen.[4]

Heavy particles such as mesons or protons can be accurately identified if their energies are not too high. The μ-mesons, either positive or negative, have a mass 206 ± 3 times the electron mass, unit charge, and spin $\frac{1}{2}$. They have little interaction with nuclear matter, and make up the bulk of the "hard component" at sea level. A μ-meson decays with a mean lifetime $T_0 = 2.15 \times 10^{-6}$ sec into an electron and at least two neutrinos. The decay of the μ-meson has provided the basis for a useful technique where positive identification of particles as μ-mesons can be made. The mass of μ-mesons is large enough so that they rarely produce bremstrahlung in passing through matter. Thus in a cloud chamber the interaction of particles with lead plates often serves to distinguish mesons (or protons) from electrons.

Protons or π-mesons can be identified if their mass can be measured, but this is often difficult if the energies are large. The π-mesons have a mean life of 2.65×10^{-8} sec, and recently experiments have used the lifetime to identify the mesons. The π-mesons probably have spin 0.

The primary particles with higher charge are identified by their large ionization and δ-ray frequency.

For rough identification of particles by scattering, range, and ionization, Powell[5] has prepared some useful curves which give these quantities in terms of range in lead.

It should be emphasized that identification of cosmic-ray particles, especially at high energies, is often uncertain, and considerable care and caution should be used in interpreting results.

Detection techniques. *Geiger counters* are useful in the detection of cosmic rays. Coincidence counting is usually used, the Geiger counters arranged to detect particular types of events. For single particles a train of two or three counters is frequently used, the resulting telescope defining a solid angle and area. Accidental coincidences must always be considered, but are usually negligible for triple or higher order coincidence.

Geiger counters in coincidence are also used to detect showers of particles. A typical arrangement is shown in Fig. 1, where a shower produced by one or more particles above the plate is detected by the

[4] Rossi, B. and Greisen, K., "Cosmic Ray Theory," *Revs. Mod. Phys.*, *13*, 233 (1941).

[5] Powell, W., "A Cloud-Chamber Analysis of Cosmic Rays at 14,120 ft.," *Phys. Rev.*, *69*, 385 (1946).

counters below the plate. More complicated shower detectors are often used with large arrays of counters connected to neon tubes that indicate which counters have been discharged.

Geiger counter data have the disadvantage that any fast charged particle can trip the counters, and interpretation of the results is sometimes difficult.

Proportional counters and *ionization chambers* give information on the ionization produced by the passage of charged particles and thus provide more information than Geiger counters. They are, however, generally less convenient to use, and they present problems in interpretation. It is not always possible to decide, for example, whether a pulse is due to a single particle with large ionization or to a number of particles at minimum ionization. For an approximate method to distinguish these two possibilities, see Bridge *et al.*[6]

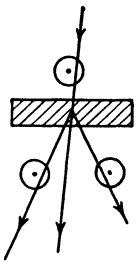


Fig. 1. Apparatus for detecting showers.

Photographic emulsions are extremely useful in the study of cosmic radiation. With the development of various sensitivities of emulsion, including those to detect particles at minimum ionization, many experiments in cosmic rays have been performed and many others have become possible. By counting the developed grains along a track in the emulsions, estimates of the ionization are made. These, combined with measurements of range or scattering, often serve to identify the particle. Stars produced in the emulsion are easily found, and even showers of mesons have been detected and analyzed. The π-meson and the heavy nuclei were first discovered in photographic emulsions. Emulsion plates are light and inexpensive, thus well adapted to balloon work.

Photographic emulsions are not well suited to determination of momentum of particles in conjunction with a magnetic field. Although some work has been done along this line, fields of the order of 100,000 gausses are required to produce magnetic curvature appreciably larger than the scattering curvature.

For momentum measurements and for many other purposes, *cloud chambers* are useful for measurements and observations of cosmic rays. Because of their short sensitive time, cloud chambers are frequently counter-controlled in cosmic ray research. The investigator must always recognize the bias that such control implies, and interpret his results accordingly. Randomly operated cloud chambers are fairly free from bias but analysis of the data may be laborious.

[6] Bridge, H. S., Hazen, W. E., Rossi, B., and Williams, R. W., "A Study of Cosmic-Ray Bursts," *Phys. Rev.*, *74*, 1083 (1948).

Cloud chambers are often used with a series of metal plates with gas between to observe the interaction of the particles with the plates. Such arrangements are particularly valuable in the study of high-energy nuclear interactions where the entire event occurs in the cloud chamber, and much information about high-energy events can be obtained. If the cloud chamber can also be placed in a magnetic field, the momenta of some of the particles may be measured. Since such events are rare, cloud chambers for cosmic-ray research must be made to operate reliably over long periods of time with little attention.

Scintillation counters are used in cosmic-ray research, especially liquid scintillators, which may be made quite large.

MEASUREMENTS IN COSMIC-RAY RESEARCH. Let us now consider some of the measurements required and the accuracies attainable.

The *momentum* of cosmic-ray particles can be measured in a cloud chamber in a magnetic field, by scattering in an emulsion, or by deflection in a magnetic field where counters are used to determine the path of the particle. The upper limit of momentum measurements with a cloud chamber in a magnetic field is about 3×10^{10} ev/c. With long magnetic paths and small counters the upper limit for the second method may be 10^{11} ev/c, while the emulsion techniques depending on scattering and ionization go to about 5×10^{10} ev/c.

Masses of cosmic-ray particles can be obtained by measurement of momentum and range, momentum and ionization, or ionization and range. All three methods have been used, with the most accurate results coming from the first. It is possible to measure the mass of a cosmic-ray meson or proton fairly easily to 10% accuracy, but to increase the accuracy to better than 5%, extreme precautions must be taken. The principal error comes from the measurement of curvature in the magnetic field, where turbulence and scattering may give spurious curvatures.

The *ionization* produced by a cosmic-ray particle can be measured by counting the drops formed along the track of the particle in a cloud chamber. It is necessary first to know that every ion forms a drop around it. This is done by separating the ion columns in an electric field and counting drops in both columns. If, in the case of an alcohol-water mixture, the negative column count is not less than 50% of the positive column count, one is sure that all positive ions have formed drops. The inaccuracies in ionization measurements come from the statistics of a limited number of events and from difficulties with extra large blobs of ionization. The photography must be exceptionally good so that individual drops are easily seen.

Grain counting in photographic emulsions also provides an indirect measure of ionization.

The *range* of a particle can be observed by its penetration of various thicknesses of absorber with detectors spread through the absorber to indicate the presence of the particle. The detectors may be particle counters or may be sections of a cloud chamber. The error in measurement of range can be made very small if the experimenter is willing to use many detectors and thin absorbers.

The *spin* of particles must usually be determined indirectly, either by experiments on nuclear interactions where the spins of the product particles are known or by electromagnetic interaction of particles. Attempts have been made to determine the spin of μ-mesons by the rate of high-energy bremstrahlung, which depends on the spin. Although the experiments have ruled out spin 1, they are not accurate enough to distinguish spin 0 and $\frac{1}{2}$.

SPECIAL TECHNIQUES. Cosmic-ray workers have found it necessary to travel rather more widely than most physicists and sometimes to send their apparatus to places where it is inconvenient for the physicist to accompany his apparatus. There are many laboratories at mountain altitudes, some well equipped with fine buildings, others merely with a temporary tent or trailer. The highest laboratory in the United States is at Mt. Evans, Colorado, with stations at 10,000 and 14,260 ft. Other laboratories near the 10,000 ft level are at White Mountain, California, and Climax, Colorado. In Europe there are several laboratories over 10,000 ft, including Jungfraujoch, Pic du Midi, Mont Blanc, and Testa Grigia.

Many experimenters have taken to airplanes, with the B-29 being the most frequently used. It is pressurized and has a fairly large cabin in the after section. Available B-29's have been known to go up to 40,000 ft, but they are much more reliable at altitudes between 30,000 and 35,000 ft. A B-29 can remain at altitudes conveniently for about 5 hr, so the time is quite limited; B-29 flights are also very expensive.

For higher altitudes balloons are used. There are two types of balloons in general use: Rubber and plastic. The rubber balloons will carry 5 lb to about 95,000 ft, so weight considerations are very important. Often strings of 10 or 20 balloons are used to lift the apparatus. The plastic balloons, which will lift 150 lb to 100,000 ft, are made and flown by General Mills Co. Data on balloon flights can be obtained by telemetering during the flight, but often the equipment must be recovered. The balloons are followed by radio tracking, and the percentage of recovery is fairly high if the flights are made in the proper region, where

the population density is high and the terrain not too rugged. Flights may last for several hours.

To carry cosmic-ray apparatus completely outside the atmosphere of the earth, experimenters have used rockets. Although the rocket may go 150 miles up, it spends only a few minutes outside the atmosphere, and the experiments must be designed with this in mind. The presence of the rocket itself, a large mass of metal, may seriously disturb the measurements on the primary rays because of interaction taking place in the rocket. Many valuable measurements have been made, however, using a rocket "platform" for cosmic ray apparatus.

USEFUL FIGURES. As in other fields of research, it is useful to have in mind certain values for physical quantities that are constantly used. Some of these follow.

Thickness of the atmosphere: 10 m water equivalent
Energy loss of a fast singly charged particle in the atmosphere: 2 bev
Flux of primaries at 50°N: 0.3/cm² sec
Total flux at sea level: 1/min cm²
Radiation length in lead: 0.51 cm
Radiation length in water: 43 cm
Radiation length in air (sea level): 300 m
Energy loss of a fast singly charged particle in lead: 1.1 mev/g/cm²

BIBLIOGRAPHY

Auger, Pierre, *What Are Cosmic Rays?* Chicago: University of Chicago Press, 1945.

Janossy, L., *Cosmic Rays.* New York: Oxford University Press, 1948.

Janossy, L., *Cosmic Rays and Nuclear Physics.* London: Chapman and Hall, Ltd., 1948.

LePrince-Ringuet, L., *Cosmic Rays.* New York: Prentice-Hall, Inc., 1950.

Montgomery, D. J. X., *Cosmic-Ray Physics.* Princeton: Princeton University Press, 1949.

Reviews of Modern Physics, January 1949. Full issue on cosmic rays.

Rossi, B., "Interpretation of Cosmic-Ray Phenomena," *Revs. Mod. Phys., 20,* 537 (1948).

Rossi, B. and Greisen, K., "Cosmic-Ray Theory," *Revs. Mod. Phys., 13,* 233 (1941).

Rossi, Bruno, *High-Energy Particles.* New York: Prentice-Hall, Inc., 1952.

Wilson, J. G., *Progress in Cosmic-Ray Physics.* Amsterdam: North-Holland Publishing Co., 1952.

22

MASS SPECTROSCOPY

GENERAL DESCRIPTION. A mass spectrograph is a device for measuring the mass-to-charge ratio of ionized atoms or molecules. Since the charge is usually known (usually ± 1), we may make an analysis of the mass of the ions formed.

Some of the uses of mass spectrographs are:

1. Analysis of samples, such as oil products, to determine their chemical content.
2. Determination of abundance of ratios of isotopes of elements.
3. Determination of the mass of an isotope exactly so that its mass defect can be determined.
4. Assignment of mass to radioactive isotopes.
5. Use as a separating device to separate usable quantities of a particular isotope.

The mass-to-charge ratio of ions is usually determined by passing them through a magnetic field.

In such a field

$$mv = He\rho \quad \text{or} \quad \rho = \frac{mv}{He} \tag{1}$$

$$\rho = \frac{144.4\sqrt{M\bar{E}}}{H\epsilon}$$

$$\rho = \frac{144.4}{H}\sqrt{\frac{M\bar{V}}{\epsilon}} \quad \text{(ions starting from rest)}$$

where v = velocity in centimeters per second, V = accelerating potential, m = mass of ions in grams, M = mass of ion in mass units (i.e., M for hydrogen = 1), E = energy of ions in electron volts, e = charge of ion in abcoulombs, ρ = radius of circular path in centimeters, H = magnetic field in lines per square centimeter, ϵ = number of electronic charges (i.e., ϵ for singly charged ion = 1).

Ions of the same charge and same energy will strike the detector at a position corresponding to their mass. Thus, in principle, separation of masses can be obtained. In practice two difficulties must be overcome. First, in ionizing and accelerating the ions, it is difficult to make all the ions entering the mass analyzer have the same velocity. The radius of

ions, and compensation for the spread in velocities must be provided. This is called "velocity focusing." Second, the width of the slit that forms the beam is finite and the particles have a finite angular spread. For accurate measurements, provision must be made to bring the ions back to a line focus even though they enter the magnetic field at different angles. This is called "direction focusing." Depending on the application, instruments have been made that provide either or both types of focusing.

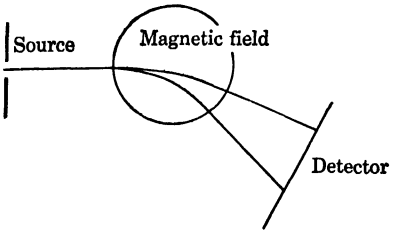


Fig. 1. Mass spectrograph.

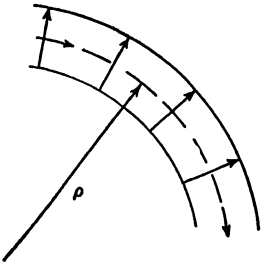


Fig. 2. Deflection in radial electric field.

VELOCITY FOCUSING. According to Eq. (1), there is velocity dispersion in a uniform magnetic field. There is also velocity dispersion in an electrostatic field. Consider a charged particle passing through an electrostatic field arranged so that it always travels perpendicular to the lines of force (Fig. 2). Then

$$Xe = \frac{mv^2}{\rho}$$

where X = potential gradient. Using the same symbols as before,

$$\rho = \frac{mv^2}{Xe} = \frac{2E}{\epsilon F}$$

where F = potential gradient in volts per centimeter, or

$$\rho = \frac{2V}{F}$$

It is seen that velocity dispersion is obtained also in such a (radial) electric field. Now if a combination of electric and magnetic fields is arranged so that the velocity dispersion from one is arranged so as to cancel the velocity dispersion from the other, velocity focusing is

obtained. Aston's mass spectrograph (Fig. 3) used velocity focusing only.

In order to obtain sharp lines with this machine it is necessary to have very narrow slits defining the beam, and the intensity at the detector becomes very small.

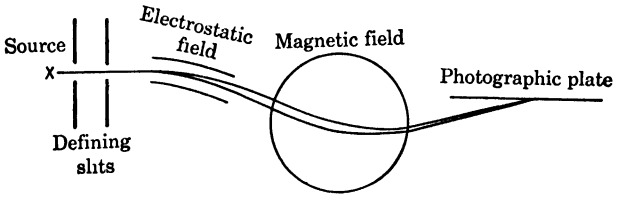


Fig. 3. Aston's mass spectrograph.

DIRECTION FOCUSING. If electric or magnetic fields are properly arranged, divergent beams that are homogeneous in velocity can be brought to a focus. For magnetic fields a simple condition obtains, shown in Fig. 4. The ion source, ion detector, and the center of curvature of the central ion beam all lie on a straight line, and this condition is independent of the angle α of the magnetic field. Angles commonly used are 60°, 90°, and 180°.

Direction focusing can be obtained also with radial electric fields.[1]

In general it is difficult to obtain high resolution and collect a large fraction of the ions formed at the same time. Resolution is usually

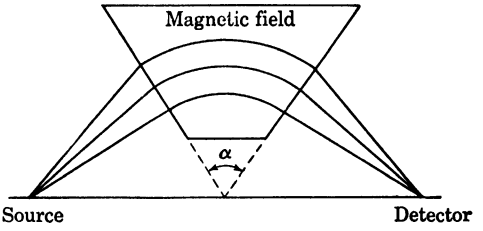


Fig. 4. Direction focusing.

sacrificed for intensity and vice versa. Inghram[2] gives an excellent list of mass spectroscopes which indicates the various types and uses to which they have been put.

The measure of the resolving power of a mass spectrograph is stated in terms of the ratio $M/\Delta M$, where M is the mass of the ion measured

[1] Hughes, A. L. and Rojansky, V., "On the Analysis of Electronic Velocities by Electrostatic Means," *Phys. Rev., 34*, 284 (1929).

[2] Inghram, M. H., "Modern Mass Spectroscopy," *Advances in Electronics*, New York: Academic Press, 1948, Vol. 1.

and ΔM is the smallest resolvable separation of two adjacent masses. Since only a fraction of the total ions formed is collected, another important quantity in mass spectroscopy is this fraction, usually expressed in per cent.

TYPES OF MASS SPECTROGRAPHS.

1. *Single focusing 180° mass spectrometer.*[2] This was first used by Dempster in 1918 and has been used in large-scale separation of uranium isotopes[3] (Fig. 5).

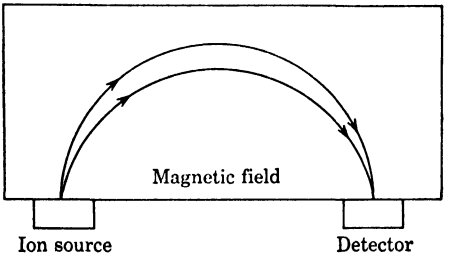


Fig. 5. Single focusing 180° mass spectrometer.

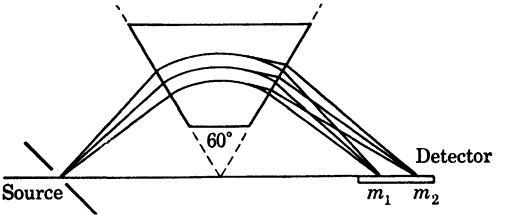


Fig. 6. Single focusing 60° mass spectrometer.

2. *Single focusing 60° mass spectrometer* (Fig. 6). Developed by Nier,[4] this instrument is used for measurements of isotope abundance and is produced commercially for use in chemical analysis.

3. *Double focusing Dempster type mass spectrograph.*[5] This type of mass spectrograph has both direction and velocity focusing. The electrostatic field separates the ions with high energy at A (Fig. 7) and low energy at B. Higher energy ions entering at A have a larger radius of curvature and strike the plate at the same point as the lower energy

[3] Smyth, H. D., *Atomic Energy for Military Purposes.* Princeton: Princeton University Press, 1946.

[4] Nier, Alfred O., "A Mass Spectrometer for Routine Isotope Abundance Measurements," *Rev. Sci. Instr., 11,* 212 (1940).

[5] Duckworth, Henry E., "A Large Dempster-Type Double-Focussing Mass Spectrograph," *Rev. Sci. Instr., 21,* 54 (1950).

ions. The 180° focusing of the magnet provides directional focusing.[6] This spectrograph, as built by Duckworth,[5] has a resolution of $M/\Delta M = 7000$.

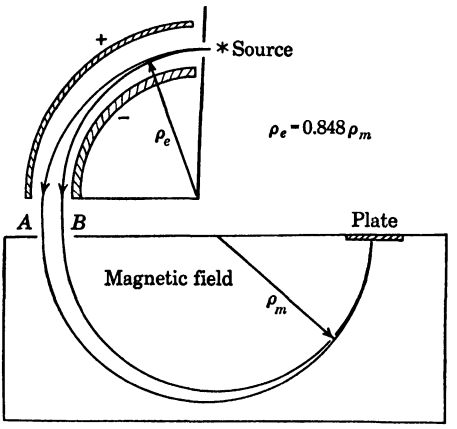


Fig. 7. Double focusing Dempster type mass spectrometer.

4. *Bainbridge and Jordan type.*[7] This is a double-focusing type of mass spectrograph, which provides velocity focusing similar in principle to the Dempster type and provides direction focusing with a 60° magnet as shown. Ions entering at *A* in Fig. 8 are directionally focused at *B*. This type has the advantage of requiring a smaller pole face on the magnet than the 180° type, and is therefore physically easier to construct. It is possible for this type to have a resolution $M/\Delta M$ of 10,000.

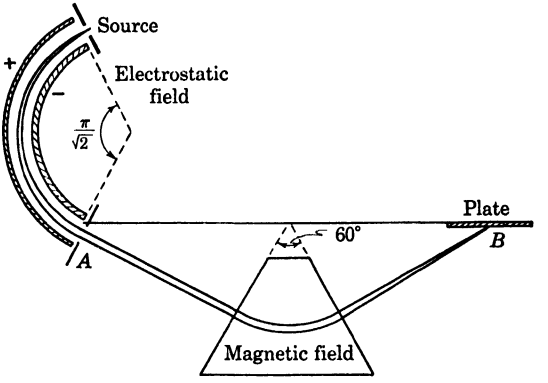


Fig. 8. Bainbridge-Jordan double focusing mass spectrometer.

5. *Mattauch type.*[8,9] This is a double-focusing mass spectrograph using a 90° magnetic field. It has the advantage of having double

[6] Mattauch, J. and Herzog, R., "Über einen neuen Massenpektrographen," *Z. Physik, 89*, 786 (1934).

[7] Bainbridge, K. T. and Jordan, E. B., "Mass Spectrum Analysis," *Phys. Rev., 50*, 282 (1936).

[8] Mattauch, Josef, "A Double-Focusing Mass Spectrograph," *Phys. Rev., 50*, 617 (1936).

[9] Shaw, A. E. and Rall, Wilfred, "An ac Operated Mass Spectrograph of the Mattauch Type," *Rev. Sci. Instr., 18*, 278 (1947).

focusing over all values of mass. A resolving power of approximately $M/\Delta M = 6500$ is claimed for this instrument.

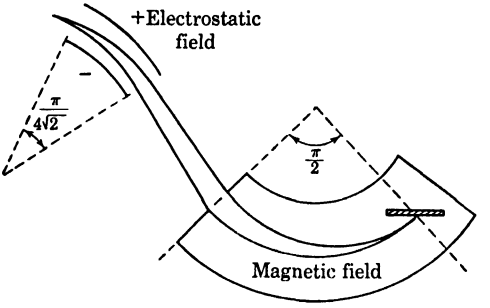


Fig. 9. Mattauch double focusing mass spectrograph.

6. *Bleakney and Hipple type.*[10,11] This double-focusing spectrograph shown in Fig. 10 has a uniform magnetic and electric field at right angles. The ions travel in a cycloidal trajectory (either prolate or curtate). Ions of a given mass leaving at A, regardless of energy or direction, come to a focus at point B. Different masses come to a different focus along the plate. This should have twice the dispersion of the 180° mass spectrograph and should have a resolution of $M/\Delta M = 25{,}000$. Experimentally, a variation of energy of 50% gave no loss in resolution in exploratory experiments.

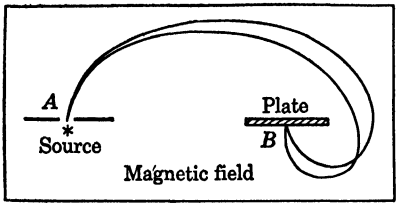


Fig. 10. Bleakney-Hipple double focusing mass spectrograph.

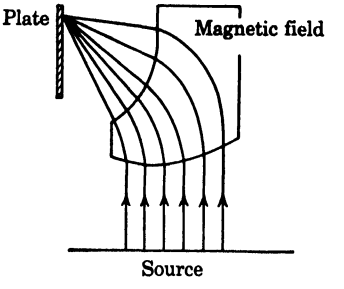


Fig. 11. Magnetic lens mass spectrograph.

7. *Magnetic-lens type.*[7] The magnetic lens mass spectrograph was designed for high intensity. It has no velocity or direction focusing; however, it requires no slit. The ions are accelerated in a parallel field from a 30 sq cm area and are focused in lines corresponding to their

[10] Bleakney, Walker and Hipple, John A., "A New Mass Spectrometer with Improved Focussing Properties," *Phys. Rev., 53,* 521 (1938).

[11] Smythe, W. R., Rumbaugh, L. H., and West, S. S., "A High Intensity Mass-Spectrometer," *Phys. Rev., 45,* 724 (1934).

mass on the plate. If the ions have a low energy compared to the accelerating potential this device can have satisfactory resolution.

8. *Time-of-flight type.*[12] This type of instrument operates on the principle that the transit time between two points in a magnetic or electrostatic field is different for a different m/e. These instruments are at present under development and may prove to be of value. The electrostatic time of flight mass spectrograph has the advantages of having a high percentage of the ions formed collected, and of being much less expensive than the conventional magnetic type. The magnetic time of flight mass spectrograph has both velocity and direction focusing.

TECHNICAL DETAILS.

Fringing. The magnetic field does not drop off sharply at the edge of the parallel plane gap. For a good approximation of the effective field, consider the pole face to be extended 1.5 times the gap width. The fringing field from electrostatic deflection plates can be considered to extend about 0.6 of the gap width.

Collection of charges on surfaces. In designing the electrostatic deflection channels it was found that the deflection plates must be placed farther apart than the ion beam width by quite a factor because of charges collecting on them. If the plates were placed too close to the beam, particles of dust, oil, or other insulating material on them would charge up and deflect the beam erratically. By placing the plates further apart there was less tendency for these spots to charge up. The effect is also reduced, since the small spots of charge tend to be averaged out because of their distance from the beam.

In detecting the ions by a photographic plate there is a tendency for the ions to charge the insulating emulsion. Since the plate charges up positively, it tends to repel the beam and "blow up" the lines on the plate. This effect can be reduced by using a lower beam intensity or by having higher dispersion of the spectrograph.

Space-charge effects. In mass spectrographs in which the ion beam passes through a magnetic field-free region, there can be dispersion of the beam from space-charge effects. The ions repel each other and tend to "blow up" the beam. This effect can be reduced by using a lower beam current.

In instruments that have the beam entirely in a magnetic field this effect does not exist, since electrons from ionized residual gas in the system get "trapped" in the magnetic field along the beam and neutralize the space charge repulsion.

[12] Goudsmit, S. A., "A Time-of-Flight Mass Spectrometer," *Phys Rev.* 74, 622 (1948).

Vacuum requirements. In a mass spectrograph the mean free path of the ions should be long in comparison with the length of the path of the ions in the instrument. Collision of the ions of the beam with residual gas molecules changes the direction and energy of the ion and decreases the resolution. For a path length of approximately 1 m a pressure of less than 10^{-6} mm Hg has been reported to be necessary.

Ion sources. The choice of an ion source is often one of the most difficult problems of mass spectroscopy. It is very difficult to find an ion source that will ionize efficiently and not give the ions too much kinetic energy. A certain amount of energy must be given to an atom in order to exceed its ionization potential. Higher ionizing energies give a higher ionization efficiency. However, high ionizing energies ionize atoms to more than a single charge, which is usually undesirable, and also give the ions a high kinetic energy. The higher the kinetic energy the harder it is to get good resolution and high collection efficiency. Many different sources are described in the literature. Three general types are described below.

1. *Surface-ionization source.*[13,14] This source consists of an electrically heated filament, with the sample to be studied coated on it. If the work function of the surface of the filament is comparable in magnitude with the ionization potential of the sample, a fraction of the sample will evaporate as ions. This type of ionization has the advantages of simplicity, usually ionizing only singly and giving the ions a low kinetic energy. Since the ions have low kinetic energy, velocity and direction focusing are unnecessary. The electrostatic field accelerating the ions from the filament should be parallel so that all the ions will be accelerated in the same direction. The ions can be made to illuminate the slit uniformly, giving better definition. In general this type of source has poor ionization efficiency except for a few elements such as the alkali metals.

2. *Bombardment source.* In this type of source the sample, as a gas or volatilized solid, is bombarded by a beam of electrons. This may ionize the gas to more than a single charge but usually does not give the ions such a high energy that velocity focusing is necessary. The ionization efficiency is usually no higher than that of the thermal source, but the source is of more general applicability and can be stabilized to a high degree.

3. *The arc source.* In an arc source a high-voltage arc is struck between the sample and an electrode. This type of source has a very high

[13] Blewett, J. P. and Jones, E. J., "Filament Sources of Positive Ions," *Phys. Rev.*, *50*, 464 (1936).

[14] Bleakney, Walker and Sampson, Milo B., "A Mass-Spectrograph Study of Ba, Sr, In, Ga, Li, and Na," *Phys. Rev.*, *50*, 456 (1936).

ionization efficiency but multiply ionizes the ions and gives them a high kinetic energy. Both velocity and direction focusing are necessary with this type of source. It is possible to get better illumination of the slit with such a source by applying a magnetic field to the source with lines of force perpendicular to the slit plate.

Detectors. The most common detector of the ions in a mass spectrograph is the photographic plate. The most sensitive plates are the Ilford x-ray plates produced in England or the Eastman equivalent produced in America. These plates have a special emulsion that is especially sensitive to the positive ions. The emulsion is allowed to dry with the emulsion side down so that the sensitive material settles to the surface of the emulsion. In this way the relatively low-energy ions can penetrate the emulsion to the sensitive material more easily.

In exposing the plates it is important that the exposure be of the proper length so that a change in intensity causes a significant change in density. This is particularly true when the ratio of abundance of two masses is to be determined by measuring the density of the lines on a plate.

It is often desirable to assign a mass to a radioactive isotope.[15] One way of doing this is to count the plate with a Geiger counter or ionization chamber before it is developed. A metal partition with a narrow slit in it is placed in front of the counter, and the photographic plate is moved over this slit. In this way the position and intensity of the radiation can be determined. This can be correlated with the mass of the radioactive element after the plate has been developed.

Another way of determining the position of the radioactivity on the plate is to place another plate face to face with it after it has been exposed. The second plate is then exposed by the radioactivity on the first plate. The two plates are then developed, and one plate has an exposure corresponding to the masses of ions present plus radioactivity; the other plate indicates the position and intensity of the radioactivity.

Since detecting ions with photographic plates is usually awkward, it is desirable to have some method of detecting the ions directly. This can be done by collecting them on a plate with a slit in front of it and measuring the current collected with a sensitive electrometer. In such an arrangement the ions may knock off 10 times as many secondary electrons as there are ions and give an error in the current measured of as much as a factor of 10. If this multiplication is not desired it can be remedied by collecting the ions in a Faraday cup instead of on a plate.

[15] Lewis, Lloyd G. and Hayden, Richard J., "A Mass Spectrograph for Radioactive Isotopes," *Rev. Sci. Instr.*, *19*, 599 (1948).

Another more sensitive method of measuring the current directly makes use of an electron multiplier as a detector. This technique is much more sensitive than use of the electrometer, but it has only recently been developed and is not yet fully perfected.

In measuring the current of ions of a certain mass with an electrometer or electron multiplier it is usually convenient to sweep the electrostatic accelerating potential or magnetic field so that the fixed detector can detect different masses. The electrostatic field is somewhat easier to handle, but the intensity of a beam may be slightly dependent on the accelerating voltage. For this reason the magnetic field is usually changed and the electrostatic field held constant.

In changing the magnetic field it is important that the field always sweep in the same direction (from low to high flux or vice versa), since the field is not exactly proportional to the magnet current because of hysteresis of the magnet core.

Stabilization of fields. It is important that the electrostatic and magnetic fields of the mass spectrograph be stable and uniform to a factor better than the accuracy of the machine. This may require a constancy of better than 1 part in 10,000 to 25,000 over the period of the exposure of the photographic plate. Such accuracy is difficult to obtain by ordinary electronic methods. The electrostatic field voltage and current through the magnet may be compared electronically with a voltage standard and automatically stabilized to such accuracy, but obtaining a constant standard is somewhat of a problem. A standard cell is diffi-

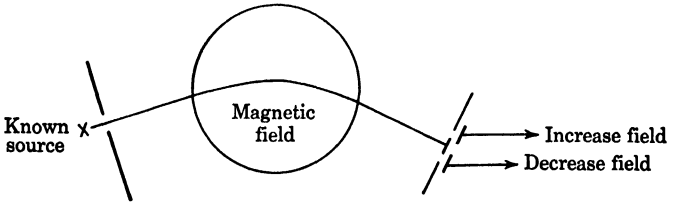


Fig. 12. Stabilization of magnetic field.

cult to handle in an electronic circuit since its voltage is so low. Voltage regulator tubes (preferably a voltage reference tube) or batteries under constant temperature and ideal conditions are practical standards for most purposes. Fortunately, the time of an exposure is usually short enough that the drift of the standard during the exposure will be small.

A more accurate method of stabilization has been used by Nier with very satisfactory results. In his system he uses the trajectory of a known sample in the spectrograph to stabilize the field. A beam of ions of known mass is accelerated by the same potential that accelerates the ions of unknown mass. These ions pass through the same magnetic

field as the unknown ions and strike one of two adjacent plates connected electrically to the field regulating circuit.

If the field is too weak, the beam strikes a plate, causing an increase in field. If the field is too strong, the beam will strike the other plate and decrease the field. This type of stabilization will also compensate for drifts in the electrostatic field. It provides a stability much greater than that obtained using ordinary voltage standards as reference.

BIBLIOGRAPHY

Aston, F. W., *Mass Spectra and Isotopes*. London: E. Arnold & Co., 1942.

Inghram, M. H., "Modern Mass Spectroscopy," *Advances in Electronics*. New York: Academic Press, 1948, Vol. 1.

Jordan, E. B. and Young, Louis B., "A Short History of Isotopes and the Measurement of their Abundances," *J. Appl. Phys.*, *13*, 526 (1942).

Mayne, K. I., "Mass Spectrometry," *Reports on Progress in Physics*. London: The Physical Society, 1952.

23

BETA-RAY SPECTROSCOPY

INTRODUCTION. A beta-ray spectrograph is a device that permits determination of the momentum spectrum of beta rays. The differentiation of beta-ray momenta is usually accomplished in a magnetic field, although some consideration has been given to electrostatic field instruments.[1] Provided the field distribution is known, it is always possible to trace the path of flight of an electron of specified momentum and entrance orientation. However, for other than certain very simple and symmetric field distributions it may be difficult or even impossible to solve the extended source problem in a general way. Nevertheless, experimenters have not confined themselves to entirely homogeneous fields. In this chapter six representative types of beta-ray spectrographs are described, each being a separate approach to the solution of the focusing problem.

An important part of the study of artificial radioactive isotopes is the establishment of a detailed term scheme for the disintegration in question. The intensities as well as the energies of the components of the beta and gamma radiation are of interest, for it is possible from these to draw conclusions regarding the mechanism of the disintegration. The beta-ray spectrograph is a particularly valuable tool for the investigation of these matters since the beta-ray momenta and intensities are yielded directly, or after small corrections to the observed data. Gamma-ray spectra and intensities are also conveniently observed, for it is possible to confine the sample to a small copper capsule thick enough to stop the nuclear beta rays, place this capsule adjacent to a foil radiator, and observe the momenta and intensities of photo electrons or Compton electrons as the situation requires.

In the development of term schemes, it is necessary to make assumptions concerning the sequence of events or to study in some fashion this sequence directly. The beta-ray spectrograph again serves well, for it is possible to measure beta-gamma coincidences as a function of the beta energy interval by mounting a Geiger tube or other counter behind the sample.

Much of the raw data from the beta-ray spectrograph is interpreted

[1] Rogers, F. T., Jr. "On the Theory of the Electrostatic Beta-Particle Energy Spectrograph," *Rev. Sci. Instr., 11,* 19 (1940).

in terms of the Fermi[2] theory of beta decay. This theory was developed to its present form by Konopinski.[3] The nucleus will emit an electron (or positron) in unit time in momentum interval p to $p + dp$ with probability

$$P(p)dp = \text{const. } F(Z,p)p^2(W_0 - W)^2dp$$

where p and W are the momentum and energy of the electron in units of mc and mc^2, respectively, and W_0 is the maximum energy of the beta-ray spectrum. The number of observed electrons in a momentum interval Δp is proportional to $P(p)\Delta p$. Also p (the momentum) is proportional to H, and in turn to the current I in a beta-ray spectrograph. Hence the number of electrons for a given current I may be written

$$N(I) = \text{const. } F(Z,I)I^3(W_0 - W)^2$$

If a plot is now made of $N(I)/I^3F(Z,I)^{\frac{1}{2}}$ versus W, one should obtain a linear function with intercept of $W = W_0$. Typical beta-ray spectra, when analyzed in this fashion, yield segments of straight lines which, when extrapolated to the base line, give values of W_0. Bethe[4] gives tables of $F(Z,p)$ that are good to a few per cent for all values of Z.

Details of beta-ray spectrographs. Beta-ray spectrographs are of two general types—high resolution and high intensity. The former are semicircular in form and lend themselves to both constant field variable radius, and constant radius variable field operation. The first method is especially suited to photographic survey work, but current practice involves the variable-field design with fixed Geiger tube as the detector. On the other hand, the high-intensity spectrographs use solenoidal magnetic lenses, also operating in variable-field fashion with fixed detector.

Lawson and Tyler instrument. The semicircular spectrograph of J. L. Lawson and A. W. Tyler[5] will be discussed first, since it is historically earliest of the instruments reported here and involves the simplest electron flight paths. Consider the diagram of Fig. 1, where H is normal to the paper and everywhere uniform. The central electron starts off perpendicular to the diameter and describes a semicircle of radius ρ

[2] Fermi, E., "Versuch einer Theorie der β-Strahlen," *Z. Physik, 88,* 161 (1934).

[3] Konopinski, E. J., "Beta Decay," *Revs. Mod. Phys., 15,* 209 (1934).

[4] Bethe, H. A. and Bacher, R. F., "Nuclear Physics," *Rev. Mod. Phys., 8,* 195 (1936).

[5] Lawson, J. L. and Tyler, A. W., "The Design of a Magnetic Beta-Ray Spectrometer," *Rev. Sci. Instr., 11,* 6 (1940).

where $H\rho \propto p$, the momentum. Other electrons of the same p, making small angles α with the central ray describe different semicircles and arrive π radians away as measured for the central ray. The resultant line width for a source slit of infinitesimal thickness is

$$d = Z\rho(1 - \cos \alpha) \approx 2\alpha^2$$

The computed intensity profile of such a line is illustrated in Fig. 2. If the source slit is of finite width d', the resultant width of the focused line is

$$d = d' + 2\rho(1 - \cos \alpha)$$

The resolution is defined in terms of the ratio of the half-width of the

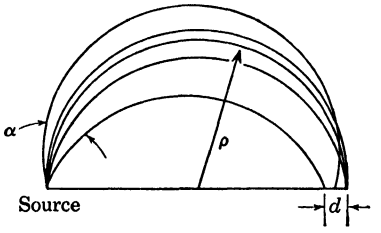


Fig. 1. Semicircular spectrograph.

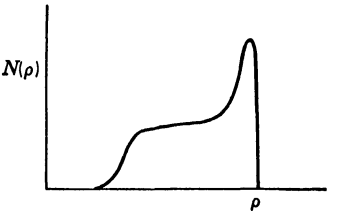


Fig. 2. Line shape.

resolved momentum interval to the momentum at that point.

$$\frac{\Delta p}{p} = \frac{\Delta \rho}{\rho} = \frac{d}{2\rho} + (1 - \cos \alpha)$$

From this it can be seen that the contribution to line width due to the extended source may be reduced to a very small amount for ρ large enough. The contribution from α is not reduced thereby. Also, since there is no transverse focusing, the number of out-of-the-plane electrons reaching the focus point is diminished with increasing ρ. Thus with this instrument resolution can be increased indefinitely but always at the sacrifice of intensity. A practical limit is reached when the count due to a sample of interest is obscured by the no-sample background count.

In the Lawson and Tyler instrument the pole pieces were 16 in. in diameter, tapered to 15 in., with a gap between poles of $2\frac{9}{16}$ in. An exploration of the field showed a region 12 cm in radius, having a variation in field strength of less than 0.5%. A cast aluminum box 16 by 9 by $2\frac{1}{2}$ in. was designed, as indicated in Fig. 3.

Although the form of this spectrograph is different from that of each of the other spectrographs to be discussed, certain design requirements of this instrument are typical of all. These and the details of designs

employed in their satisfaction will be discussed more fully at this point than in connection with the other instruments.

The beta rays from the source at *S* proceed along a semicircle through a defining slit *D* and to the detector slit *d*, behind which is mounted a Geiger counter tube. The other baffles present at *A*, *B*, and *C* are for the purpose of trapping scattered electrons from the box wall. These baffles are made of aluminum plate. All lead or brass surfaces in this and the other beta ray spectrographs are customarily covered with aluminum or some other low-*Z* material such as polystyrene to minimize scattering. Originally the source was located behind a simple slit in the same fashion as the counter tube. A check with a 165 kev particle emitter showed 3.9% of the total count to have been due to scattered electrons. The source was then rebuilt in the fashion shown, the emitter

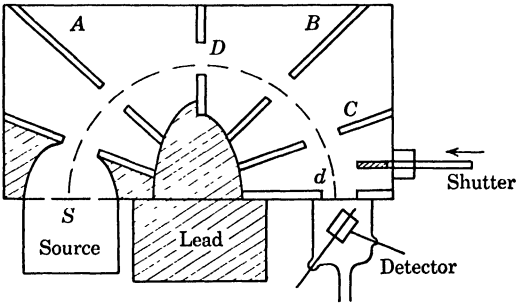


Fig. 3. Lawson-Tyler instrument.

being deposited on a thin film of zapon. Back scattering from the source support and scattering from the source slit edges was thereby virtually eliminated, bringing the scattering count down to 0.1%. This was barely above the background count of 3 or 4 per minute, determined by moving the brass shutter into place over the counter window.

The source was prepared by depositing a thin layer of the emitting material on a thin film 1 or 2 microns thick. These films are frequently prepared by dropping collodion, zapon, or nylon in a volatile solvent onto a water surface, then lifting the resultant film off by means of a wire loop. The emitter is deposited from solution by evaporation or precipitation, the latter producing more uniform deposits. A thin source is necessary to reduce absorptions of beta particles in the emitter itself and to reduce back scattering. Since a thin source is required, it must be fairly large in area to produce sufficient intensity.

The Geiger tube design is indicated in Fig. 3. The entrance slit, 4 by 16 mm, also supported by the window, is formed of 1.5 micron cellulose film. The tube pump-out was manifolded to the box pump-out so that the tube and box could be evacuated simultaneously, thus per-

mitting no large pressure differential across the counter window. The tube was then filled with hydrogen for counting. It was found that the thin window could support up to 25 cm hydrogen before failure. A Neher-Harper counter circuit was used in conjunction with the tube. A large lead block was interposed between counter and source to absorb gamma radiation from the emitter.

The vacuum problem for beta-ray spectrographs is not great. Pressures of 10^{-4} mm Hg to 1 or 2 microns are quite acceptable. Generally, a good mechanical pump is adequate, although many investigators use oil diffusion pumps as well. Sections of the apparatus are rubber gasketed together and moving parts are operated through Wilson or stopcock grease seals. Lawson and Tyler make no mention of their specific evacuating system, but certainly they did not strive to achieve very low pressures. They remark that 2.7 cm of air in the box decreased their intensity by a factor of 2 only for the 165 kev beta rays.

The resolution achieved by Lawson and Tyler was about 1.6%, but the transmission factor was low. From 0.5 to 0.1% of all source particles in the 4π solid angle were received by the Geiger tube. This is generally regarded as too low to permit much work with artificial radioactive material, especially where coincidence methods are required, although for some survey work the high resolution is a desirable feature.

Langer and Cook instrument.[6] This instrument represents a more modern and refined approach to the problem of designing a very high resolution spectrograph. It was found by analysis by Beiduck and Konopinski[7] that practically perfect focusing at π radians from the source could be achieved by introducing a field with the proper radial inhomogeneity. The field distribution is represented as follows.

$$H = H_0\left(1 - \frac{3}{4\delta^2} + \frac{7}{8\delta^3} - \frac{9}{16\delta^4} + \frac{51}{320\delta^5} \cdots\right)$$

where $\delta = (\rho - \rho_0)/\rho_0$ and ρ_0 is the radius of curvature of the $\alpha = 0$ particle.

A cross-sectional view of the Langer and Cook spectrograph magnet is shown in Fig. 4. On the basis of experiment with a quarter-scale model, a distribution was achieved that very closely approximated that of Konopinski for perfect focusing. The ρ_0 in this instrument is large—40 cm—and the magnet itself measured 41.5 in. along the

[6] Langer, Lawrence M. and Cook, C. Sharp, "A High Resolution Nuclear Spectrometer," *Rev. Sci. Instr.*, *19*, 257 (1948).

[7] Beiduck, R. M. and Konopinski, E. J., "Shaped Field for 180° Spectrograph," *Phys. Rev.*, *73*, 1229(A) (1948).

diameter. Since ρ is so large, only a moderate field is required in order to measure the highest energy beta particles. The field current is electronically regulated. The field is calibrated by means of the photo electron ejected from the K level of lead by the 0.5108 mev annihilation radiation.

The spectrograph box was very similar in principle to that of the Lawson-Tyler instrument although made in this case from five 30° wedges of cast aluminum joined together with rubber gaskets. The defining slit was made of polythene and adjustable through a Wilson seal. The detector slit was made of aluminum-clad tungsten.

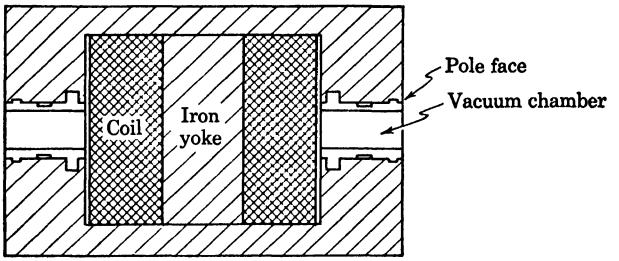


Fig. 4. Langer-Cook instrument.

The counter windows were made extremely thin, being a stack of 5 double films each 0.01 mg per cm^2 thick of zapon, supported on a drilled grid of lucite. The counter was filled with a mixture of argon and alcohol to a total pressure 9 cm Hg. The counter window could pass beta particles of energy as low as 2.5 kev.

The source film was supported from a thin mica member, in turn supported by a skeleton framework at the end of a rod. The source area was usually made to be 0.4 by 2.5 cm. For the measurement of gamma-ray energies by means of the photoelectric effect, the source material was placed in a copper box, on the front face of which was a thin lead radiator 0.4 by 2.5 cm. The copper box absorbed nuclear beta rays from the source material.

The defining slit was usually adjusted to accept a total angle of 32° at the source. This arrangement permitted a transmission of about 0.1% of all beta rays in the 4π solid angle. An analysis of the In^{112} spectrum resolved the internal conversion line at 173 kev with a full width at half maximum of 0.5%.

SIEGBAHN-MCMILLAN INSTRUMENT. Another type of high-resolution beta-ray spectrograph utilizing a radially inhomogeneous field was proposed by McMillan in private communications. Published[8]

[8] Svartholm, N. and Siegbahn, K., *Arkiv. Mat. Astron. Fysik, 33A*, No. 21, 28 (1946).

theoretical investigations were made by Svartholm and Siegbahn. An instrument based on these considerations was built at the University of California Radiation Laboratory.

The field distribution for axial focusing in the betatron is given by

$$H = H_0\left(\frac{r_0}{r}\right)^n, \qquad 0 < n < 1$$

Under this field condition, the frequency of radial or axial oscillation from the median path is given by

$\nu_r\sqrt{1-n} = \nu$ radial $\qquad \nu_r$ = rotational frequency

$\nu_r\sqrt{n} = \nu$ axial

Now, if in the beta-ray spectrograph $n = \frac{1}{2}$, then $\nu = \nu_r\sqrt{\frac{1}{2}}$ in each case. Hence a focus at $\pi\sqrt{2}$ is achieved for electrons starting out of the median path. A sectional view of the magnet is shown in Fig. 5.

In this apparatus the magnet could also be made the vacuum envelope. The resultant transmission was 0.5% of 4π solid angle. The resultant resolution was 0.5%.

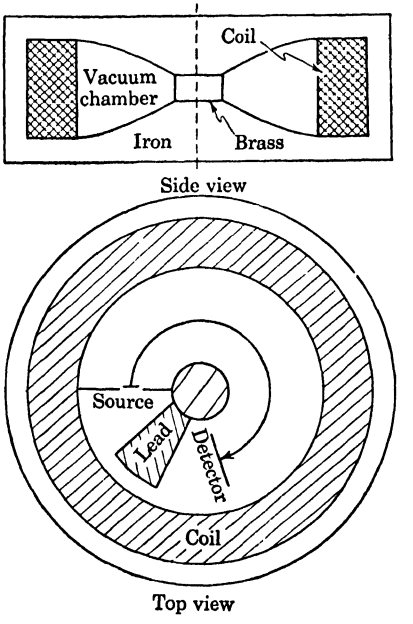


Fig. 5. Siegbahn-McMillan instrument.

WITCHER INSTRUMENT.[9] This is one of three high-intensity spectrographs to be discussed, all of which are solenoidal in general configuration. These instruments employ a field that is largely coaxial with the source-detector axis of the instrument. These instruments all accept beta particles leaving the source in a hollow cone, a section of which is a ring. It is reasonable to expect that this geometry would yield a much greater solid angle of acceptance than the semicircular type.

The long electron lens spectrograph of Witcher will be discussed first since it is historically earliest and since it employs an axially

[9] Witcher, Clifford M., "An Electron-Lens Type of Beta-Ray Spectrometer," *Phys. Rev., 60,* 32 (1941).

homogeneous field. This instrument is diagrammed in Fig. 6. Note that the axial velocity component is unaffected by the field, while the transverse component is affected. A simple derivation shows that

$$L = p \frac{2\pi \cos \alpha}{He}$$

where L = focal distance, p = momentum of electron, H = magnetic field.

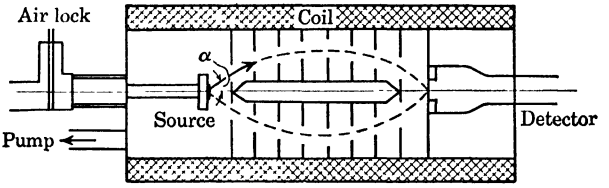


Fig. 6. Electron lens spectrometer.

Thus the focus point at L is dependent not only on p and H, but on α. For high resolution α again must be confined to a small range about some median cone. In this instrument, for an extended source, there is again a compromise needed between resolution, efficiency, and source area. A complete discussion of the design considerations is contained in Witcher's paper.

The instrument took the form of a cylinder of brass 66 in. in length and 10.5 in. outside diameter, upon which was a six-layer solenoid of copper tubing. A current of 600 amp produced a field of 1115 gausses homogeneous to 1%. The working region of the instrument was 90 cm long and was equipped as shown with a series of defining baffles lying on sine curves of 7.16 and 12.5 cm amplitude for the inner and outer baffles, respectively. The inner baffle was supported on a 1 in. diameter aluminum rod of ample length so that no gamma ray from the source could reach the counter. The source of ≈ 2 cm. outside diameter was mounted on a rod and had a window of cellulose film supported on a drilled metal disk. The background count was about 10 counts per minute.

This instrument had an effective solid angle of acceptance of about 0.122 steradian or 1% of the 4π total. The resolution was tested at $1387.5H\rho$ and found to be $\pm 3\%$. The upper range of particles brought to focus in this field was 4.44 mev. No trouble was experienced with particles down to 100 kev. The author felt that with compensating coils the instrument should be usable down to 10 kev.

Deutch, Elliot, and Evans instrument.[10] The instrument of Deutch, Elliot, and Evans is also of the high-intensity type but em-

[10] Deutch, M., Elliot, L. G., and Evans, R. D., "Theory, Design and Applications of a Short Magnetic Lens Spectrometer," *Rev. Sci. Instr.*, *15*, 178 (1944).

ploys a short magnetic lens, coaxially mounted with the source-counter axis. This instrument focuses in a much different fashion than the Witcher spectrograph, depending for its action on the "chromatic aberration" properties of a thin magnetic lens. A fairly complete discussion of the electron optics of this lens is given in footnote 10. However, it is approximately true that f the focal length, i the coil current, n the number of coil turns, and p the momentum of the particle are related by $f = C(p/ni)^2$ where C is a constant depending on the shape of the coil. If u and v are the source and image distance respectively, electrons of one particular energy will be focused in accordance with the thin lens formula,

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v}$$

the other electrons being lost in the baffle system. Thus for a given electron energy the focal length and magnification can be varied independently in order to yield the best condition for a particular problem.

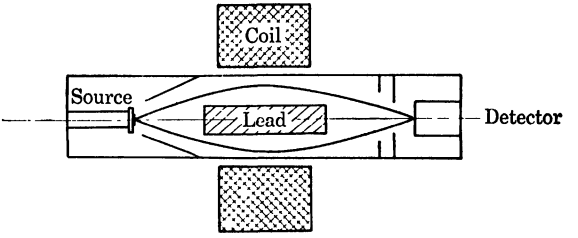


Fig. 7. Short lens magnetic spectrometer.

A schematic of the actual apparatus is shown in Fig. 7.

The chamber in this case was made of brass and had a maximum length of 200 cm with an outside diameter of 6 in. The material and small diameter were found to result in a rather high scattered electron count. The coil measured 25 cm outside diameter 16 cm long, and was movable along the vacuum chamber axis. As in the Witcher spectrograph, annular baffles defined the accepted electrons and minimized scattering from the walls. A lead baffle was inserted when strong gamma-ray sources were used. The baffle arrangement required adjustment to suit the focal length and magnification chosen for each problem. In this type of instrument both positrons and electrons are focused to the same position but rotate in helixes of opposite sense. A paddle-wheel baffle was inserted to polarize the instrument and permit only positrons or electrons to reach the counter window. Compensation coils surrounding the entire instrument minimized the effect of the earth's magnetic field. For coincidence studies a Geiger tube was mounted directly behind the sample.

The main characteristic of this instrument is its flexibility. It was possible when adjusted for high resolution for $\Delta p/p$ to be reduced to 1.7%. When adjusted for high transmission $\Delta p/p$ was increased to 6%. Magnification was adjustable over the range from 0.5 to 2.0. The maximum energy beta particles focusable was found to be 5 mev.

SIEGBAHN-HAYWARD SOLENOIDAL SPECTROGRAPH. One of the most recent and modern of the solenoidal spectrographs was proposed and built by Kai Siegbahn[11] in 1946. Almost identical to this and performing in the same fashion was the instrument reported by Raymond Hayward.[12] The field in this type of instrument is again axially inhomogeneous, being 30% stronger at the two foci than at a point halfway in between. It has been shown that an axially inhomogeneous field of infinite intensity at the foci and descending to a low value at the center (as diagrammed in Fig. 8) would act as a lens completely free of spherical aberration, leaving only chromatic aberration, upon which its operation depends. In a practical instrument the field distribution actually used gave a close approximation to this condition. A diagram of the complete instrument is shown in Fig. 9.

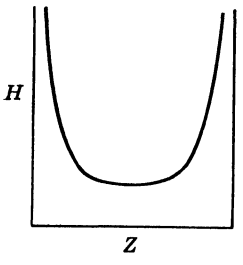


Fig. 8. Magnetic field distribution.

The coils were wound on a brass cylinder 20 cm in diameter and 60 cm in length. These were completely encased by an iron tube constructed of 2 in. square longitudinal bars. Cylindrical pole pieces of iron

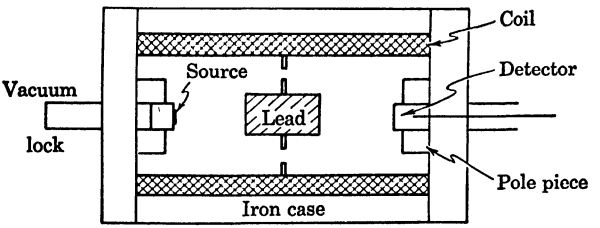


Fig. 9. Solenoidal spectrograph.

4 cm deep and 13 cm in diameter were fixed to iron end pieces 5 cm thick and 40 cm in diameter. Pole pieces, end pieces, and iron tube act to produce the desired field configuration. One central diaphragm was present with an annular defining orifice and bearing at its center a lead

[11] Siegbahn, Kai, "A Magnetic Lens of Special Field Form for β- and γ-Ray Investigations, Designs, and Applications," *Phil. Mag., 37*, 181 (1946).

[12] Hayward, Raymond, *Ph.D. Thesis*, 1949, University of California Radiation Laboratory, 582.

gamma-ray absorber. Surfaces were covered with lucite or aluminum. Holes were cut in the end pieces for the entrance of the Geiger tube and source material. The Geiger tube was of conventional design, having an 0.8 cm diameter nylon window and filled with argon and ethylene.

The source in this instrument had some interesting features. When not intended for use in coincidence counting, it was conventionally supported on a thin nylon film, in turn supported on the end of a polystyrene tube that was fixed to a metal rod. The source could be introduced through a vacuum valve. When coincidence counting was desired, a hollow support rod was substituted for the solid metal one. The hollow rod was closed at the source end with a beryllium-copper foil, which maintained the vacuum and absorbed the beta radiation from the source. Behind this absorber was placed an anthracene crystal attached to the end of a polished lucite rod that transmitted light pulses to a photomultiplier scintillation counter.

The magnet coils were powered by a 5 kw generator. The coil current was regulated to better than 0.1% by electronic regulation of the generator field excitation. In the lens there was no departure from linearity between the coil current and focused $H\rho$, as established by direct field measurement and by measuring the $H\rho$ values of known conversion lines. Occasionally a helical baffle was employed for differentiation between electrons and positrons.

The observed transmission factor of this instrument was about 2% of the 4π solid angle and the total half-width $\Delta p/p$ was measured at 5%.

CONCLUSIONS. A wide variety of beta-ray spectrographs is available, each having properties which may be most suited to the experimenter's particular problem or pocketbook. (The spectrographs of Lawson and Tyler or Witcher are undoubtedly less costly than those of Langer and Cook or Hayward.) The resolutions available in these instruments vary from a $\Delta p/p$ of 0.5% up to 6 or 8%, while the transmissions vary from 0.1% up to 2% of the total solid angle.

BIBLIOGRAPHY

Bethe, H. A. and Bacher, R. F., "Nuclear Physics," *Revs. Mod. Phys.*, *8*, 195 (1936).

Deutch, M., Elliot, L. G., and Evans, R. D., "Theory, Design, and Applications of a Short Magnetic Lens Spectrometer," *Rev. Sci. Instr.*, *15*, 178 (1944).

Langer, Lawrence M. and Cook, C. Sharp, "A High Resolution Nuclear Spectrometer," *Rev. Sci. Instr.*, *19*, 257 (1948).

Persico, E. and Geoffrion, C., "Beta-Ray Spectroscopes," *Rev. Sci. Instr.*, *21*, 945 (1950).

24

OPTICAL SPECTROSCOPY

SPECTROGRAPHS

Spectroscopic instruments can generally be classified into two main groups: prism instruments and grating instruments. Each has certain advantages over the other that makes it more suitable in its own field. These can be summarized as:

Speed. Prism instruments are, in general, faster than grating instruments.

Dispersion. Gratings can be constructed for higher dispersion.

Resolution. Resolving power is higher for grating instruments.

Linearity. The dispersion is nonlinear in prism instruments but stigmatic characteristics can be achieved.

The following is a very brief discussion of the construction and mounting of the different instruments.

PRISM INSTRUMENTS. A standard instrument consists of three main parts: the collimator, the prism, and the camera. The dimensions of each part varies from one instrument to another according to the characteristics required. The speed of the instrument is determined by the f value of the camera lens. Typical values for ordinary intensities are $f/10$ to $f/20$. For very high speeds when low intensities are encountered, special construction of fast cameras is adopted. In the Schmidt optical system a correction plate is placed at the center of a spherically concave mirror with a curved photographic plate placed at the focus (Fig. 1). Such cameras are rated at $f/0.5$. For ordinary purposes a regular achromatic lens is used. Since achromatic lenses are corrected for only two colors, an extra correction is sometimes necessary by curving the photographic plate so as to have all colors in focus. Another method of correction that is usually applied in ultraviolet quartz instruments is to have a noncorrected lens with a tilted plate to correct for the different colors, the optical axis of the crystal being parallel to the light. This method is used for ultraviolet analysis down to 2000 A.

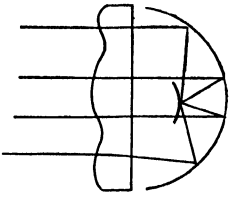


Fig. 1. Schmidt optical system.

The collimator lens does not actually give any serious difficulties.

For high-speed cameras it is not necessary at all to use a high-speed collimating lens. If we use a longer focal length lens, it is true that the light gathered by the same aperture will be less, but this will be compensated for by the image size. The image size depends on the camera lens, and for small images, high intensities are achieved.

The slit is an important part of the instrument. The movable jaw should be guided to prevent any lateral motion in the direction perpendicular to the direction of the adjustment. A spring is attached to the movable jaw so that in adjustment a screw pulls them apart and the spring pulls them together. If the split is opened gradually from zero width, the diffraction pattern will change in shape as indicated in the diagram. The intensity as indicated by the light of the central peak rises with the opening of the slit to a certain limit (Fig. 2). After that limit an increase of the opening will cause an overlapping of the patterns from the different points, and hence an increase of the width of the line without any further increase of the intensity. This limit is found to be when the pattern on lens of collimator is 1.5 to 2 times the lens aperture (see Sawyer).[1] In adjusting an instrument a small flashlight bulb is put across the room, the slit is narrowed until its image covers the lens, and the condition is satisfied (Fig. 3).

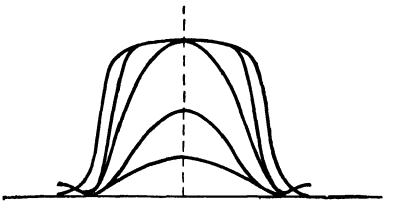


Fig. 2. Diffraction patterns from a slit.

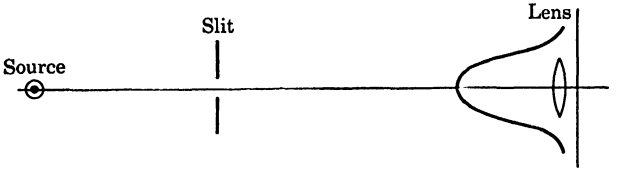


Fig. 3. Illumination of collimating lens by slit and source.

DISPERSION AND RESOLVING POWER.

$$\text{angular dispersion} = \frac{d\theta}{d\lambda}$$

$$\text{linear dispersion} = \frac{dl}{d\lambda} = f\frac{d\theta}{d\lambda}$$

where f = focal length of camera lens, or

$$\frac{dl}{d\lambda} = \frac{f}{\cos\chi}\frac{d\theta}{d\lambda}$$

if plate makes an angle χ with the normal to the axis.

[1] Sawyer, R. A., *Experimental Spectroscopy*. New York: Prentice-Hall, Inc., 1951.

The expression $d\theta/d\lambda$ can be put as $(d\theta/dn)\cdot(dn/d\lambda)$, where n is the refractive index of the material. Here $d\theta/dn$ is a property of the prism, i.e., its dimensions, while $dn/d\lambda$ is a property of the material. It can be shown that

$$\frac{d\theta}{d\lambda} = \frac{B}{A}\frac{dn}{d\lambda}$$

for minimum deviation (Fig. 4).

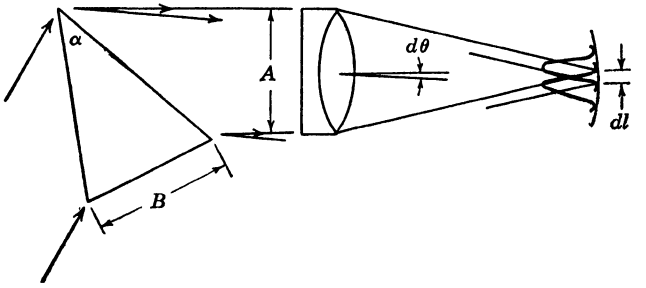


Fig. 4. Prism spectrograph.

Typical values for a prism with 5 cm faces, 60° angle, and 3 m focal length lens, $d\lambda/dl = 16$ Å/mm (1Å = 10^{-8} cm).

The resolving power is defined as $\lambda/\Delta\lambda$ where $\Delta\lambda$ is just resolved according to the Rayleigh criterion of diffraction pattern.

$$\frac{\lambda}{\Delta\lambda} \approx \frac{\lambda}{d\theta}\frac{d\theta}{d\lambda}$$

$$\frac{\lambda}{d\theta} = A \qquad \text{(from Rayleigh criterion)}$$

Thus

$$\frac{\lambda}{\Delta\lambda} = B\frac{dn}{d\lambda}$$

This is true for rectangular apertures, but actually the circular apertures of the lenses are often of smaller dimensions than the prism. The effective size of the prism then becomes smaller than the actual size and the laws of circular apertures must be used. This leads to the equation

$$\frac{\lambda}{\Delta\lambda} = \frac{B'}{1.22}\frac{dn}{d}$$

Fig. 5. Multiple prisms.

where B' is the effective base of the prism. For the previously mentioned dimensions of the instrument $\lambda/\Delta\lambda$ can approach the value 10,000.

To obtain high resolving power the prism must be large. Size is limited by inhomogeneity in the optical glass, which makes prisms larger than 10 cm base impractical. Other methods are used: a train of small prisms overcomes this difficulty (Fig. 5), but intensity is lost by

reflection from the surfaces. A large prism can be corrected by interferometer methods, correcting for the irregularities in the fringes by curving the surface. Correcting and testing every step makes the construction of such a prism very difficult.

MINIMUM DEVIATION. At minimum deviation the prism has the property of forming a stigmatic image for any beam. The prism, however, is always stigmatic with respect to parallel beams, whether at minimum deviation or not. Therefore in the case of a chromatic col-

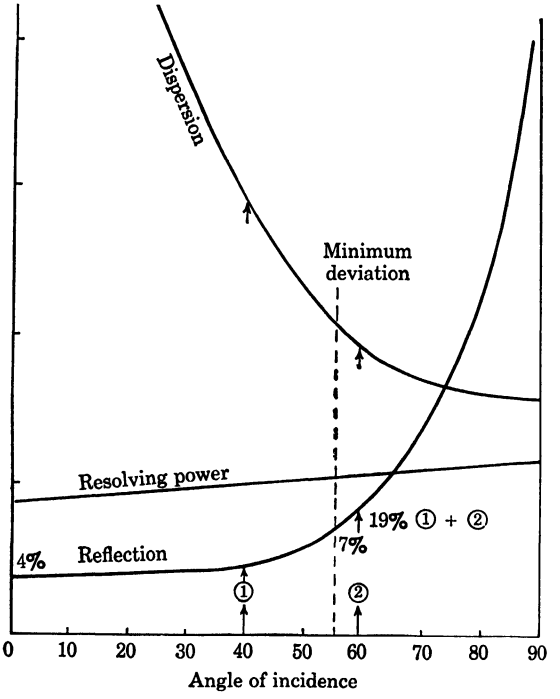


Fig. 6. Characteristics of prisms.

limating lens, it is necessary to operate at minimum deviation for lack of complete parallelism.[2]

The effect of the change of the angle of incidence on the resulting image is best discussed by help of the curve shown in Fig. 6. The dispersion, the resolving power, and the percentage of reflected rays from the surfaces are plotted as functions of angles of incidence. The resolv-

[2] By astigmatic is meant that the deviated rays, instead of meeting at a point forming a virtual image of the object, meet in two segments of straight lines at different distances from the prism. Halfway in between, the virtual rays form a circle as the best image of the point source.

ing power is nearly constant over the applicable angles. The dispersion falls off rapidly, while the reflection rises at high angles. A small angle of incidence on the first surface, for example 40°, will give a high dispersion, but the reflection from the second surface will correspond to an angle of 58° and hence to about 19% loss by reflection. Minimum deviation angle is shown on the graph, which gives a compromise between gain in dispersion and loss by reflection. However, for special purposes intensity may be sacrificed for high dispersion by using other angles of deviation.

Position of source. For maximum intensity the source should be as close to the slit as possible if all the beam entering is to be received by the lens. However, there is a limit beyond which no more gain is achieved as the beam starts to fall on the walls of the collimating tube. A lens focusing the source causes a loss of intensity due to reflection from the surfaces. Yet it is sometimes necessary to use the lens if the light source must necessarily be at some distance from the slit. In this case any focal length lens with the same aperture will give the same intensity, because as we gain in the solid angle subtended by the lens of shorter focal lens, we lose intensity due to formation of a larger image at the slit; the two effects compensate.

Grating instruments. For grating instruments all the facts and construction details of the collimation of light and the receiving cameras discussed in the section about prism instruments are valid. Characteristic features, however, reveal themselves when one considers the calculations of the dispersion and resolving power.

$$\text{dispersion} = \frac{d\theta}{d\lambda} = \frac{m}{d \cos \theta}$$

as is well known from the grating formula

$$d \sin \theta = m\lambda$$

where d is the spacing of grating lines and m is the order of the spectrum.

As the width of the beam is

$$A = Nd \cos \theta$$

where N = total number of rulings, and resolving power

$$R = A \, d\theta/d\lambda$$

Therefore

$$R = Nd \cos \theta \frac{m}{d \cos \theta} = mN$$

which depends on the total number of rulings and the order of spectrum observed. Typical values for 15,000 lines per inch, 6 in. grating, 3 m lens are

$$\frac{d\lambda}{dx} = 52 \text{ A/cm} \quad \text{or} \quad 5.2 \text{ A/mm}$$

which is much higher than the case of prism instruments. However, prisms can give higher dispersion in the 2000 A region.

CONCAVE GRATING. If rulings are made on the surface of a concave reflecting surface, the grating not only disperses the light but focuses it also, so that no lenses are necessary. The geometrical arrangement of

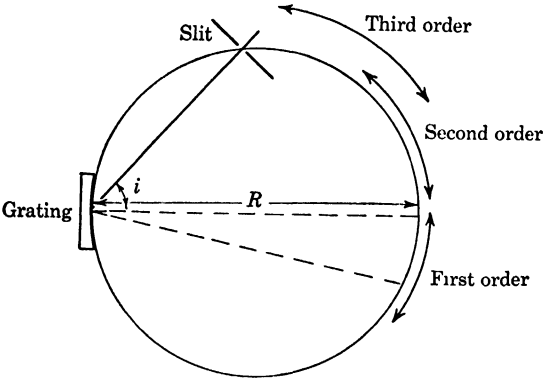


Fig. 7. Paschen mounting of concave grating.

such a concave grating is shown in Fig. 7. Let R be the radius of curvature of the surface of the grating. The large circle indicated on the drawing of radius $R/2$ is called the Rowland circle, and is drawn tangent to the surface of the grating. Under these conditions, if the slit is placed on the Rowland circle, the entire spectrum will be brought to focus on the same circle as shown.

A concave grating is astigmatic, i.e., the spectral lines, which should be simple images of the slit, are longer than the true image. The length increases as the angle of incidence increases. This eliminates the appearance of dust lines, which are caused by dust particles between the jaws of the slit.

SPECTROGRAPH MOUNTINGS. Several types of mountings have been in common use. Full description can be found in any textbook on experimental spectroscopy. Simple sketches will be given here with very short remarks.

1. *Prism mountings.* Three prism mountings are shown in Figs. 8, 9, 10. Figure 8 shows a Littrow mounting, 30° prism, platinised on surface, 3 m focal length of lens. Figure 9 shows a high-power Littrow prism, equivalent to 360° angle, usually quartz (for ultraviolet), R and

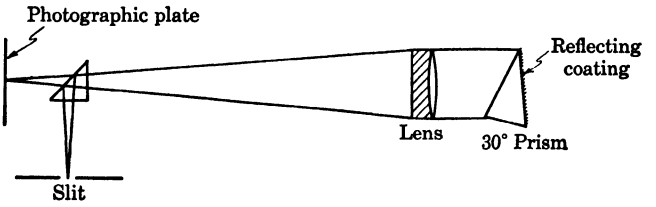


Fig. 8. Littrow mounting of prism.

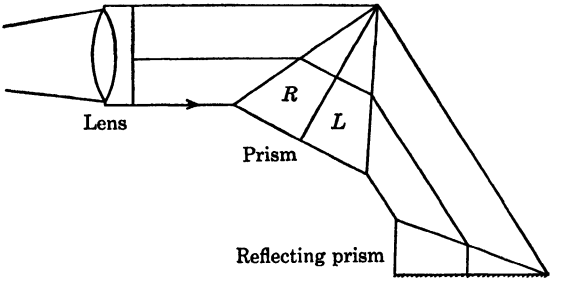


Fig. 9. 180° Littrow mounting for prism.

L for compensation. Figure 10 shows an Abbé prism, fixed arms, and movable prism; spectrum moves in observing telescope as prism is rotated about a center.

2. *Grating mountings.* Various mountings of gratings can also be made. The most accurate and versatile grating mounting is the Paschen

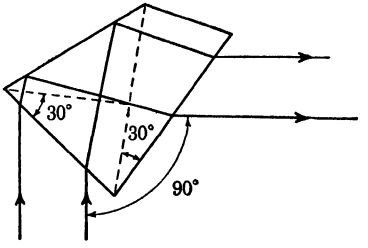


Fig. 10. Abbé prism.

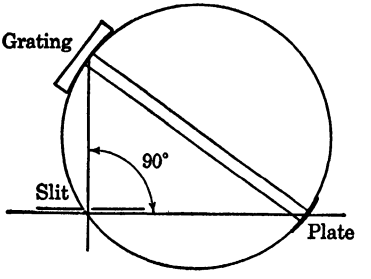


Fig. 11. Rowland mounting.

mounting, shown in Fig. 7. It has a concave grating and fixed parts with all orders of the spectrum in focus. When space is available, it is the best to use.

A Rowland mounting is shown in Fig. 11. If the center of the grating is on a circle with the slit and grating, the focus will be on the same

circle. By moving the arms and keeping the relation, one can get different order spectra.

Figure 12 shows a Littrow grating mounting, usually used by astronomers. If the focal length is large, the prism can be eliminated, putting the plate and source slightly off axis. The Mt. Wilson spectro-

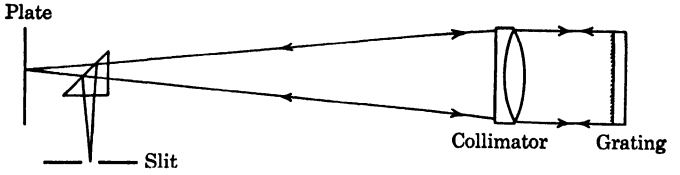


Fig. 12. Littrow mounting of plane grating.

graph has a 75 ft lens, which gives third-order visible spectrum 45 ft long, and fourth-order visible spectrum 60 ft long.

3. *Vacuum spectrograph mountings for ultraviolet.* The mounting shown in Fig. 13a is used from 2000 A to 200 A. Beyond that the reflecting power drops. A grazing angle mounting, shown in Fig. 13b, is used from 200 A to 20 A. The spectrum is not linear; overlapping occurs between different orders.

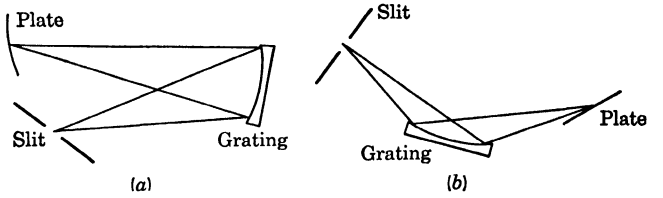


Fig. 13. Vacuum spectrograph mountings.

INTERFEROMETER INSTRUMENTS. Interference phenomena have been used already in the grating instruments. Other interferometers were developed with different arrangements and used either alone in spectrographs or as an addition providing higher resolution. The relation $F = nN_{\text{eff}}$ holds as a general relation for interferometers, where n is the order and N_{eff} is the number of interfering beams. While gratings are essentially instruments with high N and low n, on the other hand, other interferometers are usually constructed for high orders and low N.

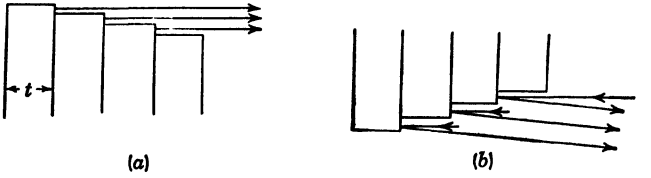


Fig. 14. Echelons: (a) transmission; (b) reflection.

Echelons are constructed both as transmission (Fig. 14a) and reflection (Fig. 14b) instruments. They are essentially a set of glass plates

exactly equal and parallel. They become, then, equivalent to a grating of very coarse ruling.

$$n = \frac{\mu t - t}{\lambda} = \frac{(\mu - 1)t}{\lambda}$$

in a typical case

$$n = \frac{(1.5 - 1)1}{5 \times 10^{-5}} = 10^4$$

For reflection instruments, the faces of the steps are silvered. In this case

$$n = \frac{2t}{\lambda}$$

One of the applications of the echelon in conjunction with a spectrograph is indicated in Fig. 15. The spectrograph separates the lines of widely different wavelengths that overlap on S_2.

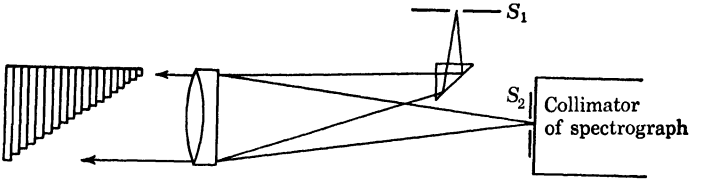


Fig. 15. Use of echelon with spectrograph.

Lummer plate depends on the principle of interference in multiple reflection from two parallel surfaces. It is also, as other interferometer instruments, used for the detailed study of one spectrum line. The glass plate is 10–20 cm long, 1 or 2 cm wide, a few mm thick, and has parallel surfaces to within a fraction of a wavelength. When the transmitted rays are brought to a focus, they give interference fringes (Fig. 16).

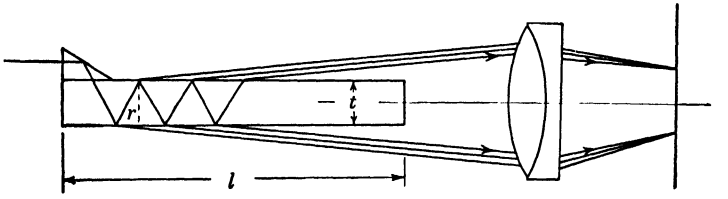


Fig. 16. Lummer plate.

If the thickness of the plate is t, then $n\lambda = 2\mu t \cos r$ and the number of effective beams $N_{\text{eff}} = l/2t \tan r$ (neglecting fall of intensity of successive beams).

The Fabry-Perot interferometer consists of two glass plates, each having exactly parallel surfaces, placed at distance t apart (Fig. 17a). For the central ray $n = 2t/\lambda$. If the inside surfaces are silvered to give a reflected fraction r,

$$N_{\text{eff}} = \frac{\pi\sqrt{r}}{1 - r}$$

For $r = 0.9$, $N = 30$. In some cases t is as high as 10 cm. The fringes then will show any hyperfine structure of the lines if the slit is made narrow enough to separate the individual lines and prevent overlapping

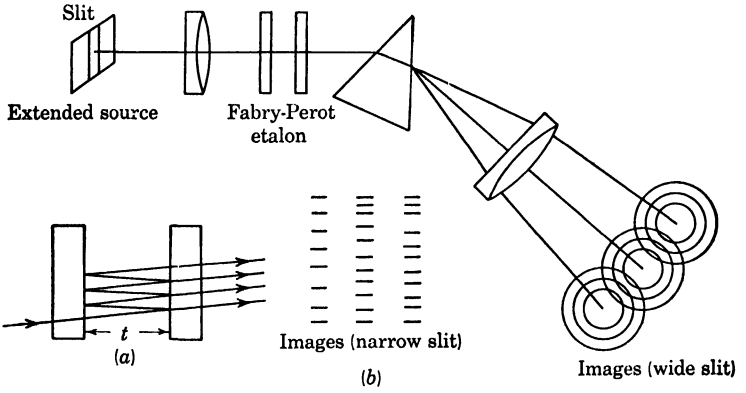


Fig. 17. Use of Fabry-Perot etalon.

of the fringes, as it will give complete circles if an extended source is used.

The position of the interferometer can be either as shown, next to the prism, or after the objective of the camera, selecting only one line.

WAVELENGTH MEASUREMENTS AND IDENTIFICATION OF LINES

STANDARDS. The *primary standard* is the red cadmium line. This line was chosen as a standard because it appears as a single line even in the highest resolutions. The wavelength was measured by absolute methods to accuracy of about 1 part in 10,000,000 through the work of the Bureau of Standards.

$$\lambda = 6438.4696 \pm 0.0005 \text{ A}$$

This is specified to be in dry air at 15°C and 760 mm Hg with 0.03% carbon dioxide content. As a matter of fact, the angstrom now, instead of being defined as 10^{-10} meter, is defined as 1/6438.4696 of the wavelength of the cadmium red line. The radiation is produced in a vacuum tube electrically.

However, there is a tendency to use another standard that is known to be absolutely free from hyperfine structure. This is the Hg^{198} obtained by bombarding gold in an ion-accelerating machine.

A few *secondary standards* have been measured in comparison with the primary standard to be reproduced in the laboratories. Usually iron, but sometimes nickel, copper, or neon has been used. They are tabulated in the *International Critical Tables*, Vol. 5, p. 275, with accuracy up to 0.001 A.

Iron arcs are usually used because they have been obtained in pure samples and give many lines, making interpolation easy. The standard arc is made of pure Swedish or Norwegian soft iron, constructed as shown (Fig. 18), and run at 220 v with a ballast resistance of about 50 ohms. If a bead of iron oxide can be collected on the top of the positive pole, it stabilizes the arc and keeps the hot point from moving around too much on the negative pole. The resistance is then adjusted with the arc to give from 4 to 5 amp. The light is taken from the middle region of the arc.

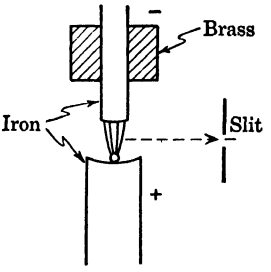


Fig. 18. Iron arc.

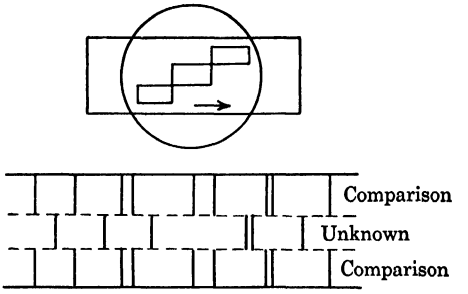


Fig. 19. Method of obtaining comparison spectra.

LINE IDENTIFICATION. The unknown spectrum is photographed between two sets of the standard spectrum. This is done by a mask with three slits covering the collimator slit as shown (Fig. 19). This prevents any shift that might be produced if we move the plate. For further precaution the condenser lens is also left unmoved when we interchange arcs. Then we obtain a set of 3 spectra, two of the standard enclosing the unknown. The slits are cut so as to make the lines slightly overlapping. The positions are then measured by a comparator. Comparators used in spectroscopic measurements usually read to 0.01 mm and can be interpolated to 0.001 mm with a total travel of 5 to 20 cm. From the readings of the comparator for the standard and the unknown together with the tables and charts of the iron spectrum, the wave-

lengths can be found.[3,4] The *Atlas*[4] gives the arc, spark, flame, and vacuum tube spectra for all known elements. Otherwise, if we know only the wavelengths of the standard, we can deduce the unknowns by essentially an interpolation method.

INTERPOLATION METHODS. A chart of the comparator readings x of the standard lines against the wavelength should determine the λ corresponding to any value of x of the unknown spectrum. However, a close investigation shows that such a graph is valueless. If $x = 0.01$ mm is the smallest division, and if we want this to be represented by 1 mm on the graph for full utilization of our readings, 1 mm of plate distance reads 10 cm of graph paper, which makes the size of paper very impractical. For this reason different methods have been adopted.

1. *Linear interpolation.* Assuming that $\lambda = \lambda_0 + c(x - x_0)$ holds for a small region, c can be determined as

$$c = \frac{\lambda_1 - \lambda_0}{x_1 - x_0}$$

if the two are sufficiently close. The relation can then be used for any line that lies close enough to them to justify the approximation.

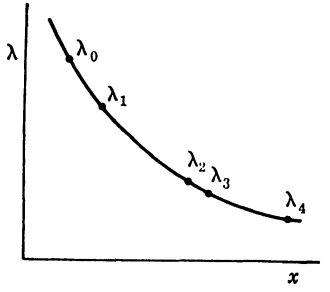


Fig. 20. Wavelength interpolation.

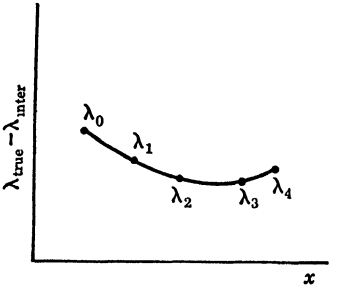


Fig. 21. Correction curve.

2. If λ_1 and λ_4 are not very close lines, we still can apply linear interpolation with correction factors. The values of λ_0, λ_1, λ_2, . . . are calculated from the interpolation formula and compared with standard values. A curve of the error ($\lambda_{true} - \lambda_{inter}$) is plotted against x, which can be used to correct any value of λ calculated from the interpolation formula.

[3] Brode, W. R., *Chemical Spectroscopy.* New York: John Wiley & Sons, Inc., 1943.

[4] Eder, J. M. and Valenta, E., *Atlas Typicher Spektrem.* Vienna: Akademie der Wissenschafter, 1911.

3. It has been found that the values of $\nu = 1/\lambda$ measured with prism instruments go quite linearly with the distance x. This is because

$$\frac{d\nu}{dx} = \frac{1}{\lambda^2}\frac{d\lambda}{dx}$$

and $d\lambda/dx$ decreases toward the violet. Since λ also decreases toward the violet $d\nu/dx$ remains almost constant. Interpolations can be made or a correction curve can be plotted, which in this case will be very flat (Fig. 22).

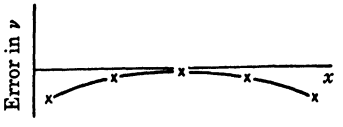


Fig. 22. Correction curve.

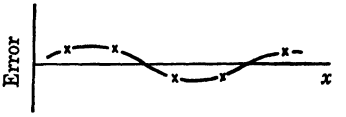


Fig. 23. Correction curve.

4. Hartel formula. An empirical formula

$$\lambda = \lambda_0 + \frac{c}{x - x_0}$$

fits very well for prism instrument data, over the whole visible spectrum. If a correction curve is plotted it will be very flat and fluctuating between negative and positive values (Fig. 23).

LIGHT SOURCES AND SENSITIVE PLATES

Different methods of excitation, different physical conditions in which a material is excited, and different requirements that a light source has to meet all interfere and affect the choice of the light source and the resulting spectrum. The general classification can be made according to four properties with subdivisions as follows.

1. By the method of excitation: thermal, arc, discharge tube, spark.
2. By the type of spectrum: continuous, line, band.
3. By the spectral region: visible, ultraviolet, infrared.
4. By the desired characteristics: high intensity, sharp lines, small quantities, controlled conditions.

These divisions are not sharp by any means and always overlap. The discussion here follows the fourth classification, mentioning the properties of the particular sources.

High intensities are convenient and may be essential in some cases. However, sometimes they are incompatible with other requirements of the three other subdivisions. Methods have been developed to produce high-intensity sources by different types of excitation, giving different characteristic spectra in each case.

Line spectra can be produced by:

1. *The direct-current arc.* Arcs can run on 110–220 v. Actually across the arc itself, only from 20 to 90 v is used, with currents of 2 to 5 amp, while the rest of the voltage is dissipated across the balancing resistor. This resistor is essential with arcs, as the arc has a negative characteristic (voltage drops as current increases). Also, a large air core inductor can stabilize the arc, because if the current drops, the induced voltage will compensate and keep the current going, reducing the fluctuations.

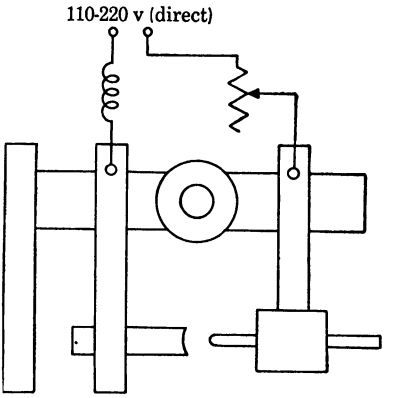


Fig. 24. Iron arc circuit.

The hot spot produced on the cathode causes thermionic emission and thus keeps the arc running. However, this hot spot shifts its position on the cathode and thus it is convenient to have a horizontal adjustment to focus the arc on the slit. If an arc spectrum is required for materials available in small quantities of salts, a carbon arc is constructed with the salt imbedded in a hole in the anode; the arc will vaporize the material and give its characteristic spectrum. Cooling will be necessary in two cases: if the arc is enclosed in glass as in case of Hg or Na arcs, or if sharp lines are required.

2. *Spark sources.* Sparking gaps give high brightness, stable characteristics, and small consumption of material. However, spark lines are broader than arc lines. The spark is usually started by about 12,000 v/cm in air. A capacitor is necessary to store enough charge to feed the spark. To suppress air lines produced in the spark, an additional choke coil is connected in series with the spark. A sparking gap can be run in vacuum by the use of higher voltage, bigger capacitor, and a secondary gap in air to help the discharge (Fig. 25).

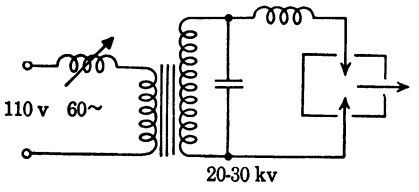


Fig. 25. Spark source circuit.

3. *Vacuum tubes* are used in many cases for the advantages they have over other excitation methods. They can be easily used for gases and can even be excited by induction from outside the glass, which avoids any contamination of the gas. For a bright spectrum, Geissler tubes, with constriction between the electrodes are used. If one electrode is replaced by a hot filament, lower voltage direct current can be

used. External excitation can use either electromagnetic or electrostatic radio frequency exciters, depending on the shape of the tube and the supply used. Various methods of excitation are shown in Fig. 26.

In internal electrode tubes, one is always faced with difficulties from sputtering. Metal evaporates from the electrodes and condenses on the glass. The best kind of electrode from this point of view has proved to be the aluminum electrode.

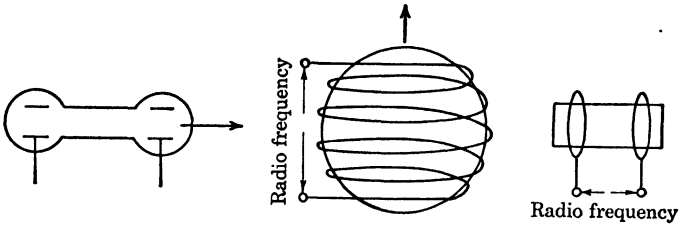


Fig. 26. Vacuum tube sources.

CONTINUOUS SPECTRA. The typical example of a continuous spectrum source is a hot tungsten filament. However, sometimes brighter sources are required and hence different methods have been developed.

The old-fashioned carbon arc is probably the brightest source in use. If a small size of source is needed, the Western Union concentrated arc gives an illuminated spot of 0.003 in. diameter for the 2 w arc. It consists of two electrodes as shown (Fig. 27a), enclosed in a bulb with low-

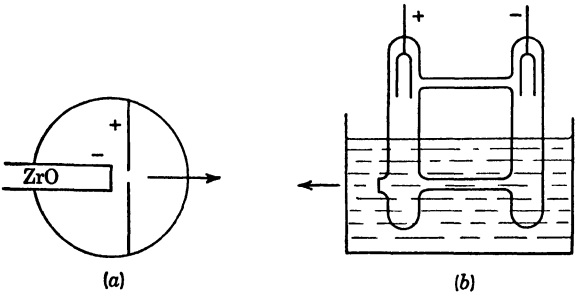


Fig. 27. (a) Point source. (b) Water-cooled continuous source.

pressure argon. Another type of concentrated arc using helium at low pressure gives a continuum in the ultraviolet region when discharged.

A hydrogen continuum source can be made with water cooling and high intensity. It has electrodes made of 5 mm inside diameter aluminum. The electrodes are placed in separate compartments connected to each other by two capillaries. The whole apparatus is placed in a water tank (Fig. 27b). It utilizes 5000 or 3000 v at 1 amp, and the pressure of the hydrogen is about 1 cm Hg. If helium is used instead of hydrogen, the spectrum extends to 1100 A.

SHARP LINES. There are several causes of wide spectral lines. The Doppler broadening is caused by the random thermal movements of the radiating atoms, and so can be reduced by decreasing the temperature of the source. High pressures also tend to broaden the lines due to the interactions between atoms during collisions. A sharp line source then must satisfy the three conditions: low current, low temperature, and low pressure.

1. *Water-cooled mercury arc.* Perhaps the simplest is the water-cooled mercury arc. The arc strikes between a metallic electrode and mercury in a small narrow vertical tube. The tube is immersed in a water cooling tank. The performance is not very satisfactory because of self-reversal. The lines are produced with considerable width in the relatively hotter inside region, and on the way out through the colder part a sharp absorption line appears.

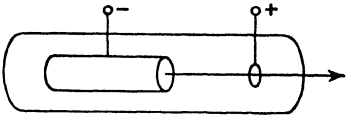


Fig. 28. Hollow cathode tube.

2. *Hollow-cathode tubes.*[5,6] If a tube is constructed as in Fig. 28 with the cathode formed of a hollow cylinder, it is found that as the pressure is lowered the glow discharge takes place *inside* the cathode. High excitation energies can be obtained with this type of source. If it is necessary to cool the cathode by liquid air, the tube is constructed as in Fig. 29, with a hollow aluminum cathode. This type of tube (Schuler tube) is very useful in connection with very small quantities of material where hyperfine structure is to be investigated.

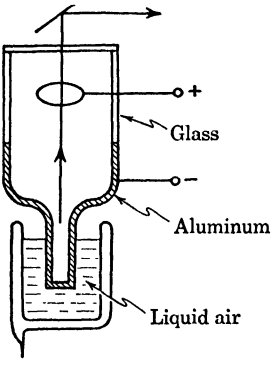


Fig. 29. Schuler tube.

SMALL QUANTITIES. The problem of producing spectra of materials of which only small quantities are available is a critical one, especially now when very small quantities of isotopes are sometimes produced by atomic bombardment and must be tested. Several methods are used, including some of the old ones. The material can be placed on top of a carbon arc and the lines observed. If the carbon is very pure, this method can make quantities as small as 10^{-8} or 10^{-9} gm detectable. The Lundegardh flame has been used for some time. An atomized solution is introduced in an oxyacetylene flame. The flame is made non-

[5] Harrison, G. R., *Practical Spectroscopy.* New York: Prentice-Hall, Inc., 1951.

[6] Jenkins, F. A. and White, H. E., *Fundamentals of Optics.* New York: McGraw-Hill Book Co., Inc., 1950.

luminous so that only the spectrum of the material will appear. Schuler tubes when constructed properly can detect quantities down to 15 μg. Radio frequency discharge tubes in helium can detect still smaller quantities.

Photographic plates are the detectors most commonly used in spectrography. They have the advantage of being integrating recorders (not like other photosensitive detectors), which makes them fit for detecting low intensities merely by increasing the time of exposure. However, choice of suitable plates and emulsion is essential.

For visible spectra and the near ultraviolet, commercial plates are suitable, unless very high sensitivity is required, which can be satisfied through some specially made plates such as IIaO of Eastman. Ordinary

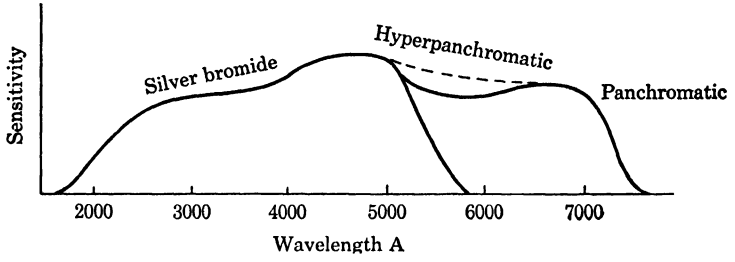


Fig. 30. Characteristics of emulsions.

plates have silver bromide emulsion on one side of the gelatin or the glass, and red antihalation coating on the other side. This coating dissolves in processing. The regular silver bromide emulsion has a cutoff at about 5000 A, with a characteristic curve as shown (Fig. 30). Plates of this kind are available commercially from Eastman in different sensitivity and grain size. The more sensitive the plate, the coarser its grains are. Slow fine grain films are used for lantern slides or copying; while medium speed is available in Kodak 33 and fast, coarse grain in Kodak 40 plates.

The introduction of coloring matters to sensitize the plates has made it possible to make emulsion with almost any characteristic required. Panchromatic and hyperpanchromatic are the best known in the field of photography. Panchromatic extends the sensitivity to about 7000 A but still retains a dip at about 6000 A, while the hyperpanchromatic sensitivity curve is almost flat over the whole visible range.

Special emulsions are made by Eastman and named commercially by a number that goes from I to IV, where I stands for fast and IV for slow, and a letter from A to Z where Z has the sensitivity in the far infrared. The characteristic curve for each is quite different from the others, and can be obtained drawn to scale from the company.

BIBLIOGRAPHY

Baly, Edward Charles Cyril, *Spectroscopy*. London: Longmans-Green & Co., 1929.

Brode, W. R., *Chemical Spectroscopy*. New York: John Wiley & Sons, Inc., 1943.

Eder, J. M. and Valenta, E., *Atlas typicher Spektrem*. Vienna: Akademie der Wissenschafter, 1911.

Harrison, G. R., *Practical Spectroscopy*. New York: Prentice-Hall, Inc., 1948.

Jenkins, F. A. and White, H. E., *Fundamentals of Optics*. New York: McGraw-Hill Book Co., Inc., 1950.

Kuhn, H., "New Techniques in Optical Interferometry," *Reports on Progress in Physics*. London: The Physical Society, 1951.

Price, W. C., "Recent Advances in Ultra-Violet Absorption Spectroscopy," *Reports on Progress in Physics*. London: The Physical Society, 1951.

Sawyer, R. A., *Experimental Spectroscopy*. New York: Prentice-Hall, Inc., 1951.

Williams, W. Ewart, *Applications of Interferometers*. London: Methuen & Co., Ltd., 1930.

25

MICROWAVE SPECTROSCOPY

INTRODUCTION. Microwave spectroscopy is similar to optical spectroscopy in that it measures the energy changes between quantum states in atoms and molecules. The microwave region extends from perhaps 20 cm down to a presently attainable 3 mm. This is the region where wave-guide components and cavity oscillators are effective and of reasonable dimensions. The absorption lines of molecules are generally stronger and more abundant in the shorter wavelength regions. Therefore the millimeter region is of most interest.

WAVE-GUIDE COMPONENTS. Many of the instruments and components used in microwave spectroscopy are commercially available because they and their associated techniques were developed for microwave radar. There are excellent sources of information in this field.

Perhaps the basic component is the wave guide. It is usually rectangular in cross section. Several sizes are required to cover the region from 3 mm to 20 cm. These are designated by the H, I, J, K, X, or S band covered. Their inside dimensions range from 0.110 by 0.050 in. for the H to about 3 by 1.5 in. for the S band. The cutoff wavelength λ_c for the lowest mode (electric field perpendicular to the widest dimension) in a rectangular guide is equal to twice the widest dimension. All higher modes cut off at shorter wavelengths. The I^2R power attenuation increases with wavelength for a given size guide and decreases with increase in guide dimensions for a given wavelength. The loss for 5 mm waves in coin silver guides ranges from about 0.65 db/ft for J-band to 0.75 db/ft for S-band guides. Brass used in the larger sizes will double the latter figure.

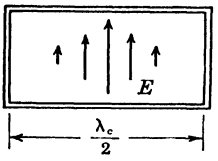


Fig. 1. Field in waveguide.

GAS ABSORPTION CELLS. Gas cells are usually made from long sections (5 to 120 ft) of rectangular wave guide sealed at the ends with thin mica windows to contain the gas while still permitting microwave energy to enter the cell. The longer lengths permit more nearly complete absorption of power by the gas and hence greater percentage power fluctuations when the frequency passes through an absorption line. On the other hand, I^2R guide losses also increase, thus decreasing

the over-all power level at the output. The longer lengths may be coiled to conserve space. It is helpful to use oversized guide to reduce I^2R losses and increase the volume of gas contained. This may be done because no large discontinuities are introduced that would excite unwanted higher order modes. A gradually tapered "transformer" section can be used to match the larger gas cell guide to smaller guide without objectional reflections. Smaller guide is used in other components to prevent unwanted higher modes.

When the Stark-effect modulation is used, the gas cell is modified by placing a brass electrode strip down the center of the guide. The strip is supported at its edges by polystyrene or teflon insulation. Teflon is better than polystyrene because of its lower dielectric loss and considerable chemical inertness. It is essential that the electrode be accurately centered within 0.001 in. or so.

A high-Q resonant cavity is sometimes used as a gas cell. Because of its small size it is particularly useful where only small quantities of gas are available. It is also useful for studies of the Zeeman effect because magnetic fields may be easily applied.

Microwave power sources. The source of microwave power used is the reflex klystron oscillator, which operates with a velocity-modulated electron stream. There are several variations on the market. Such klystrons deliver from 5 to 40 mw of power. The low power is no disadvantage because of the high sensitivity of crystal detectors available and because of the saturation effect of the molecules when more power is used. A valuable feature is the ease with which they can be frequency-modulated electrically at any rate over any range from a few kilocycles to 30 or 40 megacycles by varying the reflector voltage. They can also be mechanically tuned through a frequency change of from 10 to 45%. Klystrons are available to cover the regions from 5.7 mm to 1.6 cm (available from Raytheon Mfg. Co.) and from 2.6 cm to above 15 cm. The region below 5.7 mm is covered by using the nonlinear characteristic of crystal diodes to generate second harmonics down to 3 mm. The harmonic output is of the order of 15 μw. This is adequate for a video receiver or an evacuated thermocouple.

Saturation effects. Too much power into a gas absorption cell may actually reduce the absorption coefficient of the gas. This reduction is the result of molecular saturation, i.e., the situation where molecules, in absorbing quanta, are raised from the ground state to excited states faster than they can lose this energy, thus decreasing the number in the ground state available for absorbing power. The power at which saturation effects occur decreases as the square of the gas pressure, but increasing the pressure results in collision broadening of

the spectral lines because of very rapid induced transitions between upper and lower states. This also decreases the height of the absorption line of resonances. The best results are obtained by using a large volume of gas at low pressure (1 mm Hg or less) and low radio frequency power (of the order of 1 mw). It is necessary to work with absorption coefficients down to 10^{-8} or 10^{-9}. If the radio frequency signal level is of the order of 1 mv at the crystal detector these absorption coefficients yield power decreases when passing through absorption lines of the order of 0.05 to 0.01 μv.

APPARATUS FOR DIRECT ABSORPTION MEASUREMENTS

A typical arrangement for absorption spectra is indicated in Fig. 2. Attenuators, tapered guide transformer sections, and other details have

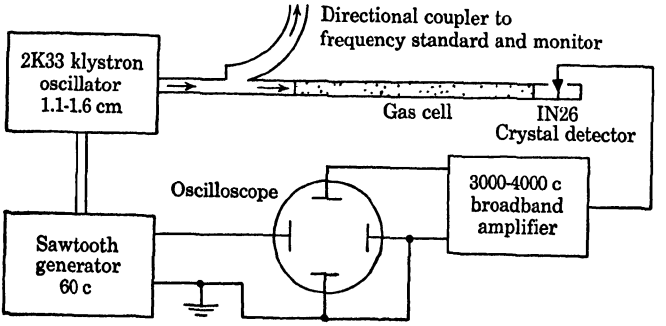


Fig. 2. Microwave spectroscope system.

not been indicated. The klystron is swept over perhaps a 10 megacycle band by the 60 c sawtooth voltage which also sweeps the oscilloscope trace horizontally. The center of the band may be shifted when desired by mechanical tuning. Whenever the radio frequency passes through an absorption line there is a decrease of power into the crystal detector. The rectified signal envelope is amplified and impressed upon the vertical sweep of the oscilloscope. Typical scope presentations are indicated in Fig. 3. Both the frequency and the relative height of absorption lines can be determined with this instrument. The resolving power is of the order 100 to 300 kc.

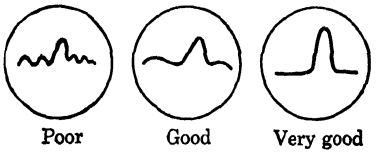


Fig. 3. Signal patterns on oscilloscope.

The height of the signal with respect to the background is limited by the strength of the absorption line, the signal-to-noise ratio, and reflections due to discontinuities in the wave-guide system. Since the signal-to-noise ratio for a given resistive impedance and temperature is inversely proportional to amplifier bandwidth, one might suppose that

a better ratio could be obtained by decreasing the sawtooth sweep rate and thus the amplifier bandwidth required. However, this is not the case, because the effective noise "temperature" of the silicon crystal detector increases with decrease in bandwidth at radio frequency excitations of more than a few microwatts.

Reflections in the wave guide result in standing waves whose mimina and maxima shift position as the klystron frequency is shifted. When these minima and maxima pass through the crystal detector's position they cause signal fluctuations that may be confused with absorption lines.

STARK-EFFECT MODULATION. A hundredfold increase in sensitivity may be obtained by using Stark-effect modulation. The circuit may be similar to that shown in Fig. 2 except that one uses a gas cell with an insulated brass electrode down its center (as described previously), and the amplifier is replaced by one of much narrower bandwidth, say 50 kc. A high direct voltage plus a small 50 kc alternating voltage is applied to the brass electrode. The direct voltage splits each absorption line into its Stark-effect components and the alternating voltage shifts these back and forth in frequency slightly at a 50 kc rate. It is this 50 kc modulation that is detected and amplified. The noise level is much reduced, and spurious lines caused by reflections in the wave-guide system are eliminated because no signal appears on the oscilloscope screen unless the gas molecules are undergoing Stark effect.

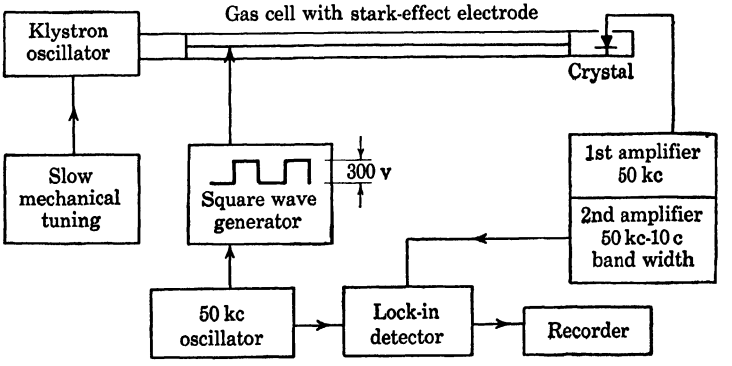


Fig. 4. Sensitive microwave spectrometer.

STARK-EFFECT MODULATION WITH LOCK-IN DETECTION. Another hundredfold gain in sensitivity can be obtained by use of Stark-effect modulation and a phase sensitive or lock-in detector (Fig. 4). The sensitivity is such that lines having absorption coefficients as low as 10^{-9} to 10^{-10} cm^{-1} may be readily detected. A high signal-to-noise ratio is obtained by using tubes hand picked for low noise in the first

50 kc narrow band amplifier and a twin T band-pass filter 100 c wide (10 c could be used with a 5 kc amplifier) to exclude noise in the second amplifier. The grid resistance to the first stage of amplification after the crystal is only 200 ohms in order to minimize noise generation. Direct current is used on the heater, and the grid is left floating (Fig. 5) so that it is self-biased and will have a minimum of noise pickup from the heater. Both the first and second 50 kc amplifier have gains of 1000.

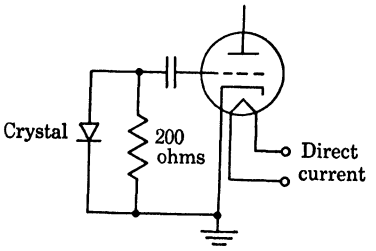


Fig. 5. Detector-amplifier circuit.

The lock-in detector requires signals both from the 50 kc oscillator and the amplifier that are exactly in phase and of the same frequency, i.e., coherent. Otherwise no signal is obtained. A typical circuit is shown in Fig. 6. When there is no signal to the grids both tubes conduct equally. Since the circuit is balanced, no current flows in the meter (or recorder) circuit. However, because the plate voltages are out of phase, a signal to the grids will be in phase with one plate while it is out of phase with the other, causing more and less plate current, respectively. This permits the phase (from the direction of current flow in the meter

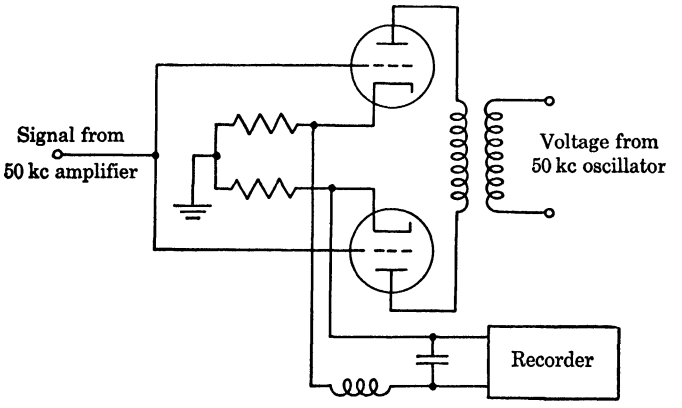


Fig. 6. Lock-in detector.

circuit) as well as the amplitude of the signal to be determined. This is important because with zero-based square-wave modulation of the gas cell the original absorption line is of one phase while the Stark-effect components are of the opposite phase. Thus the two may be readily distinguished. A typical spectrum from a lock-in detector is indicated in Fig. 7.

The long time constant of 10 sec (some investigators use 1 sec) filters out any beat frequencies from signal components that may be close to the 50 mc oscillator frequency. Such a narrow band width of 0.1 c requires a slow sweep for the klystron frequency. This is obtained by a motor drive that mechanically tunes the klystron (by bending its cavity resonator) at 500 megacycles per hour, or about 0.1 megacycle per second. The region that can be covered is 20,000 to 55,000 megacycles through a series of klystrons. This slow sweep and high sensitivity yield so much detail that it is laborious to analyze the spectrum. It is customary to use a less sensitive method to pick out the strong absorption lines and simple features of a spectrum and then to go to the lock-in detector with Stark-effect modulation and mechanical sweep to study detailed features.

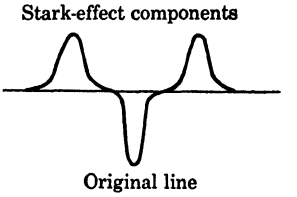


Fig. 7. Spectrum line from lock-in detector.

SQUARE-WAVE GENERATOR. The capacitance to ground of the Stark-effect electrode in a gas cell is of the order of 1000 $\mu\mu$f. It is required of the square-wave generator that it charge and maintain this capacitance at essentially constant voltage. If the rise and fall times are not rapid and the top of the square wave is not flat, the absorption lines will be undesirably broadened. Cathode followers are often used in the output

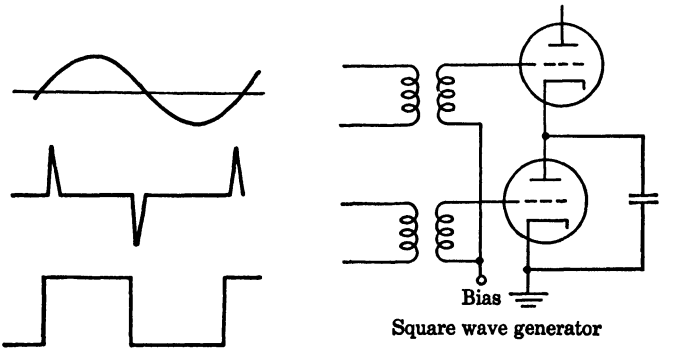


Fig. 8. Square wave generator and waveform.

to meet these requirements. If one uses a 100 kc square wave of 1000 v across a 1000 ohm load, this amounts to about 1 amp that the cathode follower must carry half the time. This necessitates high-power tubes and wastes power. A solution to these difficulties uses two tubes back-to-back as in the circuit shown in Fig. 8.

A 100 kc sine wave is put through a differentiating circuit and the sharp spikes are applied through pulse transformers to the grids of the

back-to-back tubes. The positive spike makes the upper tube conducting, causing it to "instantly" charge the capacitance load. The capacitance load remains charged until a half cycle later when the negative spike turns on the lower tube which then discharges the capacitance. The rest of the time both tubes are nonconducting. Thus little power is used, and fast rise and fall times are obtained because of the low circuit resistances. Tubes of relatively low plate dissipation are suitable, e.g., 807, 829, or four pairs of 715. Probably hydrogen thyratrons would also be suitable, but mercury-vapor thyratrons have too long a deionization time to be useful. With this circuit rise times of not over 5% are obtainable and the top of the square wave may be made flat within a few per cent.

FREQUENCY MEASUREMENTS. Precise frequency measurements are made by generating a series of marker signals of known frequency and then displaying them on the oscilloscope or pen and ink recorder along with the absorption lines being measured. A 5 megacycle thermostatically regulated crystal-controlled oscillator is used, which is compared against station WWV in Washington. WWV is accurate to 1 in 10^7. However, Doppler shift caused by changing path lengths from a varying Heaviside layer, etc., reduces the short time accuracy to perhaps 1 in 10^6. The 5 megacycles (see Fig. 9) are raised to 270 megacycles by

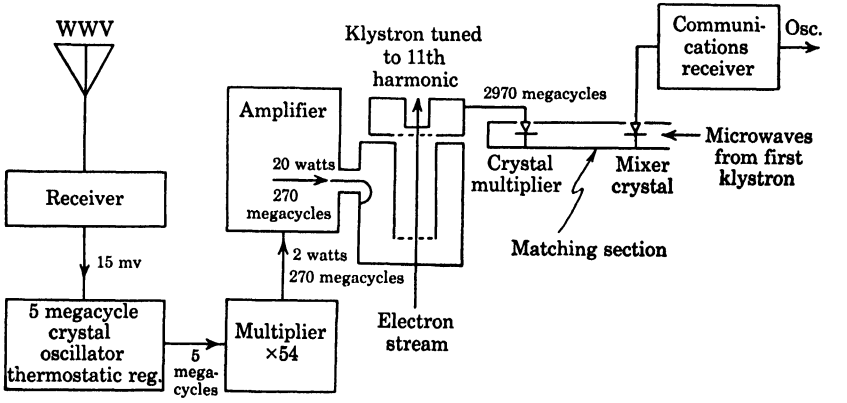


Fig. 9. Frequency calibration system.

the usual doublers and triplers. This is amplified and fed into a klystron whose output cavity is tuned to the eleventh harmonic to produce 2970 megacycles at an ample power level of 40 mw. The beam voltage on the klystron is adjustable and it should be adjusted for maximum power, i.e., so that electrons are bunched for the first time as they reach the output cavity.

A frequency of 2970 megacycles is equivalent to about a 10 cm wavelength. To obtain shorter wavelengths the klystron is fed into a silicon crystal multiplier. The crystal nonlinearity produces usable harmonic power up to the twentieth (5 mm waves) and higher. In order to obtain marker frequencies at intervals closer than 2970 megacycles it is necessary only to feed 270 megacycle and 90 megacycle power along with 2970 megacycle power into the crystal multiplier. The crystal's output will then contain all frequencies at 90 megacycle intervals starting with the wave guide cutoff frequency. These are fed through a matching section of wave guide to a second crystal, which mixes the reference (standard) frequencies and the time-varying frequency from the klystron (hereafter referred to as the first klystron) supplying microwaves to the gas absorption cell.

The output of the crystal mixer goes into the antenna input of a communications-type receiver that is tuned to beat frequencies between the reference and first klystron frequencies. The receiver output is fed to the vertical plates of the same oscilloscope upon which the gas absorption lines are displayed. If the receiver is tuned to, say, 10 megacycles, a marker pip will appear on the 'scope screen in a position corresponding to every first klystron frequency that is 10 megacycles either above or below a reference frequency. A cavity wavemeter, which absorbs some of the first klystron's power, thus producing a pip of its own on the 'scope, is used to determine to which of the reference frequencies the marker pips are close. The accuracy of a cavity wavemeter is about 30 megacycles, which is sufficient for this purpose, but of no value for actual absorption line measurements. One determines whether a marker pip is above or below a reference frequency by the direction it moves when the receiver tuning is changed. To measure an absorption-line frequency one tunes the receiver until a marker pip exactly coincides with the absorption-line pip and then subtracts the receiver frequency from or adds it to the closest reference frequency as determined by the cavity wavemeter.

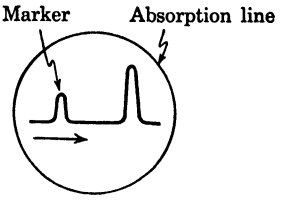


Fig. 10.

The above precise method is usually used to ascertain the position of strong absorption lines obtained by using a less sensitive detecting system, e.g., the arrangement of Fig. 2 either with or without Stark-effect modulation. The position of weak lines is determined with Stark-effect modulation, lock-in detection, and a pen and ink (chart) recorder. The marker pips (usually at 30 megacycle intervals) are "bulldozed" through the lock-in detector, i.e., they come through anyway, although on the face of it one would not expect them to do so. These marker pips

appear at regular intervals among the spectral lines and provide a calibrated scale by which frequency differences may be accurately determined. Since they provide no indication of absolute frequency, one must have recourse to previous measurements on strong lines for these.

BIBLIOGRAPHY

Coles, Donald K., "Microwave Spectroscopy," *Advances in Electronics*. New York: Academic Press, Inc., 1950, Vol. 2.

Gordy, Walter, "Microwave Spectroscopy," *Revs. Mod. Phys.*, *20*, 668 (1948).

Hamilton, D. R., Knipp, J. K., and Kuper, J. B. H., *Klystrons and Microwave Triodes*. New York: McGraw-Hill Book Co., Inc., 1948.

Lamb, Willis E., Jr., "Theory of a Microwave Spectroscope," *Phys. Rev.*, *70*, 308 (1946).

McAfee, K. B., Jr., Hughes, R. H., and Wilson, E. B., Jr., "A Stark-Effect Microwave Spectrograph of High Sensitivity," *Rev. Sci. Instr.*, *20*, 821 (1949).

Pound, R. V., "Electronic Frequency Stabilization of Microwave Oscillators," *Rev. Sci. Instr.*, *17*, 490 (1946).

Sharbauch, A. Harry, "The Design and Construction of a Stark-Modulation Spectrograph," *Rev. Sci. Instr.*, *21*, 120 (1950).

Unterberger, Robert and Smith, William V., "A Microwave Secondary Frequency Standard," *Rev. Sci. Instr.*, *19*, 580 (1948).

26

MOLECULAR BEAMS

Molecular beams provide a direct means of examining the properties and behavior of individual atoms and nuclei under the action of fields. They have also been used to verify statistical predictions of kinetic theory and space quantization of quantum mechanics. Here an atomic beam is considered a particular type of molecular beam, and all discussions apply to both. Molecular beams usually correspond to thermal beams with velocities of the order of 10^4 to 10^5 cm/sec.

A molecular-beam apparatus is comprised of four basic elements: a source, a collimating system, a detector, and a surrounding vacuum envelope.

MOLECULAR-BEAM SOURCES. The source generally consists of a vessel containing a gas or vapor and provided with an aperture through which the molecules effuse. The term "oven" refers to the source of a beam even when the vessel is not heated, as for permanent gases, or when the vessel is cooled. The general conditions for beam formation are the same for gases and vapors.

The total rate of effusion of a gas into a vacuum from an ideal (infinitely thin-walled) aperture is given by kinetic theory.

$$N = n\bar{v}a/4 \quad \text{molecules/sec} \tag{1}$$

where n = number of molecules per unit volume, a = aperture area, $\bar{v}$ = average velocity. In terms of pressure and temperature, using $p = nkT$ and $\bar{v} = \sqrt{8kT/\pi m}$ this may be expressed as

$$N = \frac{pa}{\sqrt{2\pi mkT}} \quad \text{sec}^{-1} \tag{2}$$

where m = mass of molecule, T = absolute temperature, k = Boltzmann's constant. The angular distribution of the molecules is

$$\frac{dN}{d\Omega} = \frac{pa \cos\theta}{\pi\sqrt{2\pi mkT}}$$

where θ is the angle between the solid angle $d\Omega$ considered and the normal to the aperture. These equations assume that no collisions take place in or beyond the aperture.

This places the restriction that $d < \lambda$ on the dimensions of the aperture relative to the mean free path λ of the particles at the pressure used in the oven, where d is the width of the slit. Slit widths of the order of 10^{-2} to 10^{-3} mm are generally used. If we require that the ratio λ/d be a certain value K, the maximum value of the pressure is determined by the slit width since $p \propto 1/\lambda$; therefore $p = \text{const.}/Kd$. Also, since the intensity is proportional to the product pd, we see that we cannot improve the intensity by manipulating the pressure and the width of the slit beyond a certain point without violating the $\lambda > d$ condition. It was pointed out by Stern, however, that the length of the slit could be increased beyond this condition if we were interested only in displacement normal to the slit length, since collisions in the direction of the slit length would not effect the results. This permits a considerable increase in intensity. If the condition that $\lambda > d$ is not fulfilled, the flow through the slits approaches the conditions for viscous flow, and a "cloud" tends to form on the detector side of the slits. This will tend to diffuse the beam if the oven slits are a part of the collimating system, and hence decrease resolution. If a separate collimating system is being used, there is no objection to viscous flow other than the possibility of an undesirable increase in base pressure of the envelope containing the beam.

For the slit widths quoted above, it is necessary to have oven pressures of the order to 10^{-1} mm Hg or less. Molecular beams have been produced with substances ranging from those for which this pressure requires temperatures higher than 1000°C to permanent gases where the oven may be cooled with liquid air. Consequently, the design of the ovens depends upon the nature of the substance to be investigated. For low temperatures, copper and glass have been used; for higher temperatures, steel, molybdenum, and tantalum have been used. The oven must not react or alloy with the substance. The slit jaws must be kept hotter than the rest of the system for condensable vapors in order to prevent clogging the slit jaws.

The equations for beam intensity were derived on the assumption of ideal slits. In actual practice, the slit walls are generally thick, and the flow is through a canal. This causes the total rate of effusion to be decreased, but the angular distribution (similar to a higher power of the cosine) is such that most of the decrease in intensity is at wide angles. This is an advantage, since only the small angle particles are usually used and the others must be pumped out to keep the background pressure at a minimum. This throttling action of a canal is used to advantage in collimating systems and manometer detectors as will be discussed later.

COLLIMATING SYSTEMS. The simplest collimating system is a single slit, which together with the source slit defines a beam (Fig. 1). Since the oven may be heated or cooled the alignment of the slits in this single system is difficult to maintain. A better beam definition is obtained with an auxiliary slit called a fore-slit (Fig. 2). The fore-slit remains at room temperature and hence remains in alignment in use. The fore-slit also eliminates cloud formation at the mouth of the oven slit, and the latter may be widened when used in conjunction with a fore-slit as

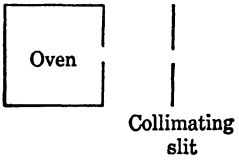


Fig. 1. Oven and slit.

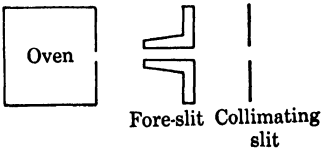


Fig. 2. Use of fore slit.

indicated previously, or the pressure may be increased in the oven, giving a corresponding increase in intensity.

Use of a fore-slit has another strong advantage. The molecular beam is scattered by any residual gases in the apparatus, and consequently it is desirable to keep residual pressures as low as possible over most of the length of the beam. The fore-slit permits separation of the chamber between the part containing the oven and the rest of the system. The fore-slit may be placed close to the oven, and if the two chambers have separate pumps, the region of high pressure due to the non-usable fraction of molecules effusing from the oven may be restricted to be a small part of the total distance traveled by the molecules. If the fore-slits are in the form of long canals, this isolation of the two pressure levels is enhanced further without any decrease in intensity of the usable portion of the beam, since the canal increases the flow impedance for all angles except those comprising the usable beam. The intensity at the detector is inversely proportional to the area of the collimators; consequently these must be designed for each experiment, using as large an area as is compatible with the deflection to be expected from the fields imposed on the beam. The problem involved is identical with that encountered in optical experiments, where intensity and resolution have an inverse relation, and the intensity that is permissible is governed by the magnitude of the separation to be resolved. This may be stated in terms of the height of the slit for the case of an inhomogeneous magnetic field. The magnitude of the inhomogeneity is inversely proportional to its height. Since the deflection is inversely proportional to the magnitude of the inhomogeneity and directly to its length squared (see p. 260), we get a deflection $\delta \approx L^2/h$, but assuming h is also the

width of beam and collimator, we get an intensity $I \approx h/L^2$. Thus $\delta \propto 1/I$. As an example, when Stern and Gerlach sought to verify the space quantization of the angular momentum of the silver electron structure, the deflections to be expected necessitated such small slits that an 8 hr run on a condensation target did not yield a visible trace.

DETECTORS. Most work with beams has been done using condensation targets, surface ionization detectors, and manometers. Chemical targets that react with the beam to give a colored trace, thermal detectors that measure the energy transferred to the target by the beam, and space charge detectors are mentioned in the literature, but will not be discussed here. The intensity to be expected at the detector may be calculated by multiplying (3) for the angular distribution per unit solid angle of the emitted particles by the solid angle a'/r^2 subtended by the detector. Setting $\cos\theta \approx 1$ ($\theta \approx 0$) we get for the number incident on the collector

$$N = \frac{pa \cos\theta\, a'}{\pi \sqrt{2\pi mkT}\, r^2} = \frac{paa'}{\pi r^2 \sqrt{2\pi mkT}} \quad \text{sec}^{-1} \tag{4}$$

where p = oven pressure, a = oven aperture, r = path length, T = oven temperature, a' = detector aperture.

CONDENSATION TARGETS. Condensation targets for beams of condensable vapors were used in the early experiments of Stern and others. This method is not very sensitive, and high-intensity beams or long runs are necessary to produce visible images. Techniques of developing a deposit invisible to the eye by chemical processes improve the method, but it remains more or less qualitative, as it is practically impossible to measure the quantity deposited and hence relative beam intensities.

SURFACE IONIZATION DETECTORS. If an atom impinges on a surface whose work function is greater than the ionization potential of the atom, it becomes ionized. Further, if the surface is hot, the ionized atom may be evaporated off and picked up on a negative collector. This is the basis of the surface ionization detector. In practice, a hot tungsten wire of a fraction of a millimeter diameter is used as a detector, and is surrounded by a negative electrode, which collects the positive ions. By using current amplifiers, beams of as few as 2500 atoms per second can be made to give a 1 mm galvanometer deflection. Hence this type of detector is extremely sensitive, gives immediate values, and has excellent geometrical characteristics for studying small deflections, since the wire may be made as small as 0.01 mm in diameter. The principal limitation of the surface ionization detector is that only a few

elements have sufficiently low ionization potentials for the method to be applicable. The alkali metals and their compounds can be detected by a pure tungsten detector, and a few others with an oxidized tungsten filament, which gives a work function of about 6 ev. Oxidized filaments cannot be run at too high a temperature, and this limits their use. For those elements to which this method is applicable, it is by far the most accurate and sensitive.

MANOMETER DETECTORS. For noncondensable vapors the beam can be detected by collecting the beam in a small chamber with a narrow slit opening. The pressure will rise in the chamber until equilibrium is reached between the particles effusing from the slit and the incoming particles of the beam. If we call p the pressure in the manometer, using (2) and (4) we get

$$N'_{\text{incident}} = \frac{p_0 a a'}{\pi r^2 \sqrt{2\pi m k T_0}} \qquad N_{\text{effusing from manometer}} = \frac{p_m a'}{\sqrt{2\pi m k T_m}}$$

Setting these equal at equilibrium in the manometer,

$$p_m = \frac{p_0 a}{\pi r^2} \sqrt{\frac{T_m}{T_0}}.$$

Hence, measuring p_m and T_m we have for the incident beam $N' = a' p_m / (2\pi m k T_m)^{\frac{1}{2}}$ at equilibrium. Since the partial pressure of the entire system at the detector is added to this pressure, it is necessary to compensate for the background. This is done by using another manometer identical with the first in the circuit but so oriented that the incident beam does not reach it. This may then be made to compensate automatically for the base pressure and fluctuation in the base pressure. The usual method of measuring the pressure in the small manometer chamber is by means of a Pirani gage. The pressures involved are very small for the use of a Pirani, and a canal entrance on the manometer increases the pressure by a factor of 10 over that just given, for the same incident intensity.

VACUUM SYSTEMS. The entire apparatus described so far must be maintained in a region of high vacuum. Since the beams are extremely narrow, any collision removes a particle from the beam and hence the beam falls off exponentially with distance; $I = I_0 e^{-r/\lambda}$ where λ is the mean free path under the existing background pressure in the instrument. It is desirable to make the path long to increase the resolution of small deflections. At the same time lack of intensity is also a constant problem, and hence the vacuum must be such that the mean free path is large compared with the length of the beam, to maintain good

intensity. As mentioned earlier, the use of a canal fore-slit permits separation of the higher pressure in the vicinity of the oven from the region containing the major portion of the beam path. For permanent gases, separate pumps may be used in each region. For the case of condensable vapors, no trouble is generally encountered, since cool surfaces have extremely high effective pumping speeds for vapors.

MOLECULAR-BEAM EXPERIMENTS. One of the earliest experiments with molecular beams was the direct measure of the velocity of silver atoms by Stern in 1920. A silver-plated platinum filament from which silver atoms were evaporated served as source, and the collimator slit and condensation target revolved around the source with a known angular velocity. Very short beams were used and deflections of the order of 1 mm were obtained by reversing the direction of rotation.

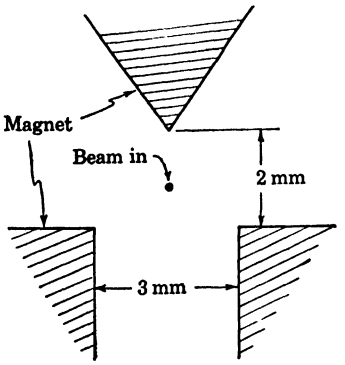


Fig. 3. Inhomogeneous magnetic field pole tips.

The experiment of Stern and Gerlach[1] gave the first direct demonstration of space quantization. A beam of silver atoms was passed through an inhomogeneous magnetic field produced by pole pieces shaped as shown in Fig. 3. If the atoms have a magnetic moment $\mu_B = e\hbar/2mc$, they experience a force $F = (\mu_B \cdot \nabla)H \approx \mu_H(dH/dz)$, where μ_H is the component of μ_B in the field direction z. According to quantum mechanics this component could take only certain discrete values given by $\mu_H = gm_j\mu_B$ where g is the Lande factor and m_j is the magnetic quantum number $m_j = J, J - 1, \ldots, -J$. For the Stern-Gerlach experiment the beam consisted of silver atoms and $J = \frac{1}{2}$, $g = 2$, $m_j = \frac{1}{2}$. Therefore the beam should split into two parts equidistant from the undeflected position. This was observed, and the right order of magnitude for the Bohr magneton was obtained. An example of the orders of magnitude involved in such cases may be given by a numerical example for the Stern-Gerlach experiment with potassium atoms. Here $dH/dz = 5 \times 10^4$ gausses/cm, length of field $L = 10$ cm, $\mu_B = 10^{-20}$ emu, $J = \frac{1}{2}$, $g = 2$, $M = 6.7 \times 10^{-23}$ g. For $T_0 = 450°\text{K}$ the most probable velocity is $v = 4.3 \times 10^4$ cm/sec. With these values the deflection at the end of the pole pieces is

$$\delta = \tfrac{1}{2}at^2 \approx \frac{\mu_B}{M} \cdot \frac{dH}{dz} \cdot \frac{L^2}{v^2} \approx 0.83 \quad \text{mm}$$

[1] Gerlach, W. and Stern, O., *Ann. Physik*, *74*, 673 (1924).

NUCLEAR MAGNETIC MOMENTS. The early experiments of Stern and his collaborators were concerned with *atomic* magnetic moments, due to the atomic electrons. The magnetic moments of nuclei, which are of the order of 1000 times smaller, can best be measured in the absence of the atomic magnet moments and with specially designed apparatus of high sensitivity. The presence of nuclear moments can, however, be detected as hyperfine structure on the atomic levels when apparatus of extremely high resolution is used. Stern and his co-workers used hydrogen molecules cooled by liquid air. Hydrogen molecules have no net electronic magnetic moment. By use of parahydrogen, which has no net nuclear moment, they evaluated the molecular rotational magnetic moments. Then in experiments with ordinary hydrogen they were able to interpret the deflections due to the double moment of orthohydrogen molecules. They obtained a value for the proton moment of about 2.5 nuclear magnetons.

Nuclear moments are also observable in hyperfine structure in Zeeman spectra. Quantum mechanics, however, places a limit on changes in energy which can be detected, as given by the uncertainty principle. From $\Delta E \Delta t \approx \hbar$, and with $\Delta t \approx 10^{-8}$ sec for the lifetime of optically excited states, we get 10^{-19} ergs as the smallest energy difference detectable. In the molecular beam the lifetime of the state is about 10^{-4} sec; thus energy differences of the order of 10^{-23} ergs should be resolvable by the molecular beam method.

Full use of this important advantage of molecular beams was made by Rabi and his collaborators[2] when they developed the magnetic resonance method of analysis of molecular beams. This method is based on the fact that angular momentum is associated with the magnetic moment. When the nucleus with its magnetic moment finds itself in a uniform magnetic field, it experiences a force tending to align it along the direction of the field. The angular momentum of the nucleus makes it act like a small gyroscope, however, and the resultant motion is a precession around the magnetic lines of force with the Larmor frequency. Because of the principles of space quantization, only certain orientations are allowed in the magnetic field. These different orientations possess different energies, and transitions between orientations can be made if radiation of the correct frequency (in this case radio frequency, and properly polarized so that it will interact with the magnet moments), is applied to the nuclei while they are in the steady magnetic field. If these transitions can be detected by some means, the magnetic moments can be obtained. This basic idea is also

[2] Rabi, I. I., Millman, S., Kusch, P., and Zacharias, J. R., "The Molecular Beam Resonance Method for Measuring Nuclear Magnetic Moments," *Phys. Rev., 55,* 526 (1939).

used in experiments on nuclear induction and nuclear magnetic resonance absorption, and is discussed in Chapter 27.

Rabi used a combination of the previously known techniques of deflection in inhomogeneous fields with the magnetic resonance absorption in a homogeneous field. The experimental arrangement is shown in Fig. 4. The source and slit arrangement produce a beam of atoms or molecules, which enter an inhomogeneous field produced by magnet A. They are deflected in A just as atoms were deflected in the original Stern-Gerlach apparatus. If there were no disturbance at C, the beam would enter magnet B, which has the magnetic field in the

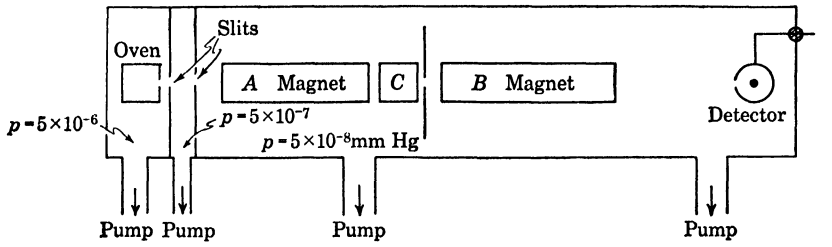


Fig. 4. Molecular beam resonance apparatus.

same direction as A, but the inhomogeneity reversed. Magnet B then cancels the deflection produced by magnet A, and the particles reach the detector almost as if no magnets were there. If, however, transitions are introduced at C by a combination of a steady magnetic field and a radio frequency source interacting with the magnetic moments, the beam can no longer be brought back in line by magnet B, and a dip in intensity is noticed at the detector. In principle the frequency or the magnetic field is varied until such a dip is obtained, and from a knowledge of the field and frequency at resonance, magnetic moments can be calculated.

Let us now consider in somewhat more detail the experimental apparatus required. The entire apparatus is enclosed in a vacuum chamber, and for proper operation it is important to have a vacuum of about 5×10^{-8} mm Hg in the main chamber. The differential pumping system indicated is necessary because of the large amount of gas given off near the source. The high vacuum in the main chamber is required because of the long path the molecules must traverse, where any collision with a residual gas atom would throw them out of the beam.

If the substance to be investigated is a solid it is vaporized in an oven, similar to one shown schematically in Fig. 5. The slit is usually 0.001 in. wide. Additional slits in front of the source (Fig. 4) define the

beam and provide a vacuum block so that differential pumping may be used. The final slit may be about 0.001 in. wide and 0.080 in. long.

After leaving the final slit the beam passes into the A magnet, which has the cross section and dimensions shown in Fig. 6. It produces an inhomogeneous magnetic field, which deflects the particles as in the original Stern apparatus. A geometric factor that expresses the degree of inhomogeneity is $(1/H)(\partial H/\partial y)$. In the field arranged as in

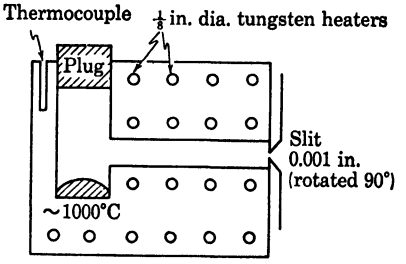


Fig. 5. Oven.

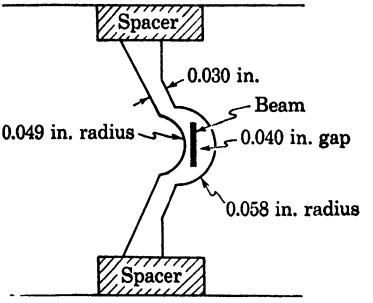


Fig. 6. Inhomogeneous magnetic field pole tips.

Fig. 6, $(1/H)(\partial H/\partial y) = 8$ cm^{-1}, and H is about 10,000 gausses. The deflection in this field amounts to about 0.001 in. It is to be noted that in the usual drawings of the apparatus the deflection is much exaggerated, and the small deflection points up the need for extremely accurate alignment of the magnetic field, slits, and detectors.

Magnet B is arranged so that the inhomogeneity is reversed, but the magnetic field is in the same direction as in magnet A. The magnetic

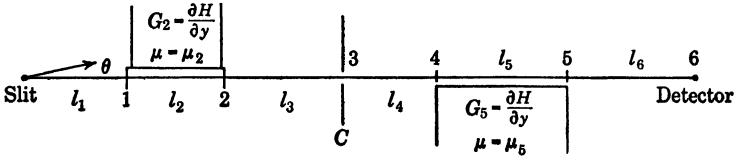


Fig. 7. Diagram of apparatus showing various lengths.

field throughout the apparatus must never reverse or even go to zero, as otherwise the magnetic moments lose their sense of direction.

In Fig. 7 a schematic diagram of the apparatus with path lengths for various parts of the beam is shown. Some of the design considerations for such an apparatus are as follows. Let y_p be the deflection at each point of the apparatus. Then

$$y_6 = (l_1 + l_2 + l_3 + l_4 + l_5 + l_6)\theta + \frac{1}{2}\frac{a_2}{v^2} l_2(l_2 + 2l_3 + 2l_4 + 2l_5 + 2l_6) + \frac{1}{2}\frac{a_5}{v^2} l_5(l_5 + 2l_6)$$

In this equation the l's represent the lengths of the various parts of the apparatus, v is the velocity of the particle, and

$$a_2 = \frac{\mu_2 G_2}{m}, \qquad a_5 = \frac{\mu_5 G_5}{m}, \qquad G = \frac{\partial H}{\partial y},$$

m = mass of particle

where μ is the magnetic moment at the particular point. It is clear that for maximum deflection G must be as large as possible. The various values of l are at our disposal. Suppose we set $y_3 = 0$, $y_6 = 0$, $\mu_2 = \mu_5$ (i.e., no disturbance at the C magnet), $G_2 = -G_5$ so that $a_2 = -a_5 = a$. Now $y_3 = (l_1 + l_2 + l_3)\theta + \frac{1}{2}\frac{a}{v^2} l_2(l_2 + 2l_3) = 0$. So, using the equation for y_6, which we have set equal to zero, we get the focus condition,

$$\frac{l_5(l_5 + 2l_6)}{l_4 + l_5 + l_6} = \frac{l_2(l_2 + 2l_1)}{l_1 + l_2 + l_3}$$

We notice that the focus condition is independent of v, the initial velocity, although the deflections occurring are not. Thus particles with different initial velocities have different trajectories but all curve back to focus at the detector. The focus condition indicates that a symmetrical apparatus can be used. Generally, however, it is more practical to make an unsymmetrical arrangement, using l_6 much smaller than l_1.

The apparatus is lined up optically as well as possible, then the beam itself is used for the final lining up process. After this is done, the C magnet is turned on and a radio frequency field is applied to cause transitions of the magnetic moments. The C magnet produces a homogeneous field of several thousand gausses. Placed in it is a hairpin shaped copper tube that carries the radio frequency current (Fig. 8)

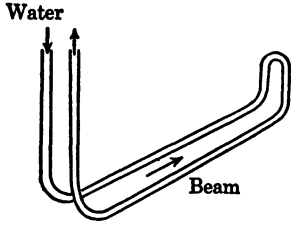


Fig. 8. Hairpin carrying radio frequency current.

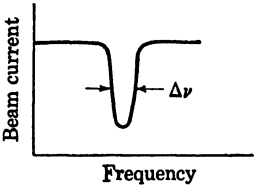


Fig. 9. Beam current near resonance.

and produces a peak radio frequency magnetic field of the order of 20 gausses. This requires a fairly high radio frequency current, and tubing is used so it may be water cooled.

As the frequency of the radio frequency field is changed, the beam current will change as a resonance is passed (Fig. 9). The accuracy of

this method is shown in a calculation of the width of the resonance curve. The width of the resonance is $\Delta\nu \approx 1/t$, where t is the time spent in the C field. For $v = 50{,}000$ cm/sec, and the length of the C field 10 cm, $\Delta\nu = 5000$ cycles, out of 5×10^6 cycles for a typical resonance. Thus the dips are very narrow and the method is capable of great precision.

Note on references: the articles by Estermann and by Kellogg and Millman are excellent review articles covering the material of this chapter in detail.

BIBLIOGRAPHY

Bessey, W. H. and Simpson, O. C., "Recent Work in Molecular Beams," *Chem. Revs., 30,* 239 (1942).

Esterman, I., "Molecular Beam Technique," *Revs. Mod. Phys., 18,* 300 (1946).

Fraser, R. G. J., *Molecular Beams.* London: Metheun & Co., Ltd., 1937.

Hamilton, D. R., "Molecular Beams and Nuclear Moments," *Am. J. Phys., 9,* 319 (1941).

Kellogg, J. B. M. and Millman, S., "The Molecular Beam Magnetic Resonance Method," *Revs. Mod. Phys., 18,* 323 (1946).

27

MAGNETIC RESONANCE TECHNIQUES

The molecular beam experiments of Stern and Gerlach and of Rabi yield valuable data concerning the magnetic moments of atoms and nuclei. In any beam, however, the density is so rare that the magnetic moments are independent of each other, and generally the atoms do not interact with each other. Recently techniques have been devised to study the magnetic properties of atoms and nuclei in solid or liquid materials. In this chapter, after a brief discussion of some of the physical principles involved, we will describe a few of the techniques used in this rapidly growing field of physics.

The mathematical theory used in interpreting the experiments will not be given but can be found summarized in the excellent review articles by Pake.[1,2] The basic idea is to orient the elementary magnetic moments in a homogeneous magnetic field, then to cause transitions between the various quantum states by applying radio frequency magnetic fields of the proper phase and polarization. This is exactly the same idea as that used in the magnetic resonance molecular beam method. There are, however, two important differences in the two cases.

1. Interactions of neighboring atoms.
2. Detection methods.

As we have said, in molecular beam experiments there is practically no interaction with the surrounding atoms. In a solid or liquid or even in a gas at atmospheric pressure such interactions cannot be ignored. In a ferromagnetic material, for example, the magnetic moments interact strongly and assist each other in lining up with the applied field. This effect is so strong that in a ferromagnetic single crystal saturation can be obtained with very small applied fields.

Ferromagnetic effects are caused, of course, by the magnetic moments of atomic electrons. Nuclear magnetic moments also interact with each other, but the interactions are much weaker since the magnetic moments of nuclei are so small. Thus a given nuclear magnetic

[1] Pake, G. E., "Fundamentals of Nuclear Magnetic Resonance Absorption. I," *Am. J. Phys.*, *18*, 438 (1950).

[2] Pake, G. E., "Fundamentals of Nuclear Magnetic Resonance Absorption. II," *Am. J. Phys.*, *18*, 473 (1950).

moment μ_n produces a field μ_n/r^3 at a neighboring nucleus at a distance r, causing it to precess with the appropriate Larmor frequency. This effect works both ways and provides a method of energy interchange between nuclear magnetic moments. The presence of this interaction has the effect of spreading the resonance curve and is one cause of line width. A quantity T_2, called the spin-spin relaxation time, is associated with the energy of interaction by the uncertainty principle $\Delta E \Delta t \approx \hbar$. The value of $1/T_2$ is of the order of the Larmor frequency of μ_1, in the field of μ_2.

$$\frac{1}{T_2} \approx \frac{\Delta E}{\hbar} = \frac{\mu^2}{r^3 \hbar}$$

This turns out to be of the order of 10^{-4} sec for nuclear spins in solids. In the absence of other effects, $1/T_2 = \Delta\nu$ gives the *line width* of the resonance.

In addition to the spin-spin interaction there is also the interaction of the magnetic moment with the crystal lattice. The effects to be considered here are:

1. Brownian motion.
2. Lattice vibrations.
3. Paramagnetic impurities.
4. Conduction electrons (in metals).

The first two are thermal effects present in any substance, and they are of extreme importance in dealing with nuclear magnetic resonance. Consider what happens to the nuclear magnetic moments when a sample is placed in a strong magnetic field. The moments are at first randomly oriented but occasionally through interaction with the lattice they acquire the right amount of energy to orient properly in the magnetic field. The length of time required to do this is called T_1, the spin-lattice relaxation time, and such relaxation times may be of the order of seconds to several minutes in pure substances. More exactly, T_1 is the characteristic time constant of the approach of the spins to thermal equilibrium in a magnetic field. Addition of small amounts of paramagnetic materials reduces this relaxation time considerably, and some substances may have thermal relaxation times as short as 10^{-5} sec. The time $T_1 = T_2$ for liquids at high temperatures and gases.

Once the substance is in the magnetic field it is necessary to disturb the distribution of moments in the various quantum states, and hence the net macroscopic nuclear magnetization, by applying perturbing magnetic fields. The detection apparatus, by detecting changes in the net macroscopic nuclear magnetization, can give us information about the transitions between various quantum states. Fortunately, the

molecular magnets are constantly making transitions because of the thermal effects. Consider, for example, the simple case where only two orientations are allowed. The probability of transition in the presence of radiation of the resonance frequency from state E_1 to state E_2 (Fig. 1) is exactly equal to the probability of the reverse transition. In equilibrium, however, more nuclei are oriented in state E_1 than in E_2, because of the Boltzmann factor in the statistical distribution of energies with temperature. Pake[1] calculates for a particular case that for every 1,000,000 protons in the upper energy state there are 1,000,014 in the lower energy state. It is with these 14 nuclei that we must work to cause detectable transitions. Thus if a disturbing magnetic field is applied it will have the effect of causing transitions of the excess nuclei in the lower energy state to the upper energy state. Although most samples contain perhaps 10^{21} nuclei rather than a million, the relative effects are small and the experimental problem in nuclear magnetic resonance is to detect a small effect.

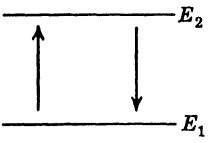


Fig. 1. Energy level diagram.

There have been two principal experimental approaches to the problem. In the Purcell[3] method of nuclear magnetic resonance absorption, the sample is surrounded by a coil that supplies radio frequency energy. As the frequency is varied through resonance, the effect of the applied field is to cause in the sample a net rotating macroscopic nuclear magnetization, which varies depending on the difference between the applied frequency and the resonant frequency. Now a changing magnetization produces an emf, in this case a back emf on the exciting coil, and this emf is then detected by its reaction on the source of exciting voltage. The resonance curves for the nuclear magnetization can thus be traced out by varying the frequency of the applied radio frequency field and detecting the emf induced in the coil by the sample. The resonance curves can equally well be found by varying the homogeneous magnetic field and keeping the frequency constant. The direction of the vector representing the net nuclear magnetization does not in general coincide with the applied field. Thus the applied emf and the induced emf are not in phase, and depending on the arrangement of the detecting apparatus, either the absorption resonance curve or the dispersion resonance curve can be obtained. The absorption effect is analogous to the resistance in an ordinary resonant circuit, whereas the dispersion effect is analogous to the reactance.

[3] Purcell, E. M., Torrey, H. C., and Pound, R. V., "Resonance Absorption by Nuclear Magnetic Moments in a Solid," *Phys. Rev.*, *69*, 37 (1946).

In the Bloch[4] method of nuclear induction, the induced emf is picked up by a detector coil, and a transmitter coil is used to supply radio frequency energy to the sample. Both the absorption and dispersion curves can be found in the Bloch method, just as in the Purcell method.

NUCLEAR MAGNETIC RESONANCE ABSORPTION. The principles of the method can be understood by considering the first experiment of Purcell, Torrey, and Pound.[3] Figure 2 shows their arrangement. A cavity resonator filled with paraffin was placed in a magnetic field of about 7000 oersteds. Power from the signal generator at 30 megacycles was fed in one side and coupled out the other to a coil. Part of the power from the signal generator was sent through an attenuator and phase shifter, and the phase was adjusted so that no signal appeared on the grid of the amplifier when H_0 was off resonance. Then H_0 was

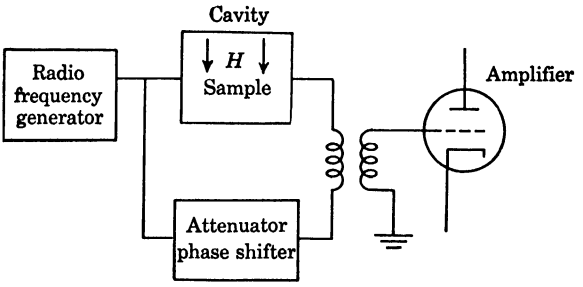


Fig. 2. First experimental arrangement of Purcell.

varied slowly, and when resonance occurred the absorption in the cavity due to the proton moments caused the two signals to the coil to be out of balance, and a signal appeared on the grid of the amplifier. They found that resonance occurred at 7100 oersteds and a frequency of 29.8 megacycles, with the full width of the resonance at half value about 10 oersteds. It should be said that the width depends not only on the factors discussed in connection with relaxation times, but also on the variation of the magnetic field over the volume of the sample. Thus it is important to use very homogeneous magnetic fields in these experiments. The radio frequency power used in the first experiment was only 10^{-11} w, to make sure that thermal equilibrium was not destroyed by the applied field.

Techniques have advanced rapidly since this first nuclear absorption experiment, and the sensitivity of various types of apparatus has

[4] Bloch, F., Hansen, W. W., and Packard, M. E., "Nuclear Induction," *Phys. Rev.*, *69*, 129 (1946).

been increased tremendously. A block diagram of an apparatus developed by Bloembergen, Purcell, and Pound[5] is shown in Fig. 3.

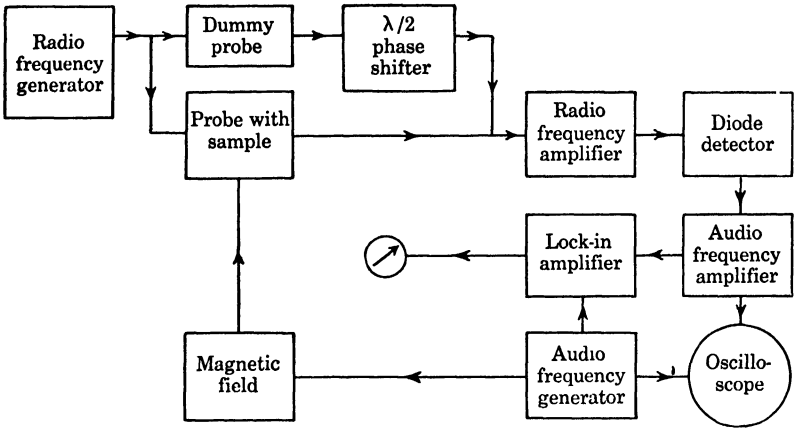


Fig. 3. Magnetic resonance absorption apparatus of Bloembergen, Purcell, and Pound.

Many features of this apparatus are found in other arrangements, and indeed the block diagram of the apparatus used in microwave spectroscopy has several similar elements. The radio frequency generator sends power into the probe placed in the magnetic field and into the dummy probe, which is identical to the real one except that it is not in a magnetic field and has no sample in it. A phase shifter

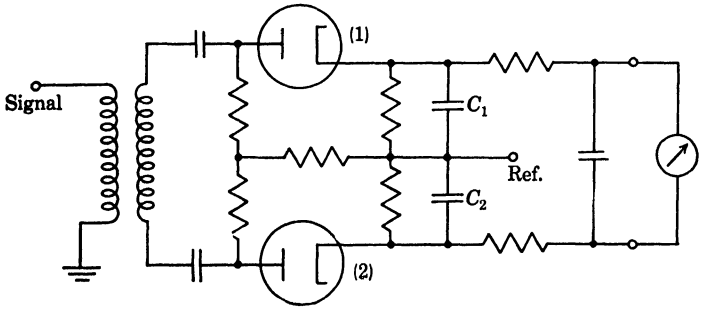


Fig. 4. Lock-in amplifier.

made of a length of transmission line or a properly designed transformer shifts the phase of the signal from the dummy probe by 180° before it enters the radio frequency amplifier. The difference radio

[5] Bloembergen, N., Purcell, E. M., and Pound, R. V., "Relaxation Effects in Nuclear Magnetic Resonance Absorption," *Phys. Rev.*, *73*, 679 (1948).

frequency signal is then detected in the diode detector. An audio frequency component of 30 c is added to the radio frequency signal by modulating the magnetic field by an audio frequency generator. Thus after detection the signal is an audio frequency signal, which is amplified and sent into the oscilloscope for visual observation. The audio frequency generator provides the sweep frequency for the oscilloscope. The audio frequency signal is also fed into the lock-in amplifier, which is sometimes called a phase-sensitive detector (Fig. 4).

In the lock-in detector the signal is applied in push-pull to the plates of symmetrically arranged diodes, so that when one plate is positive the other is negative. If diode 1 has a positive voltage on it, capacitor C_1 charges. On the reverse cycle diode 2 conducts, charging capacitor C_2. The voltage across the meter M is zero, however, because C_1 and C_2 charge up to equal and opposite potentials and are in series. Suppose that a reference voltage is fed in as indicated, in phase with the signal. Plotted in Fig. 5 are the signals on the plate of diodes 1 and 2 and the reference voltage on the cathodes. We have taken the reference voltage equal to the signal voltage for simplicity. Between a and b, neither tube conducts. Tube 2 has a negative plate, and tube 1 has no net voltage between plate and cathode. Between b and c, tube 2 conducts, but tube 1 has a negative plate, so does not conduct. Thus under these circumstances only capacitor C_2 charges, and a voltage is built up that is measured by the meter M. It is not necessary to have the reference voltage exactly equal to the signal voltage to get an output voltage, but it is necessary to have the correct phase relative to the signal voltage. Actually, for proper operation of the lock-in detector, the reference signal should be much larger than the true signal. The time constant of such a circuit may be made quite long, and this is equivalent to decreasing the bandwidth and therefore the noise. Usually the output of the lock-in detector is fed into a recording milliammeter to make a permanent record of the resonance curves. The lock-in amplifier is usually used with low magnetic field modulation, and the oscilloscope trace is observed directly when the modulation is greater than the bandwidth.

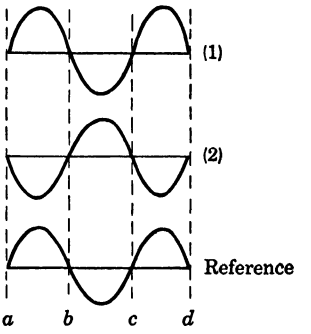


Fig. 5. Voltages on lock-in amplifier.

A somewhat different type of apparatus has been developed by Pound and Knight in experiments with metals. A block diagram is given in Fig. 6. The operation of this equipment is based on the fact

that the oscillator providing the radio frequency and the probe are tightly coupled together, so that when resonance occurs the operation of the oscillator is affected.[6] For example, the sample may be placed in the coil of the tank circuit of the oscillator, and the Q of the tank circuit changes when resonance with the nuclear moments is found. In such a circuit the capacitance in the tank circuit may be slowly changed to search for possible resonances in a fixed magnetic field. The remaining

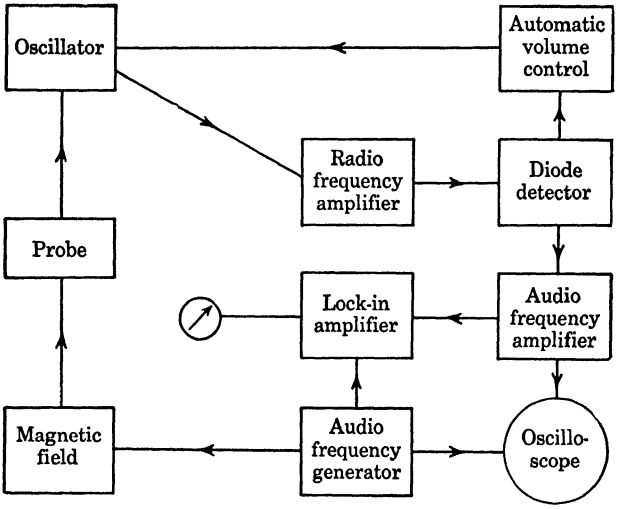


Fig. 6. Apparatus of Pound and Knight.

part of the apparatus operates in very much the same way as that in Fig. 3.

NUCLEAR INDUCTION. Bloch and his collaborators at Stanford have used a different method of detection. In the Bloch method the induced emf's caused by changes in nuclear magnetization are picked up by a receiver coil, amplified, and then detected. A transmitter coil produces the original transitions and must also be near the sample. To keep radio frequency energy from going directly from the transmitter coil to the receiver coil and blocking the amplifier, the two coils are wound at right angles to each other. A block diagram of the nuclear induction method is given in Fig. 7.

A problem specific to this method is that of reducing the direct coupling between the transmitter and receiver coils to a very low value.

[6] Pound, R. V. and Knight, W. D., "A Radiofrequency Spectrograph and Simple Magnetic-Field Meter," *Rev. Sci. Instr.*, *21*, 219 (190).

It is in practice impossible to wind the coils so that no coupling exists. Instead, a "paddle" made of a strip of copper is placed near the two coils and adjusted so that the eddy currents in the paddle distort the radio frequency field lines and reduce the coupling.

The voltage induced in the receiving coil may be of the order of 4×10^{-3} v, so it is fairly easy to amplify and detect. The power required in the transmitter coil is of the order of 10 mw which gives about 0.2 gauss for the perturbing field in a particular case.

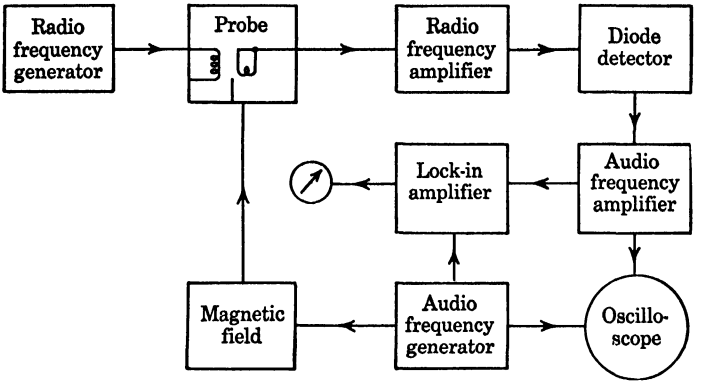


Fig. 7. Nuclear induction apparatus of Knight.

MAGNETIC FIELD MEASUREMENTS. Many techniques have been omitted from this chapter because in order to understand them, a much more detailed theory of the phenomena must be given. However the use of nuclear resonance in the measurement of magnetic field strength is becoming a fairly common procedure, and some remarks can be made concerning this technique.

1. The technique works well for fields over 500 gausses; it can be used for lower fields but is considerably more difficult.

2. The inhomogeneity of the field must not be great. If the field varies by more than 10 gausses over the volume of the sample, the line may be so broad as to be lost in the noise.

3. Paramagnetic salts are added to widen the proton resonance and make it more easily detectable. A rule for adding paramagnetic salts is the following: 1 molar manganese sulfate will produce an equivalent inhomogeneity over the sample of 1.2 gausses. Other concentrations will give roughly proportional results.

4. Microphonics in the radio frequency probe may give serious difficulty, especially in nuclear induction apparatus. These components must be mechanically rigid.

The magnetic field may be obtained with great accuracy by this method. The magnetic field is[7]

$$H = \frac{h\nu}{g\mu_p}$$

where ν = resonant frequency, h = Planck's constant, $g = 2$, $\mu_p = (1.4100 \pm .0002) \times 10^{-23}$ dyne cm/gauss. The frequency can be measured accurately by comparison with WWV, easily to 1 part in 10^5, so the accuracy of the method is determined by the accuracy of μ_p, which in turn depends on absolute measurement of a magnetic field. The method is even more accurate for comparison of magnetic fields.

Another application of this technique is the stabilization of magnetic fields. A nuclear resonance probe may be inserted in a magnetic field the value of which is required to be held constant. The signal at resonance can be used to correct the magnetic field by changing the d-c generator field voltage slightly if the magnetic field has any tendency to change its value.

Nuclear magnetic resonance techniques have been used in many different kinds of experiments in the last few years. Bloch and his collaborators have studied the nuclear moments of over 50 substances, and the techniques have also been applied to solid state physics: crystal structure, chemical binding, and the electron structure of metals by interaction of nuclear moments with internal fields.

MAGNETIC RESONANCE IN PARAMAGNETIC AND FERROMAGNETIC MATERIALS. Paramagnetic and ferromagnetic effects are caused by the magnetic moments of *electrons*. It is possible to obtain information about the structure of solids and liquids by using magnetic resonance techniques with such materials. The fundamental ideas are similar to those already discussed in connection with nuclear magnetic resonance, and many of the techniques used are quite similar. A sample is placed in a magnetic field, the electronic magnetic moments assume their quantized directions, and perturbing fields are applied to cause transitions between various states.

There are some important differences between the phenomena encountered in nuclear magnetic resonance and electronic magnetic resonance. First, the electronic magnetic moment is about 1000 times larger than the nuclear moment. Energy differences are correspondingly larger, and thus since $h\nu = \Delta E$, the frequencies of perturbing fields must be 1000 times larger. Nuclear resonance frequencies are of

[7] Thomas, H. A., Driscoll, R. L., and Hipple, J. A., "Measurement of the Proton Moment in Absolute Units," *Phys. Rev.*, *78*, 787 (1950).

the order of 30 megacycles, so electronic resonance frequencies are in the microwave range, say 30,000 megacycles. Thus microwave techniques, wave guides, etc., are used in the radio frequency system. The samples are placed in small resonant cavities such as can easily be made with wave guides.

With large magnetic moments, the electrons also have large interactions with each other, and very wide lines are expected. This is overcome by *diluting* the paramagnetic material and so increasing the distance between the magnetic moments. The loss of signal by dilution is partly made up by the sharpening of the resonance. Various other interactions are quite large. In a crystal, the electronic magnetic moment interacts strongly with the lattice by means of exchange interactions with other electrons and with the electric field of the atoms in the lattice. The latter may produce a Stark-effect splitting of the magnetic resonance lines, so that often the resonance patterns obtained are quite complicated. Analysis of these patterns can be made to yield valuable information about the structure of the electric field in the crystal and hence the structure of the crystal itself.

An introduction to the techniques and ideas used in these experiments can be found in an article by Kikuchi and Spence.[8] In the case of paramagnetic resonance the number of moments lined up in any particular direction is nearly equal to the number lined up in any other allowed direction, just as in nuclear resonance. There are, however, slightly more of the moments in lower energy states than in higher energy states because of the Boltzmann factor, and disturbances in this equilibrium distribution can be detected. Aside from the microwave techniques used in the radio frequency system, the other detection techniques are similar to those discussed under nuclear magnetic resonance, and the microwave techniques are quite similar to those used in microwave spectroscopy.

In the case of ferromagnetic materials the exchange interaction between the electrons is so strong that they all line up in the same direction. Here the interaction energy is much larger than kT, so that the Boltzmann factor is not important in the distribution. Thus the signal to be expected is much larger. The only difficulty is that ferromagnetics are metals, and the skin effect in metals at these frequencies prevents penetration of the microwaves into the body of the sample, and the signal is correspondingly reduced. Even so, however, signals from ferromagnetic materials are quite large.

Another phenomenon of interest in ferromagnetic materials is the relation of the resonant frequency to the orientation of the crystal.

[8] Kikuchi, C. and Spence, R. D., "Microwave Absorption in Paramagnetic Substances," *Am. J. Phys.*, *18*, 167 (1950).

Ferromagnetic materials have "hard" and "easy" directions of magnetization. For example, a resonance curve (Fig. 8) may be plotted for the absorption at a given frequency. If the crystal is rotated through an angle θ, the position of H_0 changes (Fig. 8b). Information

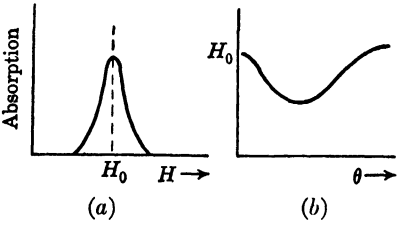


Fig. 8. (a) Resonance curve for ferro-magnetic resonance. (b) Variation of resonant frequency with orientation.

about the magnetic properties and structure of ferromagnetic materials can be obtained in experiments of this kind.

Ferromagnetic resonance was first found by Griffiths,[9] and the theory has been worked out by Kittel.[10]

BIBLIOGRAPHY

Bloembergen, N., Purcell, E. M., and Pound, R. V., "Relaxation Effects in Nuclear Magnetic Resonance Absorption," *Phys. Rev., 73,* 679 (1948).

Cooke, A. H., "Paramagnetic Relaxation Effects," *Reports on Progress in Physics,* Vol. 8. London: The Physical Society, 1950.

Kikuchi, C. and Spence, R. D., "Microwave Absorption in Paramagnetic Substances," *Am. J. Phys., 18,* 167 (1950).

Pake, G. E., "Fundamentals of Nuclear Magnetic Resonance Absorption. I," *Am. J. Phys., 18,* 438 (1950).

Pake, G. E., "Fundamentals of Nuclear Magnetic Resonance Absorption. II," *Am. J. Phys., 18,* 473 (1950).

[9] Griffiths, J. H. E., *Nature, 158,* 670 (1946).

[10] Kittel, Charles, "On the Theory of Ferromagnetic Resonance Absorption," *Phys. Rev., 73,* 155 (1948).

28

PILE TECHNIQUES

INTRODUCTION. The essential features of slow-neutron piles are now rather widely known. The pile consists of an aggregate of materials in which slow neutrons can multiply. The pile contains two basic materials. The first of these, which may be called the nuclear fuel, is a material that undergoes fission upon absorption of a slow neutron, with emission of more than one fast neutron. The second material, called the moderator, serves the function of reducing the energies of the fast neutrons to make them slow neutrons. If the complete process that started with one slow neutron results in slightly more than one slow neutron on the average, the neutron density in the pile will increase, and we have a self-sustaining neutron chain reaction.

The most available nuclear fuel is uranium. Since only a small fraction of normal uranium (the isotope U^{235}) has the property of undergoing fission after the absorption of a slow neutron, the pile must be very carefully designed if it is to have a self-sustaining reaction, in order to maximize the effect of the U^{235} and minimize the effect of U^{238}. We will discuss the design in somewhat more detail subsequently.

The moderator must have the property of reducing fast neutrons (with kinetic energy of about 2 mev) to slow neutrons (kinetic energy about 0.03 ev) with the minimum probability that a neutron be absorbed in the moderating substance. The usual mechanism of moderation is by elastic collisions of neutrons with the nuclei of the moderator. In such collisions some energy is lost by the neutron and appears as the recoil energy of the struck nucleus. In order that the largest energy be lost in such a collision the moderator should consist of light elements. Some light elements with low absorption coefficients for neutrons are deuterium, beryllium, carbon, and oxygen.

OTHER TYPES OF PILES. Besides the slow-neutron pile first mentioned, piles can be made that operate with higher energy neutrons. An intermediate-energy pile might be based on absorption of neutrons of energy of, say, about 10 kev. Such a pile would require less moderator, and as we shall mention more specifically later, would require a higher grade of nuclear fuel. A high-energy pile can be made without any moderator, based on the absorption of high-energy neutrons just as they come from a fission process. One example of a high-energy pile

is the atomic bomb. Actually the transition from the slow-neutron pile to the high-energy-neutron pile may be considered gradual, there being possible pile constructions for all intermediate cases.

In general the slow-neutron piles are usually of the largest physical dimensions, because a neutron, in being slowed down in the moderator, travels a long distance. The pile must be considerably larger than this distance if neutron loss out the side faces of the pile is not to keep the reaction from being self-sustaining. Piles based on fission caused by high-energy neutrons, on the other hand, may be considerably smaller in typical cases, for a neutron will be utilized after only one or a few nuclear collisions.

The present discussion will be confined to slow-neutron piles, with emphasis on the graphite-uranium piles in which the nuclear fuel is uranium of normal isotopic constitution and the moderator is graphite. Only a slow-neutron pile can be made using normal uranium, which is less than 1 per cent U^{235}.

AUXILIARY APPARATUS USUALLY CONNECTED WITH PILES. The products of the fission process are elements of intermediate atomic weight that are highly radioactive. In order to protect personnel from the effects of this radioactivity it is necessary to have radiation shields around piles. Rather thick concrete shields are used in typical cases.

To control the neutron density within a pile, and hence to control the rate at which fissions are occurring, it is necessary to control the reactivity of the pile. Let us define reactivity. Consider the reproduction cycle outlined above, started by the absorption of one slow neutron. Let us say that on the average k slow neutrons result from this cycle. Then each basic cycle accounts, on the average, for a net change of $k - 1$ slow neutrons. We say that $k - 1$ is a measure of the reactivity of the pile. If $k - 1$ is positive, the neutron density in the pile increases with time. If $k - 1$ is negative, the neutron density decreases; and if $k - 1$ is zero, the neutron density is constant in time.

To control the neutron density of the pile, then, we must be able to vary k at will. In a slow-neutron pile this is usually done by moving cadmium rods into or out of holes in the pile. Cadmium has the property of absorbing slow neutrons very strongly, so that insertion of the cadmium rods increases the loss of neutrons, hence decreasing k.

Obviously it is necessary also to have instruments available to register the neutron density within the pile, so that the pile operator may know the level at which the pile is operating. Such an instrument might be an ionization chamber with a high-voltage power supply and a sensitive ammeter, connected in the usual manner to register the rate of ionization of the gas of the chamber.

FISSION AND NEUTRON MODERATION. In a fission of U^{235}, approximately 200 mev is released and is distributed in the following way:[1] 80% goes into the kinetic energy of the fission fragments and this in turn is dissipated as heat and is lost for nuclear processes; 2.5% is contained in the kinetic energy of the neutrons that are released in the fission, and this energy is taken up by the moderator in slowing the neutrons to thermal energies; 9.5% of the energy released goes into the radioactivity of the fission fragments; and 2.5% appears as prompt gamma energy at the time of fission. A popular set of numbers is arrived at, using 200 mev per fission: "If all the U^{235} in 1 pound of the normal isotopic uranium mixture is fissioned, the energy released is equivalent to the obtainable heat energy from 500 tons of coal."

For the purposes of this discussion, a pile with graphite moderator will be considered. The neutrons collide with the carbon nuclei in passing through the moderator and the neutron loses some energy at each collision.[2,3] In due time it becomes a neutron of thermal energy and thereafter exists in thermal equilibrium in the carbon until it recycles by inducing another fission or until it is lost by leakage or other absorption. If the uranium and graphite are mixed uniformly, the pile fails to operate, because of the failure of U^{238} to fission when thermal neutrons are absorbed. The net result is that on the average each neutron is unable to reproduce itself and hence no activity results in the pile. If it were required to make a pile based on a *uniform* mixture of uranium and graphite, it would be necessary to use uranium enriched in U^{235}.

The fact that it is indeed possible to make a successful pile using uranium of normal isotopic constitution is based on the following special properties of the isotopes U^{235} and U^{238}. Isotope U^{235} has a very large fission cross section for slow neutrons, while U^{238} has a much lower absorption cross section for slow neutrons. In the energy region from 1 ev to 1 kev U^{238} exhibits many resonances—narrow energy regions in which the capture cross section is extremely high. The approximate energy dependence of the cross section is shown in Fig. 1. In energy regions between resonances the capture cross section is

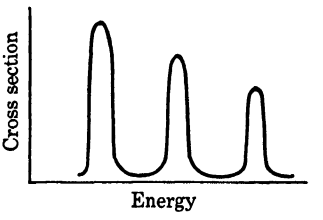


Fig. 1. Resonances in U^{238}.

[1] Ohlinger, L. A., "Engineering of Nuclear Reactors," *Nucleonics*, Dec. 1949, Jan., Feb., March 1950.

[2] Bothe, W., "On the Theory of the Slowing Down of Neutrons," *Z. Physik.*, *125*, 210 (1948).

[3] Marshak, R. E., Brooks, H., and Hurwitz, H. Jr., "Introduction to the Theory of Diffusion and Slowing Down of Neutrons," *Nucleonics*, I, *4*, (May 1949), II, *4*, (June 1949), III, *5*, (July 1949).

quite low. The capture of neutrons by U^{238} in this energy region will be called resonance capture.

In a pile that was a homogeneous mixture of uranium and graphite, the resonance capture of U^{238} would be very effective in absorbing neutrons and preventing the pile from being self-sustaining. However, it is possible by a trick to decrease enormously this resonance capture. The method involves putting the uranium in the form of lumps within a pure graphite moderator.

To see that the undesired effect of U^{238} has indeed been decreased by this step, let us examine the situation as neutrons of various energies enter a uranium lump. The size of the uranium lumps will be approximately equal to the mean free path of a thermal (slow) neutron in uranium. Thermal neutrons, when absorbed in the uranium, will be absorbed preferentially in U^{235} because of its large cross section. As far as thermal neutrons are concerned, the whole lump of uranium will be effective. Next consider a neutron in the range 1 ev to 1 kev, but of an energy between two resonances of U^{238}. For this neutron the mean free path in uranium will be quite long, and the whole lump will be effective; however, the cross section of U^{238} is very low between resonances, and the neutron in all probability simply leaves the uranium lump unaffected. Now, consider this same general energy range, but consider a neutron just at the energy of a resonance. The U^{238} capture cross section is very high, and this neutron will certainly be absorbed immediately on entering the lump. Its mean free path in uranium is exceedingly short, and it is absorbed very close to the outside surface of the uranium lump. The point now is that for producing fission in U^{235} the whole lump is effective, while for resonance absorption only a thin skin near the outside of the lump is effective. Thus there is a large self-shielding as far as the resonance absorption is concerned. This reduces the absorption drastically relative to the fission processes, and allows $k - 1$ to be made greater than zero and the pile reaction to be self-sustaining.

Carbon is a good moderator because its scattering cross section for thermal neutrons is 4.8 barns and its absorption cross section for thermal neutrons is only 0.0048 barn. Thus a given neutron, on the average, makes 1000 collisions in carbon before being absorbed by the moderator. For thermal energies, the mean free path λ of a neutron in carbon is about 2.5 cm; hence the neutron travels $\sqrt{1000}\,\lambda \approx 75$ cm before absorption in the moderator. The distance traveled by a neutron arising from a fission, during its process of slowing down to thermal energies, can be calculated as follows. In the range of 1.5 mev (fission energy) to thermal energy, it is assumed that the carbon scattering cross section for neutrons is of the order of its value at 1.5 mev, namely

2 barns.[4] For a fast neutron, the mean free path is 6 cm, and the logarithmic decrement of the energy per collision is $\eta = \ln \Delta E = 0.16$, on the average. From 1.5 mev to 1/30 ev then, $\ln \Delta E = 17$, and hence about 100 collisions are required to reduce a 1.5 mev neutron to thermal energy. In suffering this energy degradation, the neutron travels $\sqrt{100}\,\lambda_f = 60$ cm, on the average.

PILE BEHAVIOR. In the simplest theory the reproduction rate of neutrons in the pile is proportional to the numbers of neutrons present, that is, where n is the neutron density, $dn/dt = an$. The value of a will depend on the type of pile, the positions of the control rods, and the amounts of other absorbing materials that may be in the pile, as well as on the amount of neutron leakage out through the side walls of the pile. The solution of the differential equation is of the form $n = n_0 e^{at}$, where n_0 is an arbitrary constant representing the neutron density at time zero. One consequence of this differential equation and its solution is that by allowing a to be positive for a sufficient length of time, n can be made to reach an arbitrarily large value. This is in fact a rather good description. The neutron density is limited only by the consideration of thermal damage to the pile constituents. With each fission a fixed amount of energy is released as heat, and if allowed to remain within the pile this heat would eventually cause damage by melting the pile parts or otherwise changing them. The pile level is then limited by the cooling facilities within the pile. If heat can be taken from the pile at a large rate, the pile may be run at high neutron intensities without harm.

It is interesting to note that up to a certain point the permissible power level of a pile can be increased by installation of greater cooling facilities—for example, by installation of more cooling coils. However, if too many cooling coils were placed in the pile, they would increase the neutron absorption and would keep the pile from being made critical, and no self-sustaining reaction would be possible. In fact, the pile would not operate at all under these circumstances.

The level of activity of a pile is usually indicated by the power—rate of dissipation of heat—rather than by the neutron density. The two quantities are proportional for a given pile.

The position of the control rod determines only the time derivative of the pile power, as is shown in Fig. 2. The pile is by definition just "at critical" when the reactivity is zero and the pile power is constant in time, as it is in Fig. 2 to the left of t_1. At time t_1, the control rod was withdrawn from the pile by an amount to change the reactivity such that $k - 1 = +0.0001$. Now the time required for the elementary

[4] A barn is 10^{-24} cm^2.

reproduction cycle is known to be about 10^{-3} sec, so we would expect that $a = 0.0001 \times 10^3 = 0.1$ per second. In that case the pile power should increase by a factor of e during each 10 sec interval. Our descripion is basically correct, but the situation is slightly more complicated.

After fission processes, about 99% of the neutrons appear immediately, and these are termed "prompt" neutrons. However, about 1% of the neutrons appear with half-lives ranging from a fraction of a second to the longest, which appears to be about 55 sec. (The neutron-emitting nuclei are believed to have exceedingly short lives—the apparent lifetimes of the delayed neutron periods are presumably the lifetimes of beta emitters whose daughters contain neutron emitting nuclei.) Thus for 99% of the neutrons the fundamental cycle takes only 10^{-3} sec, while for the remaining neutrons the cycles require varying times, some of the order of 55 sec and longer.

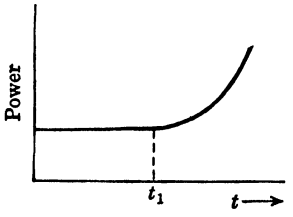


Fig. 2. Effect on power in pile when control rod is withdrawn.

The pile equations are somewhat complicated by the delayed neutrons, but help a great deal in controlling the pile level. Providing the reactivity is near zero and not over $k - 1 = 0.01$ (at which point the pile could be self-sustaining, using prompt neutrons only) the pile must essentially wait for the delayed neutrons to be emitted before the power can increase. That is, the actual rate of rise of pile power must be small enough so that the delayed neutrons are contributing appreciably, for without the delayed neutron contribution the pile would not be self-sustaining under these circumstances.

DETERMINATION OF NEUTRON ABSORPTION CROSS SECTION BY THE EFFECT OF A SAMPLE ON PILE REACTIVITY. If the absorption cross section of an element for slow neutrons is large, it can be measured in many ways outside the pile. However, if the absorption cross section is rather small, one of the best ways of measuring it is by the effect on pile reactivity when a sample of the material in question is placed within the pile.

NEUTRON DISTRIBUTIONS AND SAMPLE IRRADIATION. Research piles are provided with holes, running through the structure, into which samples can be introduced for irradiation. The samples are usually placed in cans of appropriate size and shape. The introduction of an absorbing material with a total absorption area of 10 cm^2 will have a large effect on the pile reactivity.

The scattering of neutrons by the samples introduced into the pile affects the reactivity by altering both the spatial and energy distribution of the neutrons in the immediate neighborhood. If any pure scatterer, i.e., a nonabsorber, is placed near a lump, it will decrease the neutron diffusion into the lump and hence decrease the reactivity. If the sample is near the edge of the pile in a position allowing it to scatter neutrons back into the pile, the reactivity of the pile is increased. If a pure scatterer is placed at a point within the pile where the gradient of the neutron density is zero, no change occurs in the spatial distribution of neutrons. In order to avoid disturbing the spatial distribution, irradiations are nearly always carried out at points symmetrically placed with respect to the lumps and near the center of the pile. Samples so placed affect both the energy distribution and pile reactivity.

The neutron absorption of the sample decreases the pile reactivity. On the other hand, the pile reactivity is increased by the following mechanism: the sample increases the slowing power of the moderator in its immediate neighborhood (which aids in the production of slow neutrons), and in so doing decreases resonance absorption of fast neutrons[5] by U^{238}. This effect usually is smaller than absorption effects.

In light elements, which generally have low neutron absorption cross sections, the increase in pile reactivity due to the moderation of neutrons by the sample (as just discussed) must be accounted for because it may be of the same order of magnitude as the decrease in reactivity due to absorption. An example of this effect is given by a determination of the absorption cross section of beryllium done at Chicago.[5] The samples were distributed in 16 cells, 456 g of metal per cell. The geometry was standardized by using cadmium wires in the same positions. In the notation of the next section, the results were: changes in reactivity of 1.46 ih due to sample moderation of neutrons, and -1.57 ih due to absorption were found. This corresponds to the relatively small observed net change of -0.11 ih in the pile reactivity. In heavier elements, the effect of sample moderation in increasing reactivity becomes negligible in comparison with the absorption effects.

Let one thermal neutron absorbed in uranium result, after fission, in k neutrons of thermal energy. If $k - 1 < 0$, the power decays, and if $k - 1 > 0$, the power grows. The unit called the "inhour," and abbreviated ih, is defined by

$$\text{ih} = C(k - 1)$$

[5] Anderson, H. L., Fermi, E., Wattenberg, A., Weil, G. L., and Zinn, W. H., "Method for Measuring Neutron-Absorption Cross Sections by the Effect on the Reactivity of a Chain-Reacting Pile," *Phys. Rev.*, *72*, 16 (1947).

Anderson *et al.*[5] give $1/C = 2.5 \times 10^{-5}$ for the first two graphite piles at Chicago. This seems to be a good value for uranium-graphite piles in general.

The inhour value corresponding to a given pile period T is

$$\text{ih} = \frac{54}{T} + \frac{20.3}{T + 0.62} + \frac{204}{T + 2.19} + \frac{535}{T + 6.5} + \frac{2036}{T + 31.7} + \frac{787}{T + 80.3} \tag{1}$$

where T is the pile period in seconds. In the limit of periods that are long in comparison with the numbers in the denominators, it is then seen that ih $= 3600/T$, or if T is in hours, ih $= 1/T$. This relation gives the inhour, or inverse hour its name. At critical, the pile period is infinite, so that ih $= 0$. The numbers in the denominators are the periods (mean lives) of the delayed neutrons. In the notation of (1) the pile period is 1 hr when ih (crit) $-$ ih $= 0.990$. The control rod, in any case, is calibrated in terms of inhours by measuring the deviation of the rod from some standard position and observing the pile period. The ih expression used is then employed to convert the observed periods into ih units. The effect of a sample on pile reactivity is usually expressed in inhours.

The value of T is negative when the pile reactivity is decaying; hence ih is also negative during decay, and positive during growth of the pile reactivity. If one inserts all the control rods to kill the pile, the reactivity will not go to zero immediately, but will decay according to the periods and intensities of the delayed neutrons that are still present when the rods are inserted.

The absorption cross section of an isotope can be measured by first making the pile critical. Then the sample is inserted into one of the holes provided for the purpose, and the pile period is noted. The sample is then replaced by thin pieces of boron (which has an accurately known capture cross section) which serve as monitors, and the period of the pile is again noted. Comparing the two periods and the relative sizes of the samples and monitors, one arrives at the capture cross section of the samples. The boron monitors must be thin foils to avoid shielding some boron nuclei by others nearer to the surfaces of the monitor foils.

Such an experiment was carried out for columbium at Chicago,[5] using the $1/v$ law to reduce the cross section to that for neutrons of 2200 m/sec, and the result was 1.4 barns. The absorption cross sections for thermal neutrons are usually quoted for an energy $E = kT$ corresponding to $T = 20°\text{C}$. This gives a neutron velocity of 2200 m/sec.

Anderson *et al.*[5] report that for precision measurements of absorption cross sections for thermal neutrons, the reactivity of the pile must

be corrected for temperature, atmospheric pressure, and humidity. For the Argonne graphite pile, the following coefficients have been determined

-0.814 ih/°C
-3.73 ih/cm Hg atmospheric pressure
-2.85 ih/cm Hg atmospheric water vapor

The humidity effect is due to the variation of atmospheric water vapor within the pile, and the more important pressure effect results from variations in the amount of atmospheric nitrogen within the pile. The air pressure must be read to 0.01 mm Hg, and because of rapid and spurious variations of this order of magnitude, the pressure does not have time to come to equilibrium over the entire pile. It is this variation that imposes the limiting sensitivity of the method for finding absorption cross sections with the pile. Temperature changes are slow and uniform, and can be nearly averaged out by performing comparison experiments.

THERMAL COLUMN. Carbon has a diffusion length for slow neutrons of about 75 cm, and a diffusion length for fast neutrons of about 60 cm because of its smaller absorption cross section for slow neutrons.

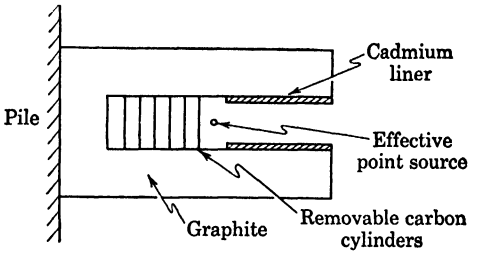


Fig. 3. Thermal column.

The result is that the ratio of slow to fast neutrons is large. The thermal column is a device for producing a very pure flux of thermal neutrons from the pile, and it is simply a graphite column of two or three meters square cross section and of a length that is calculated on the basis that the fast and thermal neutron densities decrease exponentially according to the smallest absorption cross section for each. Since the loss of thermal neutrons is smaller than that of fast neutrons, the flux from the thermal column becomes purer in thermal neutrons with increasing length of the column. Flux is defined as nv, where n is the neutron density and v is the neutron velocity. The appearance of the thermal column is shown schematically in Fig. 3.

The thermal column has a hole along its axis, that is filled by a carbon cylinder formed of many short, removable sections, thus

allowing the column to be adjusted to an effective length that gives a suitable beam intensity consistent with the fast neutron contamination that exists. An effectively point source of thermal neutrons exists at the base of the hole in the column. The cadmium liner is not inserted to the full depth of the hole, so that depression of the source strength is thereby avoided. A well-collimated beam of thermal neutrons is then available for experimental purposes at the exit of the thermal column.

The fast and slow neutron fluxes are about equal at the exit hole of the pile, but at a point in the thermal column at which the slow neutron flux has been attenuated by a factor of 10, the fast neutron flux has been reduced by a factor of about 100.

FLUXES IN EXISTING PILES. Some approximate central fluxes in various piles are given in Table 1.

TABLE 1

Argonne (D_2O)	$1 \cdot 10^{12}$	Brookhaven (C)	$5 \cdot 10^{12}$
Los Alamos (D_2O)	$1 \cdot 10^{11}$	Oak Ridge (C)	$1 \cdot 10^{12}$

MECHANICAL VELOCITY SPECTROMETER. The so-called *chopper*, or mechanical velocity spectrometer,[6,7] is used in conjunction with the thermal column to provide a beam that is still purer in thermal neutrons, which are used for experimental purposes. The schematic arrangement of the apparatus is shown in Fig. 4. The shield nearest the thermal column prevents neutron-captured gamma rays from entering the counting apparatus, and the series of cadmium slits provides adequate collimation. The boron fluoride counter is situated within a boron shield to reduce neutron background. Details of the chopper rotor or shutter are shown in Fig. 5. The rotor is a cadmium aluminum sandwich, cylindrical in shape, which rotates with its axis parallel to

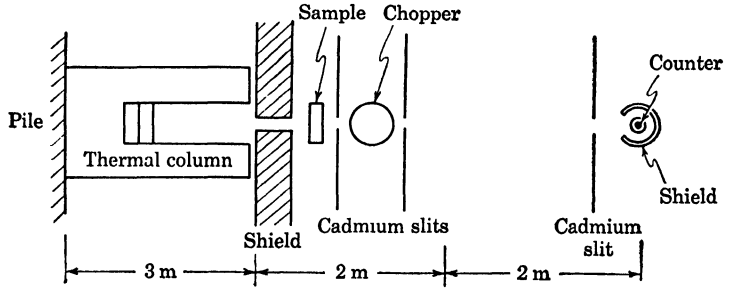


Fig. 4. Chopper.

[6] Fermi, E., Marshall, J., and Marshall, L., "A Thermal Neutron Velocity Selector and its Application to the Measurement of the Cross Section of Boron," *Phys. Rev., 72,* 193 (1947).

[7] Brill, T. and Lichtenberger, H. V., "Neutron Cross-Section Studies with Rotating Shutter Mechanism," *Phys. Rev., 72,* 585 (1947).

the collimator slits. The speed of the rotor is commonly 60, 120, or 240 rps, and the upper limit is imposed essentially by the mechanical strength of the rotor assembly under centrifugal force. Neutrons of energies from 0.003 to 0.3 ev are selected by the chopper. When the cadmium layers of the chopper are parallel to the neutron beam, a pulse of neutrons is allowed to pass. The path of a neutron in the aluminum of the chopper can be calculated by transforming into the frame of the rotor and noting that the centrifugal force is negligible in comparison with the Coriolis force, since the neutrons are close to the axis of rotation. The path of a neutron is shown approximately in Fig. 5 by the dotted line. For very slow neutrons the shutter is never open, and for neutrons of energy greater than about 0.3 ev the cadmium becomes transparent, resulting in a background of fast neutrons in the counters. This intensity can be checked by closing the shutter and taking a background count. The rotor thus selects neutrons of approximately 0.003 to 0.3 ev energy.

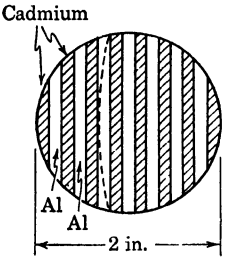


Fig. 5. Arrangement of apparatus using chopper.

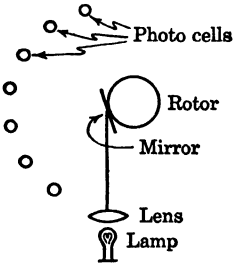


Fig. 6. Gating circuit using photocells.

The pulse of neutrons from the rotor is analyzed into several energy intervals by gating the several scalers. A mirror is attached to the rotor as shown in Fig. 6, and photocells are properly positioned around the rotor so that light reflected from this mirror results in output pulses from the photomultipliers at the appropriate times for the activation of several gating circuits. These gates control one scaler channel each, and with a known rotor speed, each scaler registers the number of neutrons within a given time interval at a measured time after the opening of the shutter. The pulse of neutrons, having been sorted out in terms of time of flight between rotor and counter, is therefore analyzed into neutron energy intervals. The diagram of Fig. 7 shows the relative positions in time of the gates and shutter opening.

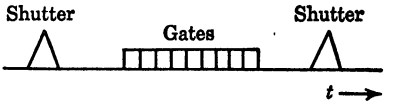


Fig. 7. Timing of gating circuit.

A CHOPPER EXPERIMENT. A typical experiment performed with the chopper is the determination of the total absorption cross section of some element or isotope, such as gadolinium.[7] A count was taken before the sample was in place, to determine the energy spectrum of the pulse of neutrons. The sample was then placed between the chopper and the thermal column, and the energy spectrum of the transmitted neutrons was determined. The total absorption spectrum of the sample was then calculated from the differences between the two runs. The energy dependence of the total absorption cross section of gadolinium is shown in Fig. 8. The resolution of the chopper-counter apparatus is given by the triangles on the energy axis.

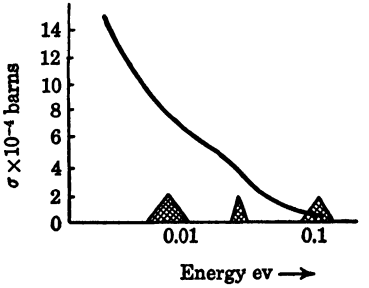


Fig. 8. Cross section of gadolinium.

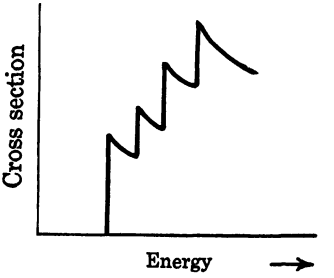


Fig. 9. Cross section for Bragg reflection.

CARBON NEUTRON FILTER. A block of carbon is made up essentially of a large number of randomly oriented microcrystals. It is well known that neutrons falling onto a single crystal obey the Bragg reflection law. In a block of carbon, neutrons that have energies greater than a certain minimum are then Bragg reflected by microcrystals that have the proper orientation. The beam intensity is thus attenuated. For energies below a certain minimum, the wavelengths of the neutrons are too long to satisfy any Bragg condition within the microcrystal, and hence neutrons of these energies pass through the filter unreflected. A filter of this type removes neutrons of virtually all energies above the minimum energy for reflection.

The behavior of the cross section for Bragg reflection in a large block of graphite is shown in Fig. 9. The initital rise from zero occurs at an energy at which the first Bragg reflection occurs. As the energy increases, the cross section falls off with $1/E$. The vertical discontinuities that occur are due to the energy reaching the minimum value for reflection from a new set of planes. Some simple crystal structures can be determined from this kind of data.

At very low neutron energies the scattering cross section of neutrons in carbon has a $1/v$ dependence. From considerations of time spent by

the very slow neutrons in the neighborhood of a carbon nucleus, this dependence can be seen to be due to fast carbon nuclei colliding with very slow neutrons.

CRYSTAL SPECTROMETER.[8,9] A large single crystal of some simple structure is mounted as shown in Fig. 10. The neutron beam from the pile is reflected according to Bragg's law from the crystal, and with appropriate collimation as shown, a monochromatic neutron beam is available for performing standard transmission experiments.

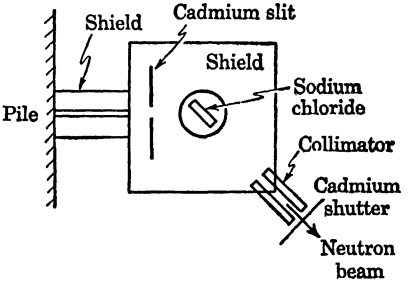


Fig. 10. Crystal monochromator.

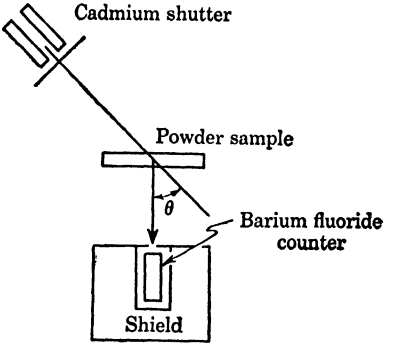


Fig. 11. Neutron diffraction by a powder.

With the apparatus shown in Fig. 11, the neutron analogs of Debye-Sherrer powder rings are easily detectable. The sample of powdered material is placed in the beam as shown, and the reflected neutrons are detected with a well-shielded and collimated boron fluoride counter which accepts neutrons coming only from the sample, at a given angle to the beam θ. The graph of counts per minute versus counter angle for aluminum is given in Fig. 12.

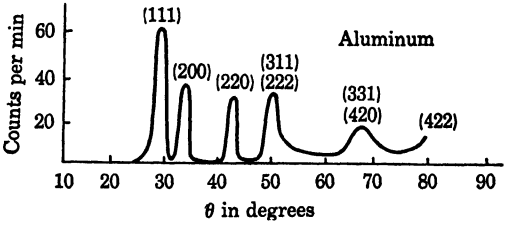


Fig. 12. Variation of counting rate with angle for aluminum.

[8] Wollan, E. O. and Shull, C. G., "Neutron Diffraction by Crystalline Powders," *Phys. Rev., 73,* 830 (1948).

[9] Shull, C. G., Wollan, E. O., Morton, G. A., and Davidson, W. L., "Neutron Diffraction Studies of NaH and NaD," *Phys. Rev., 73,* 842 (1948).

Crystal structure investigations based upon this method are of value when x-ray techniques fail because the compound contains nuclei of low Z (which give small x-ray scattering).

The scattering cross section and the amplitude of the scattered wave can be deduced from data of counting rate versus counter angle. The curves of Fig. 13 were found for sodium hydride and sodium deuteride.[9] The complete curves, including peaks for several other crystal planes, are not given. To determine the crystal structure of sodium deuteride and sodium hydride, the expected line intensities were calculated for the two lattice forms that could not be ruled out on other grounds. Comparisons with the data showed that the crystal structure is face-centered cubic. If only the sodium nuclei had contributed to the Bragg reflection, as in the case of x-ray reflection, the 111 plane would be expected to give a peak about twice as high as the 200 plane. This is not observed, therefore the hydrogen and deuterium nuclei also contribute to the reflections, or coherent scattering. If f_{Na}, f_H, and f_D are the scattering amplitudes of the sodium, hydrogen, and deuterium, respectively, a calculation leads to the expressions of Table 2 for the line intensities. An examination of the peaks and comparison with the calculated intensities and signs of the amplitudes in Table 2 shows that sodium and deuterium scatter in the same phase, and sodium and hydrogen scatter with opposite phase. The background with sodium hydride was greater than with the sodium deuteride. The coherent scattering cross sections for hydrogen and deuterium are found to be 2.0 ± 0.3 barns and 4.1 ± 1.2 barns, respectively.

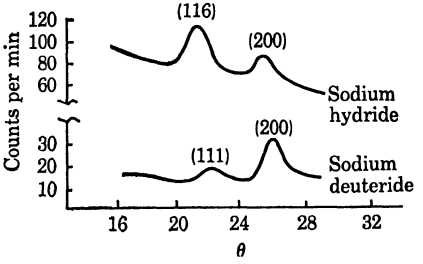


Fig. 13. Variation of counting rate with angle for sodium hydride and sodium deuteride.

TABLE 2

	NaH	NaD
111 plane	$(f_{Na} - f_H)^2$	$(f_{Na} - f_D)^2$
200 plane	$(f_{Na} + f_H)^2$	$(f_{Na} + f_D)^2$

C^{14} PRODUCTION IN THE PILE.[10] The valuable isotope C^{14} is being used in increasing quantities for tracer work, and thus it is impor-

[10] Norris, L. D. and Snell, A. H., "Production of C^{14} in a Nuclear Reactor," *Nucleonics, 5,* 18 (September 1949).

tant to secure a means for producing it in quantity. The reaction $N^{14}(n,p)C^{14}$ is used, the thermal neutron cross section for which is 1.7 barns. The neutron irradiation of the nitrogen nuclei is carried out by circulating a concentrated solution of ammonium nitrate through a tube of 2S aluminum that is built into the pile. The advantages of this system are: (1) a large quantity of nitrogen can be irradiated conveniently and with safety to personnel, (2) no parasitic neutron capture processes or side reactions are competing with the production of C^{14}, (3) extensive decomposition of the solution due to radiation or thermal effects is not likely.

The C^{14} is prepared in convenient form by chemical treatment of the irradiated solution.

CONTAMINATION OF BORON ABSORBERS. An interesting approximate calculation can be made in order to get some idea of the length of time an absorber or other reactant can be left in a pile before having to purify it. With a flux of 10^{12} neutrons per second and for a nuclear absorption cross section of the order of 1000 barns, there is $10^{12} \times 10^{-21} = 10^{-9}$ of each nucleus used up per second on the average. This results in a period of about 10^9 sec, or roughly 30 years before the boron concentration drops to $1/e$ of its original concentration in an absorber. Purification is therefore not necessary except at very infrequent intervals.

DELAYED NEUTRONS FROM U^{235} FISSION.[11] Delayed neutrons have been studied intensively because of their importance in the operation of piles. Roberts, Meyer, and Wang[12] first observed delayed neutrons in uranium fission in 1939 by bombarding uranium with neutrons from a lithium target under cyclotron bombardment. Later in 1939, Meyer and Wang[13] showed that the delayed neutrons were not due to (γ,n) reactions as had originally been proposed, but rather due to fission products in some way.

Hughes *et al.*[11] have determined the half-lives, energies, and relative yields of six different periods of delayed neutrons that arise in the fission of U^{235}. Their apparatus consisted of a pneumatic transfer apparatus, or "rabbit," in conjunction with the graphite pile at the

[11] Hughes, D. J., Dabbs, J., Cahn, A., and Hall, D., "Delayed Neutrons from Fission of U^{235}," *Phys. Rev.*, *73*, 111 (1948).

[12] Roberts, R. B., Meyer, R. C., and Wang, P., "Delayed Neutrons from U^{235} Fission," *Phys. Rev.*, *55*, 510 (1939).

[13] Roberts, R. B., Hafstad, L. R., Meyer, R. C., and Wang, P., "The Delayed Neutron Emission Which Accompanies Fission of Uranium and Thorium," *Phys. Rev.*, *55*, 664 (1939).

Argonne Laboratory. The schematic arrangement is given in Fig. 14. At the point *A* where the rabbit and target came to rest, the rabbit tube was buried in a well-shielded block of paraffin with a boron fluoride counter located within the paraffin at a variable distance from the tube. The scaler output was recorded on a standard portable electrocardiograph with time pulses forming the comparison trace. This procedure was necessary to secure as statistically accurate a decay curve as practicable for the shortest period, which was found to be 0.05 sec.

The procedure included the insertion and irradiation of the sample for various times. The sample was then blown out of the pile as rapidly

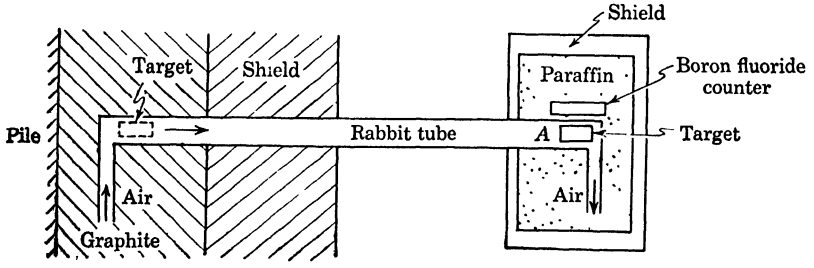


Fig. 14. Experiment of delayed neutrons.

as practicable to avoid excessive decay of short-period activities while the rabbit was in transit to the counter. The neutron periods found are given in Table 3.

TABLE 3

Half-life	*Energy*	*Yield (relative to total neutron emission)*
55.6 sec	250 kv	0 025%
22.0	560	0 166
4.51	430	0 213
1.52	620	0 241
0 43	420	0 085
0.05	...	0.025
		0.755% Total yield

The periods were determined in order, beginning with the longest. When determining a given period, the previously determined longer periods were treated as known in the graphical analysis. The intensities and times of irradiation were adjusted so that the period under study was emphasized. The periods were determined in the usual manner from graphical analysis of the decay curves, and the relative yields were determined by extrapolating the periods to zero time. An independent determination of absolute yields was made by placing a slug

of uranium metal in the beam of thermal neutrons and measuring the total neutron flux emitted by the metal under bombardment. The beam was then cut off by a rapidly acting shutter, and the decay curve of the delayed neutrons was obtained with the apparatus in the same position. This measurement gives the absolute yield of delayed neutrons directly.

All measurements were corrected for dead time of the counters and for various energies of the neutron periods, since the counter sensitivity was not independent of neutron energy.

The energies of the various periods of delayed neutrons were determined by making use of the energy dependence of the counting appara-

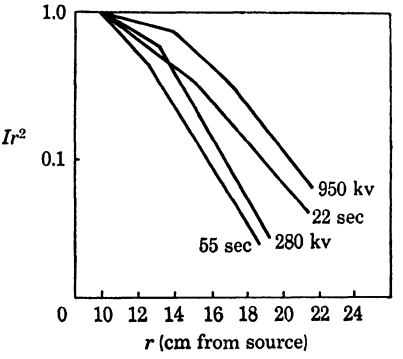


Fig. 15. Calibration curves with delayed neutron intensities.

tus. Saturated decay curves were obtained with the counter at various points in the paraffin block and the intensities of the various periods were determined as a function of the counter position. A plot of the normalized intensities of the 22 sec and 55 sec periods is shown in Fig. 15. Also plotted are the normalized intensities of a 950 kev sodium-beryllium and a 280 kev sodium-deuterium oxide standard photoneutron source. By extrapolation and interpolation, the energies of the 22 sec and 55 sec periods of delayed neutrons can be set at 560 kev and 250 kev, respectively. In this graphical manner, the data tabulated in Table 3 were obtained.

BIBLIOGRAPHY

Cochran, D. and Hansen, C. A., Jr., "Instrumentation for a Nuclear Reactor," *Nucleonics, 5* (Aug. 1949).

Cockcroft, J. D., "Peaceful Applications of Nuclear Fission," *Nature, 160,* 451 (1947).

Dunworth, J. F., "Operational Characteristics of the Harwell Piles," *Research* (London), *2,* 73 (1949).

Goodman, Clark, "Nuclear Principles of Nuclear Reactors," *Nucleonics,* (Nov. and Dec. 1948).

Goodman, Clark, ed., *The Science and Engineering of Nuclear Power.* Cambridge: Addison-Wesley Press, 1947.

Hurwitz, H., Jr., "Derivation and Integration of the Pile-Kinetic Equations," *Nucleonics, 5* (July 1949).

Marshak, R. E. *et al.*, "Introduction to the Theory of Diffusion and Slowing Down of Neutrons, IV," *Nucleonics, 5,* 59 (Aug. 1949).

Ohlinger, L. A., "Shielding from Nuclear Radiations," *Nucleonics, 5,* 4 (Oct. 1949).

Pryce, M. H. L., "Review of British Reactors," *Nucleonics* (Dec. 1948).

Ross, M. and Story, J. S., "Slow Neutron Absorption Cross Sections of the Elements," *Reports on Progress in Physics.* London: The Physical Society, 1948–1949.

Smyth, H. L., *Atomic Energy for Military Purposes.* Princeton: Princeton University Press, 1946.

Wallace, P. R., "Neutron Distributions in Elementary Diffusion Theory," I, *Nucleonics, 4,* (February 1949).

Wallace, P. R., "Neutron Distributions in Elementary Diffusion Theory," II, *Nucleonics, 4* (March 1949).

29

X-RAY DIFFRACTION

PRODUCTION OF X-RAYS. *X-ray spectra* are of two types: (1) continuous x-rays produced by deceleration of electrons in nuclear encounters in the target which by virtue of their pulse character must contain a wide range of frequencies subject to the limitation that the energy of the quantum emitted, $h\nu \leq \frac{1}{2}mv^2 = eV$, the bombarding electron energy, and (2) line spectra produced by quantum jumps following excitation of inner shell electrons by impinging electrons.

The continuous spectrum of tungsten for an incident electron that has lost no energy (by virtue of going through a thick layer of absorbing material) has the form shown in (a) of Fig. 1.

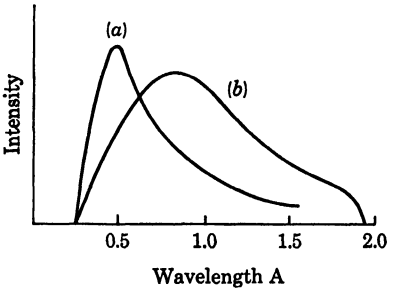


Fig. 1. X-rays from tungsten (continuous spectrum).

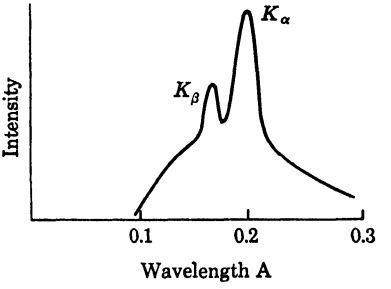


Fig. 2. X-ray line spectrum from tungsten.

If one takes into account the energy loss of the electrons, the spectrum has the form shown in (b) where the energetic limitation mentioned above accounts for the short wavelength cutoff. The cutoff at the long wavelength (≈ 2.1) is due to absorption which occurs in the thin (usually beryllium) window necessary to preserve the vacuum in the x-ray tube.

At higher voltage, the line spectrum for tungsten comes in and one gets a spectrum as in Fig. 2. As the wavelength limit becomes shorter, the more energetic transitions are reached, in the order M, L, and K.

The most prominent lines are the K and L lines and these, particularly the K lines, are used predominantly in the x-ray analysis.

FILTRATION. To obtain the monochromatic radiation necessary for most diffraction experiments two methods can be used to eliminate the background continuous spectrum and the undesired lines (e.g., the K_α lines of the above pictured tungsten spectrum).

(1) One can utilize a crystal with known plane spacing in accordance with the Bragg relation and thus select the wavelength corresponding to $\lambda = 2d \sin \theta$. The difficulty resulting from higher order (shorter λ) radiation received at the same angle, can be eliminated by operating the tube so that the high-voltage cutoff wavelength is longer than that of the higher order. Another difficulty, which cannot be obviated so simply, is that the necessarily long path as well as losses in the crystal reduce the intensity by very large factors. Only about 2% of the incident radiation is reflected from the crystal with the desired wavelength. However, this method produces sufficiently homogeneous wavelengths to allow determination of natural line widths.

(2) As the second method one can utilize the property of critical absorption, characteristic of the absorption of x rays by elements. To see how this may be achieved, consider a typical plot (for tungsten) of the absorption coefficient μ versus λ as shown in Fig. 3.

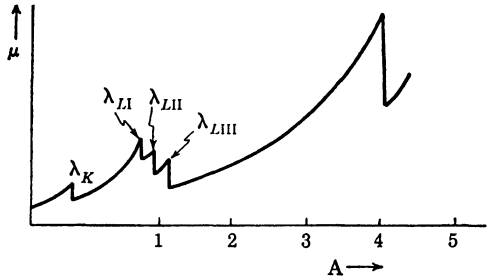


Fig. 3. Absorption coefficient for x-rays.

If we attempt to use the tungsten to absorb its own radiation, we will find that small absorbing layers transmit surprisingly large amounts of radiation, because the critical absorption wavelengths (see Fig. 3) are slightly shorter than the shortest emission wavelengths of a given series. However, if we use as absorber an element of lower atomic weight such that the K absorber wavelength falls between the K_α and the K_β emission lines, we can obtain quite monochromatic K_α radiation. (Thus hafnium, $Z = 72$, with K absorption limit at 0.190 A, or ytterbium, $Z = 70$, with K absorption limit at 0.2016 A, can serve as filters for a tungsten ($Z = 74$) target tube.) As a general rule, an element with atomic number $Z - 2$ can be used with a target of atomic number Z to give reasonably monochromatic K_α lines. Thus for a copper target, a nickel filter 0.0085 mm thick can be used; for a

molybdenum target, a zirconium filter 0.037 mm thick can be used. In each case the tube should be operated at the lowest voltage to give the K line desired.

TUBE TYPES. In general, x-ray tubes have been either of the high vacuum or gas types. The gas type tube consists only of a solid cathode (concave shaped aluminum), water cooled targets, and the residual gas, which provides positive ions that bombard the cathode when the potential is applied. These in turn produce the electrons that make up the cathode-ray stream which bombards the anode and gives rise to the x rays. The pressure of the gas is quite critical to both the potential drop necessary across the tube and the ion current that is maintained, so that these cannot be controlled independently. However, recent gas tubes have simple dynamic pressure control and interchangeable targets that are water cooled. In addition, because of the absence of the hot tungsten filament, the target can be kept free of such tungsten contaminant, and x rays are obtained from the pure target.

Most commercial tubes are at present of the high-vacuum type. In this tube, which is highly evacuated (to about 0.0008 micron of mercury), electrons are supplied by a pure tungsten filament heated to about 2000°C. The tube current under space-charge-limited conditions is controlled by varying the voltage so that the hardness of the x rays is still not wholly independent of the tube current. However, tube design allows adequate control of the two factors separately. Thus a typical high-vacuum tube, the G.E. C A-6 for diffraction use, has the following rating for continuous operation.[1] (*kvp* = kilovolts, peak.)

	30 *kvp*	40 *kvp*	45 *kvp*	50 *kvp*
Copper target	25 ma	20 ma	18 ma	16 ma
Nickel target	13	10	9	8

This tube has the structure shown in cross section in Fig. 4.

The necessity for water cooling of the target is obvious from the fact that in the above tube the size of the target spot that receives the 800 peak watts (e.g., 20 ma at 40 kvp) is 0.08 by 1.3 cm. The required water flow rate is 2 pints per minute at the minimum. Since the anode is run at ground potential (the cathode at "high negative"), the water can be taken directly from the city mains.

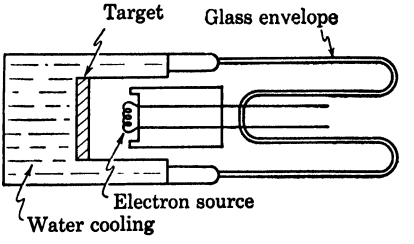


Fig. 4. X-ray tube.

[1] These data and other information about G. E. x-ray diffraction units were kindly communicated by W. E. Milburn of the General Electric Corporation, San Francisco.

To obtain an effectively concentrated focal spot (instead of the line), this tube has two beryllium windows on opposite ends of the diameter of the tube through the focal line. Thus the focal spot in this end view is approximately 1 by 1 mm.

Typical characteristics of tubes for x-ray diffraction are: large tube currents to cut down exposure time, continuous operation at steady output, fine focus, small absorption of beam in sample direction, and moderate voltages (25–50 kv). The latter is in contradistinction to the requirements in x-ray therapy work, which has made use of energies up to 2 mev.

ELECTRICAL REQUIREMENTS. The usual high-voltage source for diffraction tubes is a transformer and rectifying unit, which need only put out relatively small currents. The types of rectification and regulation necessary depend on the work to which the tube is to be applied. In general, in photographic recording of x rays it is not necessary to have a highly regulated or filtered supply, whereas in counter ("direct") measurements, a well-regulated and continuous (but not necessarily constant potential) supply is essential. Thus in the first case, a half-wave rectifier as shown in Fig. 5 may be used. The waveform is

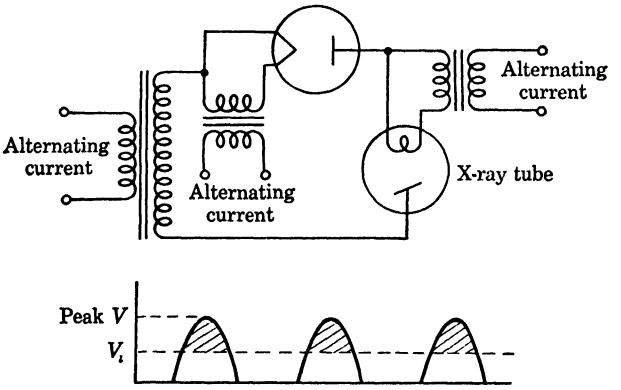


Fig. 5. X-ray tube and half-wave rectifier.

such that the x-ray beam is developed during less than half a cycle and contains the short wavelengths developed near the peak voltage for even shorter periods.

Half-wave rectification can also be obtained by using the rectifying action of the x-ray tube itself, without the use of additional tubes. However, to avoid mechanical effects that may shorten the life of the tube, it is desirable to use a separate rectifier. Full-wave rectification giving practically continuous emission is shown in Fig. 6. Regulation of

voltage may be obtained by an electronic line voltage stabilizer, which should be capable of maintaining the voltage to about 0.1% over a typical 2–3 hour run. The current also should be regulated either by an electronic method or by monitoring the x-ray beam by a second counter, which then modulates the cathode stream current appropriately. Filtering to give constant potential is fortunately seldom necessary because of the bulk and cost of the necessary capacitor banks.

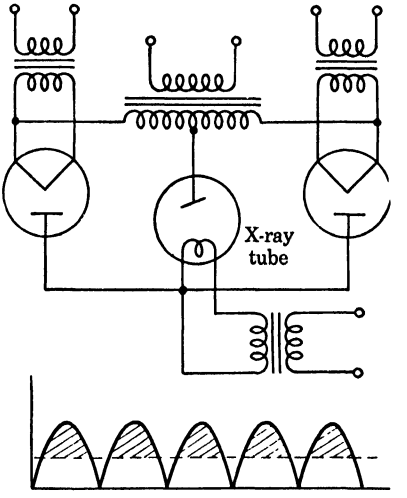


Fig. 6. X-ray tube and full-wave rectifier.

X-RAY DIFFRACTION. Review of Bragg theory of crystal diffraction. Each atom of the crystal scatters spherical waves from the direct x-ray beam. Huygens' construction will show that an envelope for those waves scattered by atoms lying in a particular plane travels as if reflected: the angle of incidence is equal to the angle of reflection. This reflection takes place no matter what the arrangement of the atoms within the plane; a perfectly random arrangement of atoms in a plane will act as a reflecting surface for x rays. The only condition for such a reflection is that the atoms all lie exactly on the plane. Any additional regularity of arrangement of the atoms manifests itself in the existence of other planes rich in scattering atoms.

A single x-ray pulse will be reflected by these planes of atoms. To detect such a reflected pulse the detector need only be placed at an angle from the normal to the reflecting plane equal to the incident angle. If there were only one plane the reflection would be the same as from a plane mirror.

However, since the planes of the crystal occur displaced from one another by characteristic distances d, (which distances depend on the particular plane), the reflected pulse will be a train of waves resulting from successive reflection by parallel planes. A train of waves of wavelength λ will experience interference among reflected waves from the same set of parallel planes. The usual construction shows that constructive interference results only when the angle θ satisfies the relation

$$n\lambda = 2d \sin \theta$$

where θ is the angle between the x-ray beam and the plane. Consequently the crystal must be held at the correct angle (for a given one of the finite planes) and even then only one order of the spectrum is obtained at a time.

VARIOUS TYPES OF DIFFRACTION PATTERNS.

1. *Powder diffraction.* To understand how powdered crystals give x-ray diffraction patterns let us imagine the following experiment. Consider first a single crystal. Send in the wavelength from a point source

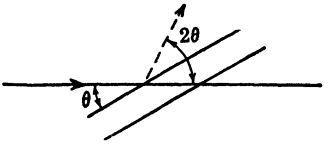


Fig. 7. Reflection from crystal planes.

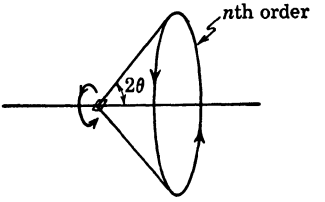


Fig. 8. First rotation.

in a given direction. Consider one set of planes of separation d, (Fig. 7). Only one angle exists for which reflection of the nth order takes place. Suppose now we rotate the crystal about an axis along the incoming beam direction. The beam sweeps out a circle as shown in Fig. 8. If there were only one plane the reflection would be the same as from a plane mirror.

This assumes, however, that we have chosen the angle θ correctly to get a reflected spot. To make sure that the angle is chosen correctly, at least during part of the experiment, we make another rota-

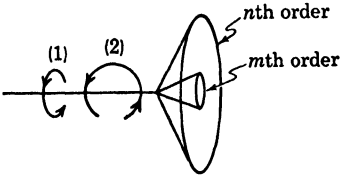


Fig. 9. Second rotation.

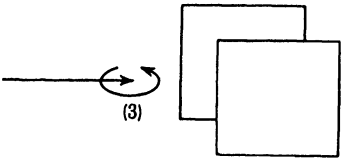


Fig. 10. Third rotation.

tion, this time about an axis perpendicular to the paper (Figs. 8, 9). We thus get a ring for each set of planes in the crystal and each order of reflection from each set of planes.

There may have been planes that were just parallel to the paper (Fig. 10), and in order to get reflections from them we must make a third rotation about an axis perpendicular to the x-ray beam and parallel to the paper, which gives, with the other two rotations, circular

patterns for every plane separation d, and every order of reflection. Now instead of orienting the crystal in every possible position in space (which is what the rotations have achieved), we can use powdered crystals that, because of their random orientation, are already oriented in every possible direction.

In powder diffraction we need monochromatic light, and we get circular patterns—one circle for each order of each d, (for one particular λ).

2. *Laue patterns.* These are produced by using a *single* crystal with continuous x-ray spectrum incident. Then all d's will find in the x-ray spectrum the requisite λ for intense reflection. (The same is true for all orders.) Since the crystal is not rotated (Fig. 11), the selectively reflected beams end up in spots on the film; each spot represents a set of planes (of particular d) and a particular order n.

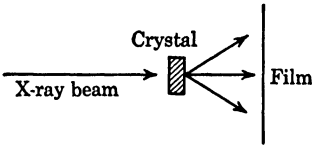


Fig. 11. Arrangement for Laue patterns.

3. *Single-crystal patterns.* A single crystal rotated about its principal axes will produce the same patterns as the powder patterns. Special techniques, however, have made use of single crystals the most powerful in the investigation of crystal structure. These techniques will be more fully described under *detection.*

The important points about x-ray diffraction are:

1. X-ray diffraction is *selective reflection.*

2. For a monochromatic beam a single crystal will reflect only if the crystal is oriented at an "allowed" position—and then only one order is reflected.

3. For powder, the pattern is a set of circles—one circle for each set of planes and each n (order of reinforcement).

4. For a continuous x-ray beam a single crystal will give *spots;* each spot equals one set of planes and one n.

DETECTION AND MEASUREMENT.

Collimation. The main purpose of the detector is to measure accurately the angles between the incident beam and the scattered beams. Consequently the incident beam must be collimated with an accuracy commensurate with the precision desired in the measurement of the scattered beams.

The usual conflicting requirements of high intensity and good collimation obtain for x rays. The collimation is provided by lead sheets of the order of thickness of 1 mm. The holes or slits in the sheets are of the order of 10^{-2} cm^2. Two such defining slits, about 5 cm apart, will serve for a moderately precise collimating system. The position of

these slits relative to the x-ray tube target determines the solid angle. The slit system can be placed within a centimeter of the tube target in a typical system, the resulting exposure time being of the order of 10 min. For better resolution smaller slits can be used, necessitating exposures that have run into hundreds of hours.

The whole region should be evacuated in order to reduce absorption and scattering by the air. The same vacuum system as the x-ray tube is sometimes convenient to use.

The *type of target* depends upon the method to be used.

For Laue patterns a continuous x-ray spectrum must be used as indicated above. A single crystal is used, as big as the x-ray beam diameter or larger. The resulting Laue spots on a film will be in a pattern of *n*fold symmetry for a crystal that has *n*fold structural symmetry about the x-ray beam direction. To learn anything about the structure from such a pattern, the intensity of each spot and the spectrum of the x-ray beam must be known. Because of the difficulty in determining these, the method is used only to determine the axes of symmetry and the order of symmetry of the crystals.

For single-crystal patterns (rotation method) a single crystal is again used. Additional resolution and consequent loss of intensity is obtained by using a crystal smaller than the x-ray beam. A monochromatic beam is used, and the crystal is rotated about an axis of symmetry. The various planes satisfy successively during the rotation the Bragg relation for the given wavelengths, and thus a pattern of spots is formed on the film. The resulting pattern is analyzed by measuring the intensity of the spots and the geometry of the curves on which the spots fall. The information so gathered gives the electron densities within the crystal and thus the atomic arrangement. It should be noted that in this method a constant-intensity x-ray beam is necessary. The modifications of this method involve the detection apparatus and will be discussed in that section.

For the powder method the sample is a tube or wire containing powdered crystals of the material to be studied and the beam must be monochromatic.

CAMERAS OR DETECTORS.

Flat film. The simplest detector is a flat film placed perpendicular to the x-ray beam. This detector can be used in all methods of x-ray diffraction mentioned above. It is usual for Laue patterns. It is also well adapted to use with single crystals and monochromatic x-rays (Bragg spectrometer), since the interpretation of the lines is so straightforward.

To avoid scattering by the film of the primary beam, a small piece

of lead should be used to absorb the primary beam—a nail with a lead head is useful for this purpose.

Cylindrical film. A strip of film bent into a circle with a specimen at the center, and the x-ray beam entering through a hole in the film, is usually used with powdered crystal diffraction and the rotating crystal method. The strip then records curved spectral lines.

Weissenberg goniometer. Since the rotating crystal method is the most accurate for the determination of crystal structure, most modifications of equipment pertain to this method. The Weissenberg goniometer is such a modification (Fig. 12). It is described briefly as follows. Cylindrical film is translated in the direction of the axis of rotation of the crystal. An aperture thus exposes for a given angle of rotation a strip on the film. The angle of reflection from the crystal then makes a spot on the strip. Thus two angles are given: the angle of rotation of the crystal is given by the strip position; the angle of reflection relative to the crystal is given by the spots' position on the strip.

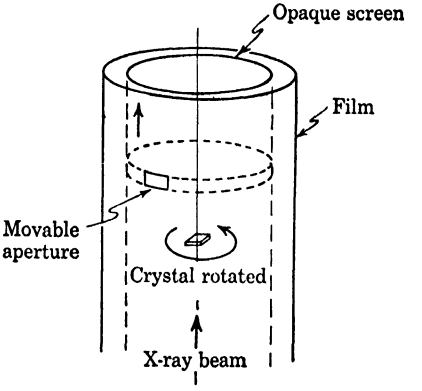


Fig. 12. Weissenberg goniometer.

Clark[2] gives a complete list of other modifications to the rotating crystal method in Chapter 13.

Focusing camera. Since the powder method requires more energy than the crystal rotation method, focusing cameras have been devised to increase the x-ray intensity at the detector.

If the pinhole or slit is located on the same circle as the sample and the film bent to fit that circle, the x-rays reflected at the angle 2θ from all parts of the sample will be focused on one part of the film.

BIBLIOGRAPHY

Clark, George L., *Applied X-rays.* New York: McGraw-Hill Book Co., Inc., 1940.

Lonsdale, Kathleen. *Crystals and X-rays.* New York: D. Van Nostrand Co., Inc., 1949.

Spiegel-Adolf, Mona and Henny, George C., *X-ray Diffraction Studies.* New York: Grune and Stratton, 1947.

Wilson, A., *X-ray Optics.* London: Methuen & Co., Ltd., 1949.

[2] Clark, George L., *Applied X-rays.* New York: McGraw-Hill Book Co., Inc., 1940.

30

LOW TEMPERATURES

INTRODUCTION. The porous plug experiment of J. P. Joule and W. Thomson in 1852 made possible liquefaction of the so-called permanent gases, oxygen, nitrogen, and hydrogen, and led to the whole field of low temperatures as we know it today. Their results are well known and can be found in any thermodynamics text, so will not be discussed here. The liquefaction of oxygen, nitrogen, hydrogen, and finally helium extended the limits of attained temperatures to approximately 1°K. Further work with paramagnetic salts cooled by adiabatic demagnetization has established the present low-temperature record of 0.0034°K.

The low-temperature region is usually understood as that below the point of hydrogen liquefaction (20.4°K), the most important section being that below the helium boiling point (4.2°K). Of primary importance is investigation of the anomalies present in specific heats, electric and thermal conductivities, viscosities, and other elastic constants, and in measuring the absolute values of certain thermodynamic functions. Measurements at low temperatures test many existing theories, and new findings are leading to further postulates.

Because of space limitations there will be no mention of the interesting phenomena that take place at low temperatures, such as the characteristics of helium II, superconductivity, second sound, etc. Instead the emphasis will be on the experimental methods by which W. F. Giauque has carried out a phase of his work, the measurement of specific heats, entropies, and magnetic susceptibilities at low temperatures.

HELIUM. Since helium is used in all cases as a refrigerant for work carried on below its boiling point, it may be of interest to know, at least to an order of magnitude, the values of some of its physical constants.

Boiling point: 4.2°K.
Heat of vaporization: 20 cal/mole.
Specific heat: 0.14 cal/g.
Viscosity (of gas): 189 micropoises at 273°K.
Viscosity (of liquid): 0.12 micropoise at 1.5°K.

Joule-Thomson inversion temperature 55°K.
Vapor pressure (4.2°K): 760 mm Hg.
Vapor pressure (0.8°K: 0.01 mm Hg.

LIQUEFACTION OF GASES. The majority of liquefiers of oxygen, nitrogen, hydrogen, and helium have three processes in common, namely compression, regenerative cooling in a heat exchanger to a point below the inversion temperature, and Joule-Thomson expansion through a valve to a reduced pressure. A simple example of this type of liquefier is the Hampson liquid air system shown schematically in Fig. 1. The incoming air is compressed in the compressor P to 150

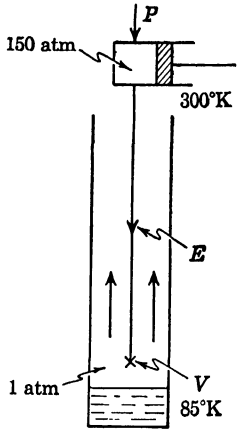


Fig. 1. Liquefier for liquid air.

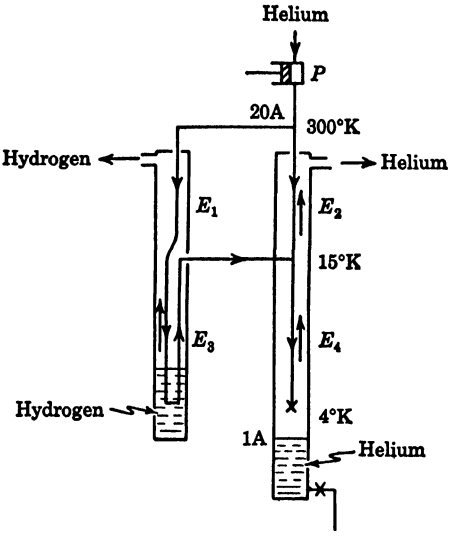


Fig. 2. Onnes' liquefier.

atm, travels through the exchanger E, and is cooled on expansion to 1 atm through the valve V. This gas then passes back through the exchanger, cooling the compressed gas as it goes. When the system has been running for some time the compressed gas at V is cooled to the temperature at which, on expansion, a part is liquefied. An equilibrium is established in the exchanger so that the returning gas gives up all its excess heat to the oncoming compressed gas.

Helium can be liquefied in any one of the following four ways.

Onnes' liquefier. Purified helium is compressed to 20 to 30 atm. (see Fig. 2), then separated into two circuits, part being sent through hydrogen cooling exchangers E_1 and E_3 and part through the helium exchanger E_2. The two parts join to enter the exchanger E_4 at 15°K,

and a fraction is then liquefied upon expansion at the valve V. The boiling point of the hydrogen is maintained at 15°K by reducing the vapor pressure above it by means of a pump. The helium reservoir and E_4 are vacuum jacketed and surrounded by liquid air. The process is continuous and can produce approximately 1 liter per hour. This is the method used in the low-temperature laboratory at the University of California.

Simon's desorption liquefier. This method and the following are examples of liquefaction without use of a Joule-Thomson expansion. Activated charcoal in an evacuated chamber C (see Fig. 3) is cooled to

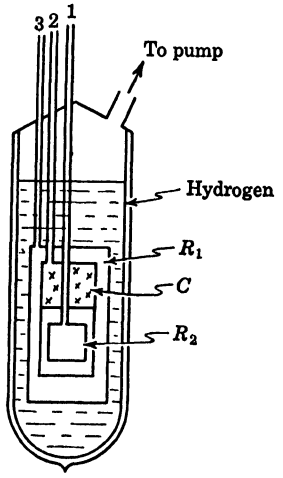


Fig. 3. Simon's desorption liquefier.

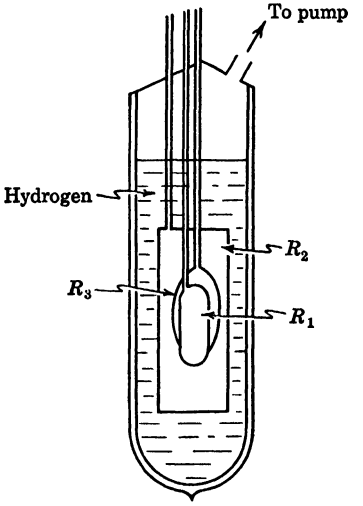


Fig. 4. Simon's adiabatic expansion liquefier.

15°K by pumping on a liquid hydrogen bath surrounding it. Heat is transferred through gaseous helium at low pressure in R_1. Helium is then admitted into C, where the carbon absorbs it, and R_1 is evacuated. The spaces C and R_2 are now thermally insulated from the outside. The helium is pumped from C, and the carbon cools by providing the energy necessary for desorption. Helium is then passed into R_2 with slight pressure and liquefies on passing through C. This method is little used except when thermal stability is desired. The temperature can be accurately and easily controlled (which is not true of other methods) by the rate of desorption. This is a discontinuous method.

Simon's adiabatic expansion liquefier. Helium is compressed into R_1 (see Fig. 4) at a pressure of 150 atm. It is cooled to 11°K by the hydrogen bath, which has a low vapor pressure produced by pumping. Gaseous helium at low pressure in R_2 conducts the heat to the bath.

When equilibrium is established R_2 is evacuated, insulating R_1 and R_3. The helium gas in R_1 is then allowed to expand adiabatically. The external work done in expanding comes at the expense of internal energy and liquefies enough helium to fill the container 85% full of the liquid. The adiabatic expansion takes several minutes. It remains adiabatic because of the high degree of insulation attained by this arrangement. This is a discontinuous method and does not achieve high productivity. The experimental apparatus is usually built into the chamber R_1, and thus being limited to small sizes, the method is restricted in use. The actual design varies according to the experiment, the liquefier being built to fit the experiment, and is desirable in that it is easily designed, built, and paid for. A temperature of 1°K can be attained by pumping on R_1 to reduce the vapor pressure. This temperature is measured by R_3, which is a thermometer reservoir attached to a manometer outside.

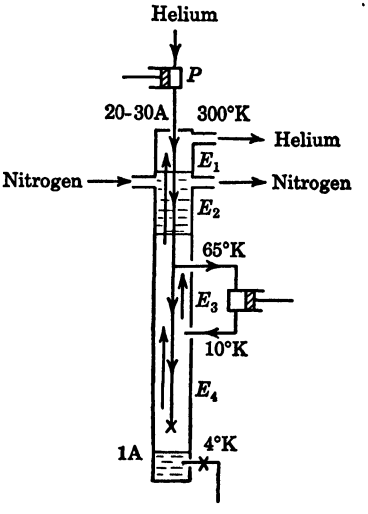


Fig. 5. Kapitza's liquefier.

Kapitza's liquefier. Helium is compressed, purified, then passed through the heat exchangers E_1 and E_2 (see Fig. 5), where it is cooled. Liquid nitrogen is used in E_2 to attain a temperature of 65°K by boiling at reduced pressure. A part of the helium is then allowed to perform work in an expansion engine and thereby cool to 10°K. This passes back through the exchanger E_3 and cools the compressed gas, which then goes through the final exchanger E_4 and is expanded at the temperature of 10°K. Part is liquefied and the remainder returns through the system of exchangers to cool the oncoming gas. This method uses no hydrogen and can produce 3–4 liters per hour after working. The chief disadvantage is that the engine must work at 10°K. No lubrication is used, and the tolerances are very close. Any impurities in the helium will freeze in the engine bearings.

HANDLING LIQUID GAS PRODUCTS. For all liquefied gases except helium a Dewar flask is capable of providing enough insulation to allow transport or use of the liquid at room temperatures for a reasonable time. However, helium has the property of diffusing through Pyrex glass. This means that the vacuum between the inner and outer layers of the Dewar will not be retained, and there will not be thermal

insulation for helium. Thus if we do not want a bulky cryostat around we must have an efficient transport system. The best method of transferring liquid helium is to use special Dewars. Jena glass is better by a factor of 1000 than Pyrex with respect to diffusion. Another gain can be made by surrounding the helium Dewar with liquid nitrogen. This reduces the temperature differential across the vacuum space. The inner wall of the Dewar should be thin; then anything in the vacuum space will be condensed on the cold inner wall, and this cleaning up creates a very efficient vacuum. The helium diffusion decreases at these low temperatures.

Another method is to use transfer tubing. Because of the low viscosity of liquid helium, small tubes may be used; the small surface area then permits effective vacuum jacketing. Helium can be transported several meters in this way. It is important to use a proper venting system when liquid hydrogen is being used. The venting hose should carry the hydrogen out of doors and be arranged with a trap or bend so that oxygen from the air cannot enter the system.

EXPERIMENT IN OUTLINE. Certain paramagnetic salts have molecular ions with magnetic moments. These moments have a random distribution at low temperatures. When a magnetic field is applied these moments line up with the field and in doing so give off heat to a liquid helium bath. The salt is then thermally insulated and the magnetic field is removed. The ions do work in the adiabatic demagnetization[1,2] that follows, and the temperature drops. The final temperature depends on the initial field, initial susceptibility, and on the paramagnetic salt. Specific heat, entropy, susceptibility, and temperature are the quantities desired in this experiment.

APPARATUS. A description of the sample and apparatus follows.

1. *Sample.*[3] The sample is an ellipsoidal crystal specimen. The shape is most important. It can be shown that this is the only shape that has a homogeneous magnetic field inside the specimen if the field is homogeneous in its absence. If the field within were not homogeneous, the different parts of the sample would cool to different temperatures. At these low temperatures the thermal conductivity is so low that equilibrium could not be reached quickly enough, and with-

[1] Debye, P., "Einige Bermerkung zur Magnetisierung beitrefer Temperatur," *Ann. Physik, 81,* 1154 (1926).

[2] Giauque, W. F., "A Proposed Method of Producing Temperatures Considerably Below 1° Absolute," *J. Am. Chem. Soc., 49,* 1864 (1927).

[3] Giauque, W. F., Fritz, J. J., and Lyon, D. N., "The Measurement of Magnetic Susceptibility at Low Temperatures," *J. Am. Chem. Soc., 71,* 1657 (1949).

out a common temperature for the sample we cannot speak of a susceptibility. The sample is mounted on Plexiglas supports. The upper support is a tube that has a small 1 cc chamber at its junction with the sample. This will be explained with the cooling system. The sample and mount are enclosed in a chamber (see Fig. 6) that can be filled with liquid or gaseous helium to form a path for heat conduction to the bath outside or that can be evacuated to insulate the sample thermally.

2. *Magnet.*[4] The experiments carried out in the low-temperature laboratory at the University of California have made use of a solenoidal air-core magnet. A high homogeneity in the field was desired

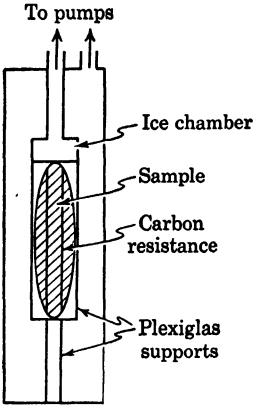


Fig. 6. Sample chamber.

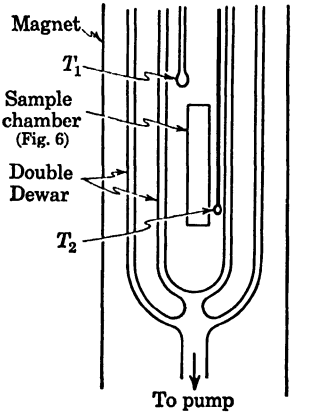


Fig. 7. Arrangement of Dewars.

and was obtainable without ferromagnetic materials. The coils are in the shape of a hollow cylinder, wound of square copper conductors in 6 independent circuits of equal resistance. Each circuit extends over the length of the magnet; they may be connected in any series or parallel combination. The coil is 91.4 cm in length, with inside diameter of 15.9 cm and outside diameter of 43.8 cm. It is enclosed in a brass case slightly bigger, which allows 15.2 cm inside for a working diameter. The cooling system utilizes kerosene, which flows through the coils from bottom to top, gaining in temperature only 4° under normal operating conditions at 50°C. The power input is 100 kw direct current, using 500 amp at 200 v. This gives a magnetic field of 8000 gausses at the center. This is homogeneous in the sample region (a

[4] Giauque, W. F. and MacDougall, D. P., "The Production of Temperatures Below One Degree Absolute by Adiabatic Demagnetization of Gadolinium Sulfate," *J. Am. Chem. Soc., 57,* 1175 (1935).

cylinder of 12 cm length and 3 cm diameter) to within 0.05% of the value at exact center. The magnet is constructed to utilize 2000 kw, which would give 25,000 gausses. The limitation is in the source of power available.

3. *Cooling system.*[5] The experiment is conducted in a double Dewar Pyrex glass, so constructed that the common vacuum space can be pumped continuously. The outside is at room temperature, the space between at the temperature of liquid nitrogen and the inner Dewar contains a liquid helium bath. Although Pyrex is permeable to helium at room temperature, any helium that diffuses through the outer wall of the Dewar is absorbed on the cold surfaces. Thus an effective vacuum is maintained without pumping. The desired temperature between Dewars is maintained by running liquid nitrogen in a brass coil wound around the helium Dewar. The coil is wound double to give both connections at the top and also to make it noninductive. (The liquid nitrogen may also be fed free into the space between Dewars.) It is supplied at a small pressure and discharges the unevaporated part from the coil into a small cup that cools the upper portion of the apparatus. The flow into this cup is regulated by an oxygen vapor pressure thermometer near the top of the cup. This opens and closes a magnetic valve that controls the pressure on the nitrogen supply Dewar in such a way as to keep the liquid nitrogen level in the cup constant. Two helium vapor pressure thermometers T_1 and T_2 (see Fig. 7) give the height of the helium bath and the temperature of the bath, respectively. This bath is pumped on by a 105 cfm Kinney pump, which lowers the vapor pressure to a value such that the helium boils at 1°K.

With 1 cc of liquid helium in the chamber mounted on the sample support and with the magnetic field on, the sample is cooled to 1°K. The enclosure around the sample is then evacuated, and with the specimen thus thermally insulated from the bath, 1 cc of liquid helium is evaporated from the small chamber. This evaporation cools the crystal still more to 0.85°K; then the magnet is cut off, and the crystal cools further by adiabatic demagnetization.

4. *Carbon resistance thermometer-heater.*[6] The low-temperature readings are made by means of a carbon resistance thermometer. Although the temperature effects of carbon are unpredictable at room

[5] Giauque, W. F., Stout, J. W., Egan, C. J., and Clark, C. W., "The Measurement of Adiabatic Differential Magnetic Susceptibility Near 1° Absolute," *J. Am. Chem. Soc., 63,* 405 (1941).

[6] Giauque, W. F., Stout, J. W., and Clark, C. W., "Amorphous Carbon Resistance Thermometer-Heaters for Magnetic and Calorimetric Investigations at Temperatures Below 1°K," *J. Am. Chem. Soc., 60,* 1053 (1938).

temperatures, in the liquid air range and lower temperatures it acts as a metallic conductor and has a large negative coefficient of resistance. A typical example is a carbon ink film which, at room temperature had a resistance of 57 ohms, but showed a value of 780,000 ohms at 1.63°K. This large resistance in the working range allows us to neglect lead resistances. This means that small external leads can be used, reducing the thermal conductivity. The carbon thermometer is mounted on the Plexiglas sample support in contact with the sample. Lampblack has been found to be a good form of carbon for this work, and the thermometer is made as follows. A layer of lens tissue paper (its loose texture gives it a small coefficient of expansion) is put on the sample with ethyl alcohol and a camel's hair brush. The saturated paper is then coated with lampblack mixed with more alcohol. This is given a protective coating of collodion. The excess tissue is cut away to leave an inverted U running the length of the sample, some 12 cm. This is then trimmed to give the desired resistance. This U shape was chosen in order to reduce the magnetic induction. A resistance of the order of 50,000 ohms is used. The strips are covered by another layer of lens tissue held on with dilute collodion. Platinum-iridium strips deposited on the glass connect the thermometer to glass knobs from which fine platinum wires run to tungsten terminals sealed through the outer glass wall.

The thermometer does not maintain its exact calibration after heating to room temperatures, but calibrations can be reproduced to 6 parts in 100,000 quite easily. The dependence of the resistance on the magnetic field is such that the corrections are easily calculated. The carbon thermometer has many advantages over other resistance type low-temperature thermometers. It has a high resistance, as stated above, and obeys Ohm's law. Where low-resistance materials such as phosphor bronze are used, large leads must be used to lower the resistance, but this raises the heat conduction of the leads. In addition, the other materials are nonohmic at low temperatures. The carbon resistance has an additional use in this experimental setup. Its high resistance is an ideal means of putting known amounts of heat into the sample by means of small currents.

5. *Fluxmeter.*[3,7] The susceptibility of the sample is obtained from a unique fluxmeter arrangement. The principle is well known. The ordinary ballistic galvanometer is not accurate enough for this work; instead a null method of balancing, using two coils in opposition, is

[7] Giauque, W. F. and Stout, J. W., "The Magnetic Flux Distribution When a Cylinder of Constant Permeability is Placed in a Homogeneous Field," *J. Am. Chem. Soc.*, *61*, 1384 (1939).

used. A double fluxmeter is set up. The magnet fluxmeter consists of an exciting coil that is long enough to establish a homogeneous field in and near the sample region, and concentric with this are measuring coils. In the magnet there are two sets of measuring coils; one coil surrounds the sample, and a system of coils below it serves as a secondary reference. These can be connected in opposition to give a nearly null reading. A more exact reference is needed, however, because of the impossibility of maintaining a standardization established at room temperature. The turn area of a coil changes with the drop to low temperatures, and calibrations of inductance made at room temperatures change by about 0.1% upon cooling. This may seem like a small change, but in the measurements involved, such accuracy can be attained that corrections must be made for steel reinforcements in the building and variations in the earth's magnetic field. This makes construction of a second fluxmeter necessary—a primary standard. Because a null method requires matching phase as well as magnitude, a duplication of metal surroundings is required. This primary standard fluxmeter is removed some 35 ft from the magnet to reduce the influence of the magnetic field. The fluxmeter has an exciting coil identical with the exciting coil in the magnet fluxmeter, and is connected in series to carry the same current. It too has a set of standard coils.

The differential adiabatic susceptibility $(\partial I/\partial \mathfrak{H})_s$ at different values of magnetic field is the quantity of primary interest, not the susceptibility, $\chi = I/\mathfrak{H}$, as such. The exciting coil is used in the magnet to give a change in flux only for zero magnetic field. For readings in the field, a 30 ohm resistor is shunted across the 0.4 ohm of the magnet coil to give a decrease of flux, and the exciting coil is not used. Under these conditions it is impossible to match the phasing of the magnet coil with the exciting coil of the external flowmeter, so in practice the primary standard, which is kept at room temperature, is used only to calibrate the secondary reference, which is in the magnet at low temperatures. The secondary reference, having the same excitation under all conditions as the sample, is used for actual measurements.

AN EXPERIMENT IN DETAIL. Because of the difficulty of maintaining low temperatures one experiment usually lasts a week, running steadily, or until the system can no longer reach the desired conditions. Thus two to four people usually cooperate on one experiment. This particular experiment was performed on copper sulfate pentahydrate ($CuSO_4 \cdot 5H_2O$). The specimen was an ellipsoidal crystal. Crystals are used rather than powder, when possible, to avoid integration over

varying magnetic moment orientations. Breaking the crystal does not matter as long as the alignment is preserved.

With the apparatus in the magnet and a vacuum seal made, liquid nitrogen was started through the cooling coil at midnight of the first day. The cooling rate was a few degrees per hour. By noon liquid nitrogen was transferred into the outer Dewar proper and the vacuum system was checked. The second morning hydrogen was compressed for use in the helium liquefier, and in the afternoon helium was liquefied. That night helium was put into the inner Dewar, and as it evaporated the Dewar was refilled. The third day the carbon thermometer was calibrated, and heat capacities and susceptibility were measured at 14 temperatures between 20.4°K and 1°K—nine points between 4° and 1° and five points between 11° and 20.4°.

The specimen was cooled to 1°K, helium was put in the 1 cc chamber, and the magnet was turned on—all this with the specimen surrounded with gaseous helium in the space between it and the bath. When the heat of magnetization was carried to the bath and equilibrium was again established, the insulating vacuum was turned on. This insulated the specimen. The 1 cc of helium was then pumped from the small chamber. When the pressure drops on the line to this chamber, the helium is gone. The temperature at this point was 0.85°K. The magnet was then turned off and the temperature dropped to 0.2°K.

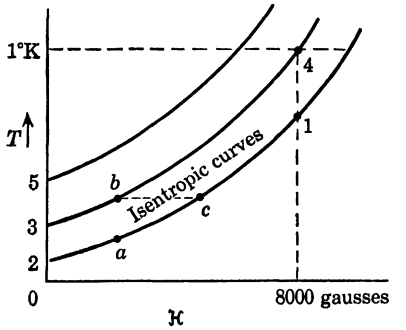


Fig. 8. Demagnetization curves.

On the graph, Fig. 8, this adiabatic demagnetization would go from point 1 to point 2. This first demagnetization is used as a check on the thermometer, and values of resistance vs. field are taken as the magnet is turned on to return to point 1. A second demagnetization is made, and this time the differential adiabatic susceptibility is measured vs. field on returning to point 1. A third time heat ΔE is added by means of the carbon thermometer to go from point 2 to 3. Readings of carbon resistance and susceptibility are taken as functions of known field intensities as the sample is magnetized on this higher curve to point 4. On further cycles ΔE is increased to change the entropy to that at 5, etc. The same readings are taken on these higher entropy curves as the magnet is energized. The measurements are repeated many times to the end of the experiment. The heat added by the carbon varies from a few thousandths to $\frac{1}{2}$ calorie.

CALCULATIONS. Measurements taken are temperature (or resistance) and differential magnetic susceptibility $(\partial I/\partial \mathfrak{H})_s$ as functions of the field. We want the specific heat and entropy difference also. The calculations follow from the first law of thermodynamics.

$$T\,dS = dE + P\,dV - \mathfrak{H}\,dI \tag{1}$$

and from the definition of enthalpy

$$H = E + PV - \mathfrak{H}\,I \tag{2}$$

The differential Gibbs function is

$$dG = -S\,dT - I d\mathfrak{H} \tag{3}$$

where the pressure term drops out because it maintained constant. From reciprocity,

$$\left(\frac{\partial S}{\partial \mathfrak{H}}\right)_T = \left(\frac{\partial I}{\partial T}\right)_{\mathfrak{H}} \tag{4}$$

To get specific heat, use

$$C_{\mathfrak{H}} = \frac{\Delta H}{\Delta T} \tag{5}$$

The temperatures and differences are given by the carbon resistance, knowing the temperature coefficient. To find ΔH between two points such as a and b (Fig. 8) it must be calculated along the path $a14b$.

$$\Delta H_{ab} = \Delta H_{s_{a1}} + \Delta H_{14} + \Delta H_{s_{4b}} \tag{6}$$

From (1) and (2)
$$T\,dS = dH + I\,d\mathfrak{H} \tag{7}$$

which is zero along the adiabatic curves. Then

$$\Delta H_s = -\int I\,d\mathfrak{H} \tag{8}$$

can be calculated from the susceptibility measurements, using

$$I = \int_0^{\mathfrak{H}} \left(\frac{\partial I}{\partial \mathfrak{H}}\right)_s d\mathfrak{H} \tag{9}$$

From (8) and ΔT the specific heat is given by (5).

The entropy between any two curves is constant for all values of temperature. It can be shown that if $I = f(\mathfrak{H}/T)$, there is no interaction between particles, as in a perfect gas, and the entropy difference can be obtained theoretically by integrating along the constant-T line, bc.

$$\Delta S = \int_{\mathfrak{H}_1}^{\mathfrak{H}_2} \left(\frac{\partial S}{\partial \mathfrak{H}}\right)_T d\mathfrak{H} \tag{10}$$

$$= \int_{\mathfrak{H}_1}^{\mathfrak{H}_2} \left(\frac{\partial I}{\partial T}\right)_{\mathfrak{H}} d\mathfrak{H} \qquad \text{from (4)}$$

This does not hold for paramagnetic salts. Instead we must use an experimental integration. From (7), at constant field,

$$dS = \frac{dH}{T} = \frac{C_{\mathfrak{H}\mathcal{C}} dT}{T} \tag{11}$$

This gives

$$\Delta S = \int_{T_1}^{T_2} \frac{C_{\mathfrak{H}\mathcal{C}}}{T} dT \tag{12}$$

An equivalent form can be obtained by using

$$\left(\frac{\partial S}{\partial \mathfrak{H}\mathcal{C}}\right)_T = -\left(\frac{\partial S}{\partial T}\right)_{\mathfrak{H}\mathcal{C}} \left(\frac{\partial T}{\partial \mathfrak{H}\mathcal{C}}\right)_S \tag{13}$$

and integrating to get

$$S = -\int_{\mathfrak{H}\mathcal{C}_1}^{\mathfrak{H}\mathcal{C}_2} \frac{C_{\mathfrak{H}\mathcal{C}}}{T} \left(\frac{\partial T}{\partial \mathfrak{H}\mathcal{C}}\right)_s d\mathfrak{H}\mathcal{C} \tag{14}$$

BIBLIOGRAPHY

Burton, E. F., Smith, H. Grayson, and Wilhelm, J. D., *Phenomena at the Temperature of Liquid Helium*. New York: Reinhold Publishing Corp., 1940.
Keesom, Willem Hendrik, *Helium*. Amsterdam: Elsevier, 1942.
Jackson, L. C., *Low Temperature Physics*. London: Methuen & Co., Ltd., 1948.

31

LABORATORY HAZARDS

RADIATION HAZARDS

EARLY HISTORY. Soon after Roentgen's discovery of x rays, the fact that ionizing radiations (which were associated with x-ray work and later with nuclear physics) could produce injuries to the body first came to the attention of the medical profession. Skin lesion abnormalities of the body surfaces, temporary and permanent baldness, were observed and reported in the literature. However, in spite of these early observations, the early pioneer researchers did not realize the insidious nature of the radiation, and consequently suffered severe injuries, as indicated by the high rate of cancer of the hands of these roentgenologists. Madame Curie herself died from cancer caused by radium poisoning; and several early cyclotron workers suffered severe eye cataracts from lining up cyclotron beams by eye. As these dangers were finally recognized, many studies were designed to set up criteria for eliminating them.

NATURE OF RADIATION DAMAGE. The exact nature of physical effects of irradiation is not as yet entirely clear. The damage is caused by the ionization arising in the interaction with tissue. Although most of the absorbed energy goes into the formation of ion-pairs, some is utilized for the excitation of atoms and the disruption of chemical bonds of complex molecules. The damage appears to be due to a large number of small energy transfers ($\approx$ few ev compared with the high-energy incident radiation).

Some of the observed general physiological consequences in the human body are:

1. Lassitude and fatigue.
2. Damage to blood.
3. Lung cancer.
4. Dermatological effects.
5. Reproductive organ damage.
6. Radio-genetic effects.

The specific ionization and ionization densities of the ionizing particles are extremely important in the evaluation of possible harmful

effects when exposed to radiation. The specific ionization is highest for alpha particles, and lowest for electrons.

TYPES OF RADIATION HAZARDS.

1. *X rays and gamma rays.* The damage is caused by secondary particles. For most practical purposes, the Compton electron effect is most important, while the photoelectric and pair-production effects are negligible, since the body does not contain much lead, tantalum, etc. For the human body, a good approximation is made by employing the well-known Compton cross section for carbon. As mentioned above, the specific ionization is comparatively low.

2. *Fast neutrons.* Damage is due chiefly to the production of recoil protons in the tissue from neutron bombardment at energies less than 50 mev. One can make approximate proton flux calculations by employing the *n-p* cross section for hydrogen. For neutron energies greater than 50 mev, higher elements are important due to star formations. The specific ionization associated with the recoils is greater than that due to Compton electrons.

3. *Slow neutrons.* The *n-p* reaction $N^{14}(n,p)C^{14}$ in the interaction of slow neutrons with the small nitrogen content of the body, is the dominant factor.

4. *Alpha particles.* These rays are highly ionizing but have small penetration. Unlike neutrons and gamma rays, which have radiation fields associated with them, alpha particles are unique as a hazard (excluding acclerator alpha beams). Because of this small range, injurious effects usually occur through ingestion, inhalation, and open wounds, although very strong alpha bombardment can harm the skin. Therefore the surface and air-borne activities in a laboratory are the significant levels. Alphas will not bring about harmful biological effects throughout the body, but rather in specific strategic regions. Once having gained access to the body, the high atomic number alpha emitters are absorbed in the bone tissue, causing great damage to the bone-forming cells. They also attack the blood-forming system with disastrous results.

5. *Beta particles.* Since these are electrons, they have essentially the same effect as x and gamma rays; exposure to a strong source of soft beta particles can, however, cause severe local damage.

6. *Protons.* The effect is essentially similar to that of the recoil protons arising from neutron irradiation.

DOSE MEASUREMENTS. Different tissues and organs vary widely in their resistance to radiation. Consequently, a dose unit based on a particular biological effect would be difficult to correlate with other

effects and with similar effects in different tissues. Hence, units based on physical effects seem to be most feasible.

From the physical point of view, the most desirable dose unit types are: (1) the radiation energy absorbed per unit mass, or (2) the ionization formed per unit mass. These types can be measured quite accurately and absolutely. Since these units can be defined independently of the absorbing medium and type of radiation (except for roentgen), they can be used in different tissues and then be correlated with chemical and biological effects. Numerous units have been proposed, each having certain advantages in specific dosimetry problems. The following represent those most commonly used.

1. *Roentgen* (r). The roentgen is that "quantity of x or gamma radiation such that the associated corpuscular emission per 0.001293 g of (dry air at STP) produces, in air, ions carrying 1 esu of quantity of electricity of either sign."[1]

It takes 32.5 ev absorbed to form one ion-pair by an electron in air. Hence, a dose of 1 r, received at any point, is equivalent to

1 esu of ion-pairs produced per cubic centimeter of air.
1.61×10^{12} ion-pairs produced per gram of air.
6.77×10^{4} mev absorbed per cubic centimeter of air.
83 ergs absorbed per gram of air

The roentgen is actually the time integral of the flux evaluated according to its ability to ionize air, and it is not simply a unit of intensity or energy flux. The relationship between energy flux and energy absorbed by the air exposed to one r depends upon the incident radiation spectrum in a complicated manner. Despite the wide variations of absorbing media, the dose is still considered as equal to one r if that same amount would produce in air 1 esu of ion pairs per cubic centimeter.

2. *Roentgen equivalent physical* (rep). "That dose of *any* ionizing radiation that produces energy absorption of 83 ergs per gram of tissue is 1 rep."[2]

In general, 1 rep does not equal 1 r since the rep is always equal to an energy absorption of 83 ergs per gram, regardless of tissue and type of radiation, while the absorbed energy under a dose of 1 r varies with the composition and density of the tissue, and, of course, is applicable to x and gamma radiation only.

3. *Roentgen equivalent man* (rem). "The rem is that dose which,

[1] Curie, M. *et al.*, "The Radioactive Constants as of 1930," *Revs. Mod. Phys.*, *3*, 427 (1931).

[2] Parker, H. M., *Advances in Biological and Medical Physics*. New York: Academic Press, 1948, Vol. 1, p. 243.

delivered to man (or mammal) exposed to *any* ionizing radiation, is biologically equivalent to the dose of 1 r of x or gamma radiation (not photoelectrically converted)."[2]

This unit is essentially a measure of the quantity of radiation that produces certain observable biological effects, rather than a measure of energy absorbed or ionization produced in tissue.

4. *Relative biological effectiveness* (rbe). The rbe is defined as the ratio of rem/rep. This term has been employed to correlate biological effects among different tissues. Hence the following relations hold

X or gamma rays:	rbe = 1 by definition.
Fast neutrons:	rbe = 5–10 from animal experiments.
Fast protons:	rbe = 5–10 from animal experiments.
Alpha particles (3 mev):	rbe = 10 from animal experiments.
Slow neutrons:	rbe = 2 from animal experiments.

DOSIMETRY INSTRUMENTATION. Since, by definition, the roentgen refers to the ionization in a region of air space, the ideal instrument to make this measurement would be a wall-less ionization chamber. The National Bureau of Standards has developed these types of chambers.[3]

Practically, however, a solid-wall chamber is easier to construct and much more manipulable. Interpretation of the measured ionization is not always obvious because of wall contributions to the observed ionization. The wall effects depend on the following factors.[4]

1. Wall material atomic number.
2. Thickness of wall relative to the range of secondary particles.
3. Dimensions of chamber relative to secondaries.
4. Attenuation by wall of primaries.
5. Possible nuclear reactions.

However, the Bragg-Gray principle enables one to use the more practical solid-wall chambers with excellent precision and interpretation of observed results (although not quite as accurate as the Bureau of Standards type).

Gray[5,6,7] has shown that for a small air cavity surrounded by a

[3] Taylor, Lauriston S. and Singer, George, "Measurement, in Roentgens, of the Gamma Radiation from Radium by the Free-Air Ionization Chamber," *J. Research Nat. Bur. Standards, 24*, 247 (1940).

[4] Siri, William E., *Isotopic Tracers and Nuclear Radiation*. 19. New York: McGraw-Hill Book Co., Inc., 1949, Chaps. 12, 13, 16.

[5] Gray, L. H., "The Absorption of Penetrating Radiation," *Proc. Roy. Soc.* (London), *A122*, 647 (1929).

[6] Gray, L. H., "The Ionization Method of Measuring Neutron Energy," *Proc. Cambridge Phil. Soc., 40*, 72 (1944).

[7] Gray, L. H., "An Ionization Method for the Absolute Measurement of γ-Ray Energy," *Proc. Roy. Soc.* (London), *A156*, 578 (1936).

medium M, the energy E absorbed in M per unit volume per second from primary radiations, is related to ionization observed in the cavity by

$$E = SJW \qquad \text{ergs or mev/cc/sec} \tag{1}$$

where E = energy absorbed per unit volume per second, S = ratio of stopping power for secondaries in M to stopping power in gas, W = average energy absorbed from secondaries to form one ion-pair, J = the number of ion pairs formed in the cavity per unit volume per second. Now

$$S = \frac{\rho_m B_m}{\rho_g B_g} \tag{2}$$

where ρ_m = density of wall material, ρ_g = density of gas, B_m = atomic stopping power of wall material, B_g = atomic stopping power of gas.

If the atomic compositions of the gas and wall are essentially similar, $B_m = B_g$. Hence

$$S = \frac{\rho_m}{\rho_g} \tag{3}$$

Equation (1) holds only for the following conditions.

1. The cavity dimensions must be less than the average range of secondaries in gas.

2. The thickness of M must be $\geq$ the maximum range of the secondaries in M.

3. Attenuation of primaries going through chamber must be small.

4. The relative stopping power is independent of velocity of secondaries.

Therefore one can readily see from these conditions that for Eq. (1) to be valid, the intensity of the secondaries in the cavity must be equal to its intensity in M when it is in equilibrium with the primaries.

Significant things to watch for with chambers following this principle are:

1. If chambers are sufficiently small, ionization per unit volume should be independent of cavity size.

2. The ionization is proportional to the gas pressure.

3. For the same primary radiation in chambers with different M, the ionization is proportional to the energy absorbed per unit volume of M and inversely proportional to the stopping power.

It is clear that chamber size is extremely significant. Theoretically, the size should be infinitesimal, since the minimum range of the secondaries is essentially zero. However, counters of the order of several millimeters do not introduce appreciable errors. Usually the largest size that still gives a constant ionization per unit volume can be used.

The accuracy of the principle also has these two factors to be considered as well as size finiteness.

1. The ionization per unit volume of M is in general not equal to the ionization per unit volume of gas.

2. The very presence of the cavity alters the energy distribution of the secondaries.

These will only be significant if the constituents of the wall and gas are markedly different (e.g., lead and air). Therefore the most accurate results can be expected when the atomic compositions of the two are essentially similar.

Chambers based on the cavity principle are applicable to all types of radiation and absorbing media.

When one is measuring x or gamma radiation doses in terms of roentgens, the so-called "air-wall" chamber is often used. Here one attempts to have M composed of constituents that give it an atomic number of ≈ 7.22 (similar to that of air). This is obviously analogous to condensing a great deal of air to make up the material of the wall.

For most accurate tissue dosimetry, the wall material and gas are made tissuelike in atomic compositions. However, for many purposes air is sufficiently similar to tissue to give good measurements. Failla at Columbia Medical School has done much research along these lines. By varying the wall thickness and extrapolating he has been able to measure the absorption of primary radiation at tissue depths before equilibrium with the secondary radiation has been reached.[8]

TOLERANCE DOSES. Medical treatment of radiation damage is, unfortunately, still in the experimental stage; hence emphasis must be placed on preventive measures. Since, in most cases, it is not possible to avoid some external irradiation while engaging in radioisotope work, one must set up a tolerance dose, which may be defined as follows: the tolerance dose is the maximum permissible dose allowable over a certain interval of time (8 hr working day or a work week) based upon statistical and biological experiments, whereby the individual receiving this dose every day or week (to the best of our present-day knowledge) will not have any detectable damage incurred during his lifetime. It is to be emphasized that there is no hazard threshold. The basis for determining the tolerance dose did not change materially for many years, and most recent revisions in safety standards have been in successive reductions of previously accepted permissible exposure levels.

[8] Moyer, at Berkeley, has used a polystyrene-lined chamber that appears to be just as effective for fast neutron tissue dosimetry.

The following is a table of tolerance dose levels based on an 8 hr day as recommended by the Atomic Energy Commission and adopted by the state of California.

Radiation	Rem	Rep	Particles per cm^2 per sec
X and gamma (2 mev)	0 05	0.05	1600 photons
X and gamma (1 mev)	0.05	0.05	3200 photons
Beta particles	0 05	0.05	40 betas
Fast neutrons (2 mev)	0.05	0.01	100 fast neutrons
Fast neutrons (1 mev)	0.05	0.01	120 fast neutrons
Fast neutrons (0.5 mev)	0.05	0 01	180 fast neutrons
Fast neutrons (0.1 mev)	0.05	0 01	500 fast neutrons
Thermal neutrons	0 05	0.025	2000 slow neutrons
α-particles	0 05	0.005	

As mentioned before, the air-borne contamination level is the most important consideration for alphas. The tolerance dose is 5×10^{-12} microcuries per cubic centimeter of air; hence, something like a few micrograms of radium can cause anemia.

It may also be mentioned that the air-borne tolerance level for beta and gamma emitters is 10^{-9} microcurie/cm^3.

RADIATION PROTECTION. Proper shielding is absolutely necessary when personnel are exposed to high-energy radiation, which is especially prevalent in the vicinity of particle accelerators. Shields must be designed with these factors in mind:

1. Attenuation of the direct radiation.
2. Emission of secondaries.
3. Scattering.

Sufficient thickness will take care of the first two. However, the scattering effect is much more unpredictable and much of it can be eliminated by adequate coverage and tight shield joints.

For large shield thicknesses, concrete is used because of its easy fabrication and comparatively small cost. Lead bricks and water tanks are also commonly employed, depending on the nature of the problem.

Despite the presence of shielding, constant monitoring is a necessity at all times. Some of the measures employed at the University of California Radiation Laboratory are the following:

1. *Area survey meters.* These are commercially built beta-gamma detectors that give constant written record on an Esterline gage.

2. *Film badges.* The emulsion response is proportional to the radiation dose. These are collected weekly.

3. *Pocket instruments.* Three forms of pocket pen-sized meters are issued to accelerator personnel daily. The first is a slow neutron boron detector. The second is an ionization chamber that is charged daily and consequently discharged by the beta-gamma radiation during the day. This chamber is then discharged through an electrometer at the end of the day, giving the dose measurement. The third is another beta-gamma counter consisting of a fiber electrometer with a scale and magnifying lens, enabling one to read his dose at all times.

4. *Air filters.* Air filters are commonly used for air monitoring. Many cubic feet of air are drawn through filter paper surrounded by a thin-walled Geiger counter. Hence, by simple calculations, one can determine the air-borne concentration in the area.

5. *Portable instruments.* Numerous commercial portable survey meters have been developed for surface activity measurements. They include thin-walled electroscopes for beta-gamma detection, air ionization chambers, Geiger counters, etc.

6. *Neutron radiation field instrumentation.* For high-energy particle accelerators, the determination of the neutron flux is most important. Fast neutrons in the range of 0.5–20 mev (most common) are detected most effectively with hydrogen-recoil proportional chambers. If the neutron flux is isotropic, it is easy to calculate, from the known n-p cross section and from the fact that the recoil proton flux is isotropic in the c.m. system, the number of protons that will be ejected. Bismuth fission chambers and stars in photographic emulsions are used for very fast neutrons. Slow neutrons can be detected employing the well-known indium-cadmium foil method, as developed by Fermi, in conjunction with standard Geiger counters. Another method measures the differential ionization between argon-filled and boron fluoride counters following zero adjustments for x or gamma radiation.

OTHER LABORATORY HAZARDS

LABORATORY HAZARDS ARE REAL. We do things in the research laboratories that would not be tolerated in industry. It is quite common to expose oneself to dangers, sometimes through makeshift experimental setups, and many times through sheer ignorance.

As science progresses, we are using more and more hazardous materials, whose dangers have too often not been recognized at the outset (e.g., beryllium poisoning). This discussion will emphasize some of the lesser known hazards, although some of the obvious and ever-present ones will be mentioned in passing.

MECHANICAL HAZARDS. Machinery or equipment should never be used without proper instruction as to operation (e.g., using a grinding

stone without eye protection, operating a drill press without clamping down the object, not tying down a high-pressure gas cylinder, etc.). The list of mechanical dangers is infinite; although most are obvious, carelessness and a lack of proper respect for them have been major causes for early receipt of insurance benefits by many a widow.

ELECTRICAL HAZARDS. The danger of electrocution is always present in the laboratory. Comparatively, this is not a great danger in industry because of the finances available to prevent electrical mishaps. Clearly, this money is not available in most laboratories. Secondly, the very nature of research lends itself to many temporary electrical setups where complete protection is often not feasible. Hence the only cure is caution and a knowledge and awareness of what one is doing.

Semipermanent and permanent laboratory installations should be well protected electrically, both from the high- and low-voltage sources. A well-trained crew is essential. Two common precautions are:

1. *Use of good ground hooks and bleeder resistances.* These should be attached to fairly flexible wire, giving them greater manipulation and less opportunity for breakage. They should be tested periodically. No high-voltage equipment should ever be touched without grounding same. Because of dielectric absorption with the consequent self-charging, high-voltage capacitors should have bleeder resistances. However, bleeders often burn out; hence a ground hook should be used on the equipment for a fairly lengthy period of time.

2. *Switch locks and interlocking relay system.* Employing a switch lock to shut off power is an excellent way of eliminating the electrical danger. In this manner the person working on the equipment knows that the voltage is turned off; it also eliminates the possibility of someone turning on the power at the wrong time. The interlocking system, as a whole, prevents power from being turned on until all personnel are out of potential danger zones and all parts of the circuit are ready to function properly. The interlocking system is usually equipped with ample time delays.

POISON HAZARDS. The vast increase in the use of chemicals in the laboratory has made a study of these chemicals and their effects on health most important. A few representative poisons are here discussed.

1 *Metals.*

(a) Mercury,[9] because of its low evaporation temperature, when spilled in an unventilated room, can bring up the concentration in air beyond the safe level in a relatively short time. Mercury is a particularly insidious poison in physics laboratories. Research rooms used by generation after generation of

[9] Hardy, Harriet L., "Mercury Poisoning," *Physics Today*, Nov. 1949.

graduate studies nearly all become contaminated from mercury spilled and left to gather in cracks in the floor. All physics laboratories should have routine checks on mercury vapor in research and shop facilities. It is possible to decontaminate a room by removing all mercury possible with a high-velocity water-type vacuum cleaner, washing the floor with a soap solution containing one-half ounce of calcium polysulfides per gallon, filling cracks in the floor, and varnishing the floor covering. This procedure seals the mercury in and prevents its evaporation. Routine checks on the mercury vapor concentration should be continued. Mercurialism causes throat discomfort, severe headaches, and alimentary canal disturbances. The allowable concentration is 0.1–0.15 mg/m^3.

(b) Lead gives rise to plumbism. The allowable concentration is 0.1–0.15 mg/m^3.

(c) Uranium. Even disregarding radioactivity, uranium can be permanently locked in the body, ultimately causing great damage.

(d) Chromium. Continual welding in a poorly ventilated region may cause chromium poisoning. It has the following toxic effects: ulcers, nasal tissue damage, severe kidney and intestinal damage. The permissible concentration is 0.1 mg/m^3.

(e) Selenium is a common constituent of photoelectric cells. Selenium poisoning gives rise to offensive breath, progressive anemia, loss of weight, abdominal pains, liver destruction, and kidney disturbances. The allowable concentration is 0.1 mg/m^3.

(f) Tellurium has toxic effects similar to selenium but is more dangerous. It causes offensive breath, nausea, constipation, and somnolence. It has a safe level of 0.01–0.1 mg/m^3.

(g) Cadmium presents a constant danger in thermal neutron research. Because of its high capture cross section, it is in constant use as a slow-neutron absorber. Cadmium is a poison analogous to arsenic and mercury. It has a safe level equal to 0.1 mg/m^3.

(h) Beryllium,[11] which is often used in accelerator bombardments, has been recognized as a poison only recently. A level of 25×10^{-6} g/m^3 can cause acute lung disease. There appears to be extreme variability in human susceptibility. Beryllium poisoning is also a danger in the zinc beryllium silicate in fluorescent lamps. The oxide, fluoride, hydroxide, and finely powdered beryllium powder are probably the most toxic. The beryllium salts have caused ulcers and dermatological damage, wherease the oxide, fluoride, and powder have damaged the lungs, giving rise to an acute disease resembling bronchitis and pneumonia. There is often a long latent period before development of the disease, usually triggered by a severe respiratory illness, or general fatigue. At present there are no therapeutic measures or methods of elimination of beryllium from the body. Contact should be minimized, if not avoided completely. As yet no dose limits have been set. Beryllium must be machined wet in a specially shielded ventilated box, utilizing a filtration system to catch the exhaust of beryllium particles.

2 *Solvents.*[10] The toxicity of organic solvents depends upon solubility, vapor pressure, volatility, chemical activity, concentration, etc. As a group these are hazardous in two ways: solvents are usually either toxic or explosive and some are both toxic and explosive.

[10] Hasterlik, Robert J., "Beryllium Poisoning," *Physics Today*, *2*, 14 (1949).

[11] Hardy, Harriet L., "Hazards of Common Solvents," *Physics Today*, July 1950.

(a) Aromatics. This group consists of benzine, toluene, xylene, and similar solvents. These are extremely toxic and the dose is cumulative. In the case of acute poisoning, symptoms range from slight dizziness to convulsions and death. For chronic exposure the toxicity varies from nausea to hemorrhages and nerve cell damage. The safe level is approximately 50–200 parts per million in air.

(b) Halogenated hydrocarbons (carbon tetrachloride, trichloroethylene, dichlorobenzene). They damage the liver, kidneys, affect the nervous system, and act as a heart depressant. Carbon tetrachloride is readily absorbed through the skin. It has a cumulative effect that is similar to diabetes symptoms. One should never use a carbon tetrachloride fire extinguisher on a fire in a small enclosed region, since highly poisonous phosgene and chlorine are liberated. The safe level is approximately 50–200 parts per million.

(c) Alcohols (esters, ketones [acetone]). These are somewhat less toxic, with safe levels at 200–400 parts per million. Alcohol is an irritant and anaesthetic. The esters can cause lethargy, loss of appetite, etc. Prolonged inhalation of acetone can cause irritation of the mucous membranes of the eye and the respiratory tract, nacrosis, skin irritation, and diabetic symptoms.

(d) Paraffins (gasoline, petroleum, naphtha, etc.). These also, are less toxic than those above, having allowable concentrations of 500–1000 parts per million. They are, however, quite explosive and inflammable. Gasoline vapors are not too poisonous except that too high a concentration will act as an asphyxiant. Naphtha can cause vertigo, nausea, irritation of the skin, and even mild intoxication.

EXPLOSIVE HAZARDS. As mentioned above, many of the less toxic solvents are explosive. For example, a 2.5 to 13% concentration of acetone has an ignition temperature of 500°C. Ethyl ether is an explosive hazard with concentrations from 1.8 to 36%, with the low ignition temperature of 180°C. Also, with this solvent, oxidation products occur that may set it off spontaneously.

Hydrogen, of course, is explosive in almost anything. Concentrations of 5 to 80% are highly explosive, with a 500°C ignition temperature. Hence, for example, one should remember that glassblowing a hydrogen system presents a true hazard.

Another possible source for an explosion in a research laboratory is the accidental dropping of hot combustible objects in liquid air, which is mostly liquid oxygen. This would be quite dangerous in a vacuum system if a bit of diffusion pump oil somehow leaked into the liquid air trap. Hence it is wiser to use liquid nitrogen.

BIBLIOGRAPHY

Brandt, A. D., *Industrial Health Engineering.* New York: John Wiley & Sons, Inc., 1947.

Cantril, S. T. and Parker, H. M., *A.E.C. Publication MDDC 1100.*

Duggar, B. M., *Biological Effects of Radiation.* New York: McGraw-Hill Book Co., Inc., 1936, Vol. 1.

Evans, R. D. in *Science and Engineering of Nuclear Power*, Clark Goodman, ed. Cambridge: Addison Wesley Press, 1949.

Goldsmith, H. H., "Bibliography on Radiation Protection," *Nucleonics*, *4* (June 1949).

Jacobs, M. G., *Poisons, Hazards, Solvents*. New York: Interscience Publishers, 1949.

Lapp, R. E. and Andrews, H. L., "Health Physics," *Nucleonics*, *3* (Sept. 1948).

Lapp, R. E., and Andrews, H. L., *Nuclear Radiation Physics*, 2d ed. New York: Prentice-Hall, Inc., 1954.

Parker, H. M., *Advances in Biological and Medical Physics*. New York: Academic Press, 1948, Vol. 1, p. 243.

Siri, William E., *Isotopic Tracers and Nuclear Radiation*. New York: McGraw-Hill Book Co., Inc., 1949, Chaps. 12, 13, 16, 19.

32

DESIGN OF EXPERIMENTS

It must be emphasized that one cannot be taught how to design experiments. Only the use of experimental techniques and the making of measurements can be taught. To acquire proficiency in planning experiments it is necessary to go through the personal experience of designing many. To this end it is highly recommended that when one hears, for example, of a proposed new particle with specified properties or a new mechanism in a process, he design an experiment to detect it and determine its properties. Proficiency is then a matter of practice. One thing that builds confidence and accelerates such a program is acquisition of a large number of odd bits of information such as, for example, the number of seconds in a year, the number of charged particles contained in a microampere, flux density at the center of a pile, sensitivities of various counters, number of charged particles in a coulomb, the expansion coefficient of steel, the maximum magnetic flux in a piece of iron, etc. An experimenter should aim at the start to develop mentally such a list, having it always as ready as the multiplication table.

In starting out to plan experiments, of course, one falls into all the usual traps that result in poor designs. It is well to have these experiments examined critically by someone experienced in planning experiments, who will in most cases point out at once the things forgotten or omitted in the design consideration. But from these early failures valuable experience can be gained, and probably more physics can be learned in this way than from many formal courses of instruction.

As an example, in looking at the design equation of a cyclotron, realizing that present machines weigh several thousand tons, one sees that the energy is related to the $H\rho$ of the particles, and notes that with enough field the machine could be reduced to desk size. Exploring this possibility, one goes through the process of eliminating the iron; just send more current into the cells, not much copper is required; pulsing the machine could reduce the heating. One could get a million gausses with no trouble at all, although the forces on the turns would make it difficult to hold together. Cooling would require too much water; it would in fact require about a car load of liquid air per minute. But just as the scheme begins to look rather hopeless, one

thinks of superconductivity. On further investigation, the huge size of the liquid hydrogen plant required might be worth while in view of the great reduction in size of the machine. However on pursuing further the problem of superconductivity, one inevitably finds a characteristic curve (Fig. 1) representing the field at which superconductivity quenches. The maximum field extrapolated to absolute zero is of the order of several thousand gausses, and there the trail ends. Nearly every one starting out to design high-field apparatus goes through the process of designing nonferrous electromagnets, only to reach a blind alley of this sort. But in the process of such an investigation much physics will be learned.

Experiments in physics may be roughly divided into two categories: (1) precise measurements of physical quantities, and (2) exploratory experiments. In the first type the experimenter asks himself the question: What is the value of a certain physical quantity? and then goes about trying to measure its value. In the second type generally the experimenter is looking for new phenomena, and developing qualitative ideas about fairly large areas of information. The two types of experiments have somewhat different design considerations, and we shall treat them separately.

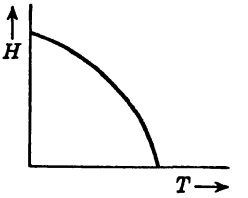


Fig. 1. Magnetic field vs. temperature.

Measurements of high precision. In planning an experiment to measure a physical quantity accurately, one needs to know how to measure certain parameters very accurately. The following are some points to bear in mind that are infrequently stressed.

Time measurements are available (free of charge) to any degree of accuracy that one might need by means of the time signals broadcast from station WWV 24 hours a day by the Bureau of Standards. The signals are good anywhere in the United States to 1 part in 10^7. For 1 part in 10^8, the motion of the ionosphere for Doppler correction may be ascertained by a telephone call to the Bureau. This service is only a few years old but is of great value.

Distance measurements can be based on interferometer methods, using cadmium red, mercury-198 green, or certain of the krypton lines, good to $\frac{1}{10}$ wavelength, with little difficulty. Moreover it is not necessary to count fringes as Michelson did. With a Fabry Perot interferometer one can take pictures of interference patterns of several lines whose wavelength is already accurately known, and from a set of simultaneous equations solve for the actual spacing. The procedure is good to 1 part in 10^7.

Mass can be measured to 1 part in 10^6 with a calibrated set of weights, though few present-day experiments in physics are based on direct measurement of mass.

For electrical measurements the problem is more difficult.

Voltage. A carefully handled standard cell is good to 1 part in 10^4, but for 1 part in 10^5 it becomes necessary to transport carefully a half dozen or so cells to the Bureau of Standards for calibration, keeping them always in an upright position and reasonably free from thermal and mechanical shock en route.

Magnetic fields can be measured at present (but only recently) to great accuracy in terms of the magnetic moment of the proton. By measuring the resonant frequency of protons, for example, in a piece of paraffin in terms of the time signals from WWV, measurement of a magnetic field is essentially reduced to a measurement of time.

Current. Measurement of current is limited by a voltage measurement plus the additional difficulty of a resistance measurement.

It is at once apparent that an accurate experiment must ultimately be based on a measurement of time, distance, or magnetic field.

There is an important factor in the making of accurate measurements that Dunnington[1] points out very clearly in his paper on the measurement of e/m. If even an approximate a priori knowledge exists in the experimenter's mind of what value of the quantity sought (say e/m) corresponds to a characteristic datum of the apparatus, proceeding with the experiment might be a complete waste of time. For the value sought can, in those circumstances, hardly be expected to differ greatly from the knowledge possessed before. Note how strikingly this subjective effect is reflected in the history of physics in the apparent secular variation of the velocity of light. Each successive determination was lower than its predecessors, but fell within the probable error of its neighbors. It even seems to exhibit typical damped oscillations about some asymtotic value that it is apparently approaching.

Dunnington said that he refused to spend five years of his life making measurements of a value he could guess beforehand. The only solution to this was not to know the answer. His experiment consisted in obtaining a minimum of intensity of electrons spiraling around through some 300° in a magnetic field. He spent two years drawing wire through his own special dies so that he could more accurately determine the magnetic field (commercial wire was too elliptical in cross section). He then gave his machinist specific instructions to make the entire apparatus to as close tolerances as possible except

[1] Dunnington, Frank G., "A Determination of e/m for an Electron by a New Deflection Method," *Phys. Rev.*, *52*, 475 (1937).

that the angle between the source and receiving slit was to be about 60°, and he asked his machinist not to measure this angle accurately. He then proceeded with his experiment, expressing his data in arbitrary units that depended on an accurate measurement of the angle. Not until after he had thoroughly investigated all possible systematic errors and written a complete report in typewritten form ready to send to the *Physical Review* did he dismantle his entire apparatus, exposing for measurement that part containing the critical angle. Only then, after an accurate measurement of the angle, was he able to complete his calculations of his value of e/m.

METHOD OF DIFFERENTIAL MEASUREMENTS. One would hardly attempt to measure the distance between two marks on a desk in Berkeley by surveying the distance from New York City to one mark, then back to New York from the other, getting the distance by subtraction. In this sense, a direct measurement of the distance is a differential measurement experiment.

E. Segrè[2] has recently employed the differential method to the measurement of the change in half-life of the Be^7 nucleus produced by varying the electron density in its neighborhood. This effect is expected in Be^7 since it decays by K capture, and the probability for decay in unit time depends both on how anxious the nucleus is to receive an electron and also on the availability of electrons to be received. (A bare Be^7 nucleus in free space with no electrons anywhere near it should not decay at all.) By comparing the rate of decay of Be^7 with that of Be^7O or the oxide with the fluoride, an effect of the order of 1 part in 10^4 in half-life should be expected. Now the direct measurement of half-lives to much better than 1 part in 10^2 is never attempted, since it would involve good counting statistics for each measurement of intensity carried out over a great number of half-lives, and maintenance of linear detection efficiency over such a wide range of intensities is prohibitive. Obviously then, a differential technique is called for whenever the difference of the half-lives is the quantity directly measured.

Two ionization chambers were employed in a balanced electrometer bridge in which differences in the ionization current between the two chambers appeared as deflections of the electrometer (Fig. 2). A plot of the unbalance will give a hyperbolic function of time if there is a difference in the decay rates. A necessary precaution in this arrangement is to make the detection geometry reproducible to a high degree, since it is necessary to interchange samples from time to time to balance

[2] Segrè, E. and Wiegand, C. F., "Experiments on the Effect of Atomic Electrons on the Decay Constant of Be^7," *Phys. Rev.*, *75*, 63 (1949).

out the effect of differences in sensitivity of the chambers and strength of the sources. To ensure this reproducibility, the samples were introduced into re-entrant hollow electrodes on the axes of the cylindrical chambers (Fig. 3). Now the $1/r^2$ dependence characteristic of an external source is replaced by an essentially even function of position that is parabolic for small displacements off the axis and reduces the placement error to second order. The batteries used must of course be

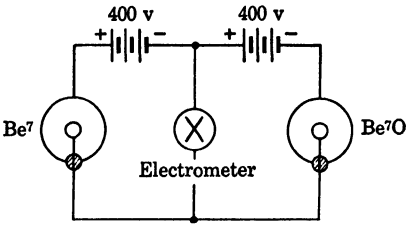


Fig. 2. Balanced electrometer circuit.

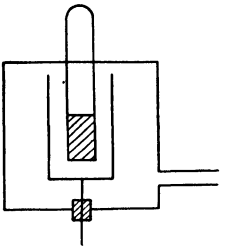


Fig. 3. Arrangement of sample in electrometer.

good, although again the effect of change in battery voltage is second order, since the chambers are loaded down with activity and only saturation currents are drawn. Moreover it is the amount of charge that enters the system possessing capacitance that determines the voltage. Another precaution is to be certain that no radioactive sources external to the system are brought accidentally anywhere near, as this would produce unequal currents in the two chambers by the $1/r^2$ factor. Common gas supply assures equal pressure in the two chambers.

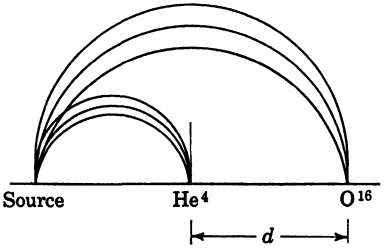


Fig. 4. Mass spectrograph orbits.

Measurements over a period of several months resulted in the actual observation of the effect. Indeed this arrangement yielded better values of the half-life (57 days) of Be^7 in half an hour than other workers were able to achieve through conventional measurements over a long period of time. For half-life measurements, comparison is made with a uranium or a radium source. One sees at once the great power of this differential method, but its use or need in some cases may not be obvious.

The atomic weight of He^4 is found to be 4.00390 by comparison with O^{16} = 16.000,000. In a mass spectrograph one might think that the mass could be determined by measuring the distance between the He^4 and O^{16} lines to 1 part in 10^6 (Fig. 4). But for the distance to have

significance it would be necessary to know the field everywhere in the neighborhood of He^4 and O^{16} paths to 1 part in 10^6. Actually, O^{16} is charged four times, giving an effective mass of 4.00000, since it is the value of e/m that determines $H\rho$. The effect of the different number of electrons carried is a higher order correction since it amounts to one electron on helium, and four electrons on O^{16}. Since one needs to measure $\Delta\rho/\rho$ to 1 part in 100, it is sufficient to measure ρ and $\Delta\rho$ separately to 1 part in 100, and the uniformity of the field is second order. Now ρ might be 20 cm, $\Delta\rho$ 0.5 mm = 500 μ, easy to measure to 1 part in 100. This gives for the mass difference 3.90 ± 0.04 millimass units, in other words, a part of measurement to 1% accuracy obtains an answer good to 1 part in 10^6.

An example similar to the above is the determination of nuclear reaction energies[3] which, known within a factor of 2, establish isotopic masses to 6 significant figures by comparison with the mass of some other isotope already well established. For example, Bainbridge gives 3.017029 for the mass of He^3 by an accurate measurement relative to 0^{16}. By measuring to within a factor of 2 the upper limit of the spectrum of H^3, one can calculate from the known mass of He^3 the mass of H^3 to one part in 10^6. That is, publishable results can be obtained (based of course on Bainbridge's work) by a relatively simple measurement of a quantity to within a factor of 2.

Accurate measurements of time are based on the differential method. Measurement of frequency differences of the order of 1 part in 10^{12} is possible. Consider two oscillators generating 10,000 megacycles $\approx 10^{10}$ c; let one put out 10^{10} c and the other $10^{10} + 0.01$ c; together these will produce beats in a crystal mixer every 100 sec, so that the difference frequency is 1 cycle in 10^{12}. Either frequency alone could not be determined to better than 1 part in 10^8.

INTENSITY VS. RESOLVING POWER. This is perhaps the most important consideration in planning experiments. By resolving power is here meant the accuracy with which one measures a particular quantity. Intensity and resolving power are both needed in all experiments, but they are always complementary. More intensity means that more signal can be discarded to give better resolution for a given ratio of signal to noise. No one is ever quite content with available intensity. At the Almagordo bomb test, some experiments had almost enough neutrons to lift the equipment bodily to one side so to speak, yet some felt that better results could have been obtained if greater intensity had been available.

[3] Tollestrup, A., Fowler, W. A., and Lauritsen, C. C., "Nuclear Mass Determinations from Nuclear *Q*-Values," *Phys. Rev.*, *78*, 873 (1950).

A striking example of the conflict of intensity and resolution was shown at Chicago during World War II in a talk on a proposed neutron velocity selector that was planned to give excellent resolution up to 6000 ev, far beyond any existing. The principle of its operation was based on the time of flight along a path length L between two perforated wheels rotating at high speed. High resolution dictated small holes, and high neutron velocities required high velocities of the holes and long path length L. Here another odd bit of information comes in that is worth remembering. There exists a sharp upper limit to the peripheral velocity of a rotating structure without bursting. This is 10^5 cm/sec, based on strength of steels available at present, and the experiment thus requires L to be large. Now if V is the radius of the hole, the intensity transmitted,

$$I \alpha \text{ (neutron flux)} \left(d\Omega \approx \frac{r^2}{L^2}\right)\left(\frac{dE}{E}\right)(r^2)$$

the last factor r^2 coming in as a result of the fact that the source area is defined by the size of the hole in the first wheel. The machine looked excellent on paper until a calculation showed the intensity to be less than the leakage flux transmitted through the shielding wall around the Hanford pile. The designer had evidently completely forgotten that it would be necessary to obtain more intensity through the apparatus than background around it. He had taken a fourth power loss in intensity to achieve the high resolution. Clearly, a compromise must always be made between intensity and resolution.

The first mass spectrograph that demonstrated the existence of isotopes was built by J. J. Thomson.[4] As an ion source he used a discharge tube with a small hole drilled through a thick cathode, out of which emerged a collimated beam of positively charged ions whose velocities depended on the position in the discharge at which they were formed. The beam then passed through a deflecting region in which it was acted upon by parallel electric and magnetic fields. In this arrangement all particles with the same e/m will strike a screen perpendicular to the beam along parabolic loci, with the position on the parabola being determined by the velocity that the ion possessed on entering the deflecting region. To obtain this form of deflection with parallel deflecting fields, it is necessary for the resolution to depend on e/m only, and that all the ions enter the field at a single point, so that actually the width of the parabolic line described on the screen depends on the size of the collimating hole or canal in the cathode. Note that effectively the canal consisted of two holes, the entrance and exit apertures.

[4] Thomson, J. J., "Cathode Rays," *Phil. Mag.*, *44*, 293 (1897); *Nature*, *90*, 645 (1913).

With this apparatus Thomson was just barely able to resolve two of the isotopes of neon by the Rayleigh criterion. To detect the third isotope the diameter of the beam would have to be reduced by a factor of 10, and this reduces the intensity by a factor of 10^4, since whenever a pair of holes is used for collimating a beam, the intensity varies as the fourth power of the diameter. The area of discharge accepted varies as r^2, and the solid angle of the second hole puts in a second factor r^2. A pair of holes is a very poor way to obtain high resolution but in Thomson's case was demanded by his use of parallel fields. Present-day mass spectrographs are designed so that slits can be used in defining the beam. The intensity transmitted by a pair of slits varies as d^2, where d is the width of the slit.

Another example contrasting holes vs. slit geometry is seen in the experiment of Stern and Gerlach compared with that of Rabi and

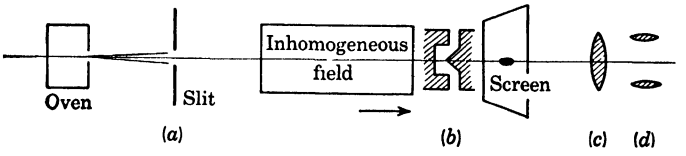


Fig. 5. Molecular beam arrangement of Stern.

others. By producing a deflection of beams of nonionized molecules possessing magnetic moment with an inhomogeneous magnetic field, Stern and Gerlach[5] were the first to demonstrate the effect of space quantization. The experiment was based on the net force on a magnetic dipole in an inhomogeneous magnetic field. Their arrangement (*a*) employed an oven for silver atoms, two defining holes, and a strongly inhomogeneous magnetic field (Fig. 5). The symmetrical form (*c*) of the deflected spot was proof of a space quantization, but the apparatus failed actually to resolve the two separate beams (*d*) because the region of strong inhomogeneity was limited in its lateral extent by the form of the pole pieces (*b*). Their work was based on the atomic moment of silver. Much later, when Rabi[6] undertook to measure accurately nuclear moments (expected to be of the order $\frac{1}{1836}$ as large) he was faced with obtaining over 1000 times the resolving power. He solved the problem, not by trying for the "world's most inhomogeneous magnetic field," but by a compromise in which by using a rather simple cylindrical geometry for the pole pieces, he obtained an extended

[5] Stern, Otto, "Ein Weg zur experimentallen Prüfung der Richtungsquantelung im Magnetfeld," *Z. Physik*, *7*, 249 (1921); *8*, 110 (1921); *9*, 349 (1922).

[6] Rabi, I. I., Millman, S., Kusch, P., and Zacharias, J. R., "The Molecular Beam Resonance Method for Measuring Nuclear Magnetic Moments," *Phys. Rev.*, *55*, 526 (1939).

ribbon-shaped region between the pole pieces in which the inhomogeneity was essentially uniform (Fig. 6). It was possible with this magnet to use slits rather than holes, and moreover to reduce their size to the order of 0.001 in., placing them 1 m apart. Thus he obtained resolution by greatly reducing r^2 but got intensity back by using slits rather than holes.

Zacharias[7] used a differential technique applied to the Rabi principle for a measurement of the moment of K^{40} which is present in natural

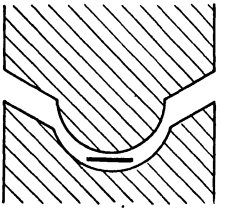


Fig. 6. Pole pieces in Rabi's apparatus.

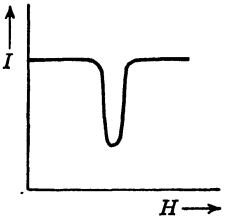


Fig. 7. Intensity of beam near resonance.

potassium in the ratio $1/10^4$ to the more abundant nonradioactive isotope. Thus from an intensity standpoint this experiment is much more difficult than measuring the moments of common nuclear species. Rabi's measurement involved looking for a minimum of transmitted intensity as the field is varied through the resonant condition (Fig. 7). But Zacharias could not measure a dip in intensity only 10^{-4} as deep, so he solved the problem by effectively turning Rabi's system over (Fig. 8). By reversing the direction of the inhomogeneity in the second

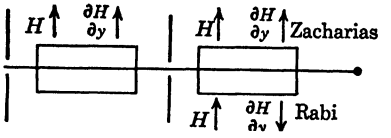


Fig. 8. Arrangement of magnetic field inhomogeneities.

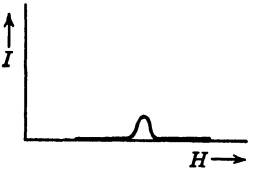


Fig. 9. Intensity of beam near resonance.

field region in Rabi's well-known arrangement he could measure a sudden increase in intensity above background as the resonant condition is passed through (Fig. 9).

More recently Zacharias has still further increased the resolution of his apparatus for use in the measurement of moments of radioactive samples of relatively short half-life, e.g., Na^{22} (3 yr). For this purpose

[7] Zacharias, J. R., "The Nuclear Spin and Magnetic Moment of K^{40}," *Phys. Rev., 61,* 270 (1942).

the beam is brought out of the modified Rabi apparatus onto a heated tungsten plate (Fig. 10) which ionizes the beam so that it may be analyzed by a mass spectrometer using an electron multiplier tube as detector. This arrangement makes use of nearly all the tricks available, and its success depends on thorough familiarity with each step along the line.

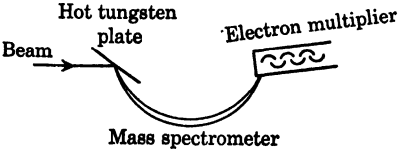


Fig. 10. Zacharias' arrangement with mass spectrometer and electron multiplier.

Dimensional analysis. Estimation of factors that go into the design of an experiment is illustrated by some examples of Alvarez's experiments. Alvarez suggests that if the experiment cannot be estimated on the back of some old envelope, "then you're not really in business."

Isotopes H^3 and He^3 were first observed only in cyclotron bombardments in, for example, the reaction

$$D + D \rightarrow H^3 + p$$

or

$$D + D \rightarrow He^3 + n$$

but it was not known which of the isobars was the heavier. It had been more or less believed that $H^3 < He^3$ from energetics, the simple theory of nuclear structure, and because people said so. Rutherford had looked for naturally occurring H^3 in tailings from the heavy water plant in Norway, making essentially a mass spectrographic analysis of 1 cc of water representing a concentration from thousands of tons and found no H^3 at all. It then occurred to Alvarez that it might be possible to make enough H^3 or He^3 to work with in the cyclotron. The plan was to bombard, say, a liter bottle of D_2 gas in the 37 in. cyclotron and then quickly admit the gas into the ion source of the 60 in. cyclotron, setting the magnet current in the latter for resonance at the $e/m = \frac{2}{3}(e/m)_\alpha$, and look for a beam of He^3. An estimate of the over-all yield expected goes as follows

Conservative average beam 37 in. machine: 50 μa.
Yield (estimated): 10^{-3} to 10^{-4} atom He^3 per deuteron.
Bombard one hour: 50 μa-hr $\approx$ 50 μa·$6(10)^{12}$ particle/μa·3600 sec/hr yields $\approx 10^{18}$ deuterons.
Total yield He^3: $10^{18} \cdot 10^{-4} \approx 10^{14}$ atoms He^3.
Concentration He^3 (for one liter target gas at one atmosphere):

$$10^{14} \text{ atoms } He^3 \div \frac{0.6(10)^{24} \text{ atoms/mol}}{22 \text{ liter/mol}} = 3(10)^{-9}$$

Typical α current: 10 μa.
Current of He^3 with a target gas as source: 10 $\mu a \cdot 3(10)^{-9} \cdot 6(10)^{12} = 2(10)^5$ He^3/sec.

Now, of course, $2(10)^5$ He^3 atoms/sec is more than sufficient to block any counter completely, and the experiment was to proceed by varying the field through the mass three condition and looking at the output of the counter with an oscilloscope.

At this point it was realized that no one had ever bothered to look for beams at the condition for mass three from either deuterium or helium gas sources. Such a beam should be expected from either a pure helium source or a pure deuteron source depending on whether He^3 or H^3 is the stable isotope. The idea was tried out with He^3 on the 60 in. cyclotron (before the machine became radioactive).

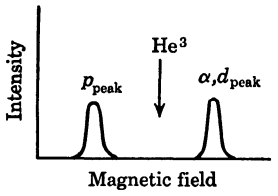


Fig. 11. Intensity of cyclotron beam.

At first nothing was observed as the field was varied between the proton and alpha-deuteron peaks (Fig. 11). This was to be anticipated from a knowledge of the structure and behavior of the 60 in. machine. The pole pieces are shimmed to raise the field at large radii over the saturation value of flat pole pieces, but the shimming is appropriate only for fields suitable for alphas. By turning the field off suddenly, after it had been set for alphas, large eddy currents develop that hold back the field at the center, relative to the outside, effectively shimming the machine momentarily for resonance of intermediate mass particles, and this procedure was successful[8] in demonstrating for the first time the existence of the stable mass three isotope of helium. Note that success of this experiment depended on a knowledge of the properties of decaying magnetic fields in shimmed cyclotrons, and this knowledge just happened to be available at the time through someone else's experience. (This experiment is a fine example of the singular importance of being alert to the significance of odd bits of information.)

The half-life of the neutron is supposed to be of the order of 15 to 30 min, according to an application of the Sargent relation to the neutron-proton difference.

$$N \rightarrow P + \beta^- \quad (800 \text{ kev})$$

In the cyclotron one can obtain neutron flux densities of the order of $2(10)^5$ neutrons per square centimeter per second (flux density is nv) but $v \approx 2(10)^5$ cm/sec, so that the neutron density n is 1 per cubic centimeter, of which nV/τ decay per second (V is volume, τ is mean life

[8] Alvarez, Luis W. and Cornog, Robert, "He^3 in Helium," *Phys. Rev.*, *56*, 879, 613 (1939).

$\approx$ 20 min). Thus in a 1 l bulb, the number decaying per second is

$$\frac{1 \cdot 1000}{1000} \approx 1 \quad (20 \text{ min} \approx 1000 \text{ sec})$$

That is, one must detect one electron per second with the cyclotron going full blast, a difficult task.

Before World War II Alvarez started out to detect not the electrons but the recoil protons. For this it was planned to use a large 15 ft evacuated pipe running well outside the shielding wall. The 400 ev protons were to be conducted down the pipe in spirals produced by a solenoid winding on the pipe. With this arrangement the protons may be transmitted with slight loss into a remote region where the background is reduced by the $1/r^2$ law. Upon nearing the end of the pipe, the protons were to have been accelerated through a thin electrostatic lens into a proportional counter. Although the war prevented the actual carrying out of this experiment, its design illustrates an important point: be sure that an experiment is designed to discriminate against background. In this experiment various tests would have had to be devised to make sure the protons had come from neutron decay. Such tests would be: changing the solenoid current; changing the pressure in the tube; use of cadmium shields to reduce the neutron density.

With the development of the pile during the war, very large fluxes of neutrons have become available and make the pile a more suitable neutron source for investigating the neutron half-life. Bottle experiments have been devised independently by many. By leaving an evacuated flask in a pile for a year, one should be able actually to measure pressure of the accumulated hydrogen resulting from neutron decay. For a pile the flux density $nv = 10^{13}$; hence the density $n \approx 10^8$; total neutrons in the 1 l bottle would be 10^{11}. The number decaying per second is 10^8. The number per year would be $3\ (10)^7$ sec/year$\cdot 10^8$ decay/sec $= 3(10)^{15}$ per liter $= 3(10)^{12}$ per cubic centimeter. One atmosphere contains $0.6(10)^{24}$ atoms/mol$/2(10)^4$ cc/mol $= 3(10)^{19}$ atoms/cc; hence the above number of protons would be $3(10)^{12}/3(10)^{19} = 10^{-7}$ atm $= 10^{-4}$ mm Hg, certainly a measurable pressure. But fast neutrons in the pile knock protons out of anything, so that not all the observed protons would come from neutron decay. Suppose the flask were copper, 63 g = 1 mole. The fast neutron flux density nv might be 10^{10} and a cross section of 10^{-3} barns for knocking out fast protons.

$$nvN\sigma = nv\Sigma = \text{number of events per second}^9$$
$$= 10^{10} \cdot 10^{-3} \text{ cm}^2 = 10^7 \text{ protons/sec}$$

[9] It is convenient in such calculations to use the macroscopic target cross section Σ, which is the nuclear cross section times the number of atoms in the target. Writing Avogadro's number ≈ 0.6 and cross section in barns, the factors 10^{24} cancel and one gets a cross-sectional area in square centimeters that can be visualized.

from fast neutron processes, as compared with 10^8 from neutron decays. Thus the bottle experiment is marginal from a background standpoint. Actually strange things happen to objects put in a pile. A bottle sent right through from one side to the other would come out absolutely black. There is very little chance of holding a vacuum of better than 10^{-4} mm for a year inside a pile.

Actual progress on the half-life of the neutron has been made both at Oak Ridge[10] and at Chalk River in Canada[11] in thermal columns.

Incidentally the thermal column is a good example of the failure of a more or less accepted law, and illustrates the importance of being critical of accepted laws. It had been widely believed ten years ago that fluxes of fast neutrons cannot be transformed into a pure thermal flux. The war introduced graphite as a moderator, which, in spite of its poor thermalizing efficiency compared with hydrogen with respect to the number of collisions required, had the peculiar property of having a greater mean free path for slow neutrons than for fast ones. Graphite then "allows the thermals to run away from the fast neutrons."

EXPLORATORY EXPERIMENTS. In this chapter we have divided experiments into two kinds: (1) exploratory, to find out what is there, and (2) question and answer, to measure a particular quantity or set of such quantities. Almost all of the first portion of these notes was concerned with experiments of the second kind in which one starts out knowing just what is to be measured and has available general orders of magnitude of the factors entering the design. It is in just such experiments that what might be called the "complementarity of intensity and resolution" imposes the greatest demands on the ingenuity of the experimenter.

A large area of experimental physics, however, is primarily exploratory. A good example is that part of spectroscopy concerned with search for hyperfine structure and regularities in spectra. Another perhaps more conspicuous example is the domain of high-energy particle physics. The laboratory production of mesons, nuclear reaction kinetics, and spallation processes at energies of hundreds of million electron volts open wholly new areas of nuclear research. The preliminary experiments in these fields are frequently of an exploratory nature. Only after some general knowledge has been gained are specific points tracked down by experiments of the question and answer type.

In designing exploratory experiments, intensity as always must be considered, but resolving power is generally less important. Perhaps the

[10] Snell, Arthur H., Pleasanton, Frances, and McCord, R. V., "Radioactive Decay of the Neutron," *Phys. Rev.*, *78*, 310 (1950).

[11] Robson, J. M., "Radioactive Decay of the Neutron," *Phys. Rev.*, *83*, 349 (1951).

most important feature in an exploratory experiment is consideration of the scope of the experiment. In the sense here used, scope is the opposite of resolving power and defines ranges or magnitudes over which effects may be observed and reliable information may be gathered. In any well-planned exploratory experiment it is of the utmost importance to know and be able to define clearly the scope in advance in order to determine those areas adjoining but lying outside the scope that are not explored. Indeed, reported results of exploratory experiments with insufficient or ambiguous statement of scope have in some cases effected serious though temporary setbacks to the progress of physics. Exploratory experiments are difficult to interpret at best, and may well be more of a hindrance and source of confusion unless the limits of the area explored are sharply defined.

An example of the effect produced by an experiment with ill-defined scope is provided in Millikan's first experiment on the latitude effect in cosmic rays.[12] If the primary cosmic rays are charged particles from outer space it is to be expected that the magnetic field of the earth should deflect those with relatively low energy in such a way that they

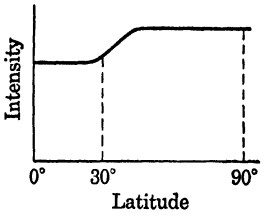


Fig. 12. Intensity of cosmic rays vs. latitude at sea level.

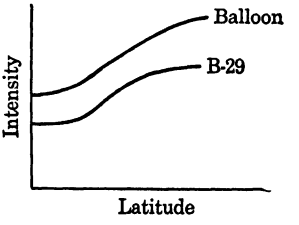


Fig. 13. Latitude effect at higher elevations.

are excluded from the earth's equational zone but may reach the earth in large numbers toward the poles. Millikan placed ionization chambers on a ship at San Pedro, California, to measure the cosmic-ray intensity en route to Panama. Unfortunately, the first good data were not obtained until the ship reached Acapulco, Mexico. No change in intensity was found southward from this point to Panama and on down to Peru. Millikan was then led to conclude that there was no latitude effect and that accordingly the primaries were uncharged. Actually, the intensity curve has a "knee" in it (Fig. 12), and by unhappy chance, Millikan had merely explored the plateau region that happens to begin at about the latitude of Acapulco. Later sea level exploration revealed a second plateau region extending from the latitude of Seattle northward to the

[12] Millikan, Robert A., *Electrons* (+ *and* −). Chicago: University of Chicago Press, 1947, p. 446.

pole. By this time the latitude effect was well established, denying the earlier uncharged primary theory, but the second plateau was interpreted as resulting from the magnetic field of the sun, estimated from the data to be of the order of 30 gausses. But the scope of these explorations was still limited and poorly defined and consequently, at present, the 30 gauss field of the sun is so widely reported in textbooks that it will be some time before this "well-established fact" is set aside. Still more recent explorations at "middle" (B-29) and "high" (balloon) altitudes provide a more complete picture of the effect. In particular, the recent work of Pomerantz[13] (Fig. 13) with balloons indicates that the sun's field is only a fraction of a gauss. Further precision experiments on this matter are needed.

The importance of scope to the interpretation of exploratory experiments is well put in the famous fishnet story due to Jeans. Two fishermen set out with seines to determine the very smallest fish that swam in the ocean. They went out with nets differing somewhat in the size of the mesh. On their return, they each reported different values for the minimum size of fish that swim in the ocean. The ensuing argument was fought out for years in journals.

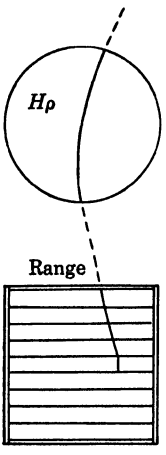


Fig. 14. Arrangement for measuring mass of mesons.

Fretter's experiment to determine the mass of the meson[14] may be considered somewhat exploratory. It was known that the μ-meson had a mass $\approx$ 200 electron masses, the π-meson, however, had not been discovered. In such circumstances it is desirable to measure the mass of any particles differing appreciably from the electron mass. As an exploratory experiment the arrangement used (Fig. 14) has received criticism because its geometry and physical arrangement excluded measurement of mass less than about 50 electron masses.

Green[15] has recently performed an experiment on penetrating showers in carbon plates in a cloud chamber. His arrangement "expands" the shower so that he can look for effects produced by π-mesons (Fig. 15), but suffers the defect that particles scattered into wide angles escape analysis. Thus in order to see clearly one area of a region of exploration, one frequently sacrifices detail in another. This effect might be classified as relative resolving

[13] Pomerantz, M. A. and McClure, G. W., "Primary Cosmic Radiation at High Latitudes," *Phys. Rev., 86,* 536 (1952).

[14] Fretter, William B., "The Mass of Cosmic-Ray Mesotrons," *Phys. Rev., 70,* 625 (1946).

[15] Green, John R., "Penetrating Showers in Carbon," *Phys. Rev., 80,* 832 (1950).

power within the scope. Careful consideration of this factor is important in defining scope.

Recognition of unusual events in exploratory experiments has led to some of the greatest single advances in physics. In some cases it places remarkably severe demands on the alertness and power of recognition of the unusual. Perhaps the classic example of this is the single photograph of a cloud chamber track that won for Anderson the Nobel Prize for his discovery of the positron in 1936. Anderson,[16] during the course of some exploratory work under Millikan was observing tracks in a cloud chamber in which he had placed a lead plate. Out of thousands of pictures he noticed one containing a track that was unusual in two respects. After recognizing that the particle came from below by the sharper curvature in the part of the chamber above the plate, he

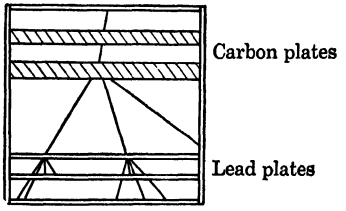


Fig. 15. Penetrating shower in carbon.

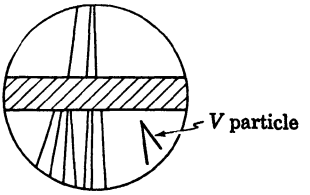


Fig. 16. V particle in penetrating shower.

then realized, as perhaps few observers might, that the direction of curvature was appropriate only for a positive charge. The momentum loss in the plate $\Delta H\rho$ and the density of ionization showed the mass to be that of an electron.

Rochester and Butler[17] in England, in observing penetrating showers produced in a thick lead plate in a cloud chamber (Fig. 16), noticed two tracks of a remarkable form that have since become known as V-particles. From energy and momentum considerations they concluded that the new tracks could not have been produced as collision recoils but must be decay processes of a new kind. Anderson[18] and others have since observed these new tracks, confirming the earlier observation. Further study of these as yet unexplained tracks might well open a whole new area of study.

[16] Anderson, Carl D., "The Positive Electron," *Phys. Rev., 43,* 491 (1933).

[17] Rochester, G. D. and Butler, C. C., "Evidence for the Existence of New Unstable Elementary Particles," *Nature, 160,* 855 (1947).

[18] Seriff, A. J., Leighton, R. B., Hsiao, C., Cowan, E. W., and Anderson, C. D., "Cloud-Chamber Observations of the New Unstable Cosmic-Ray Particles," *Phys. Rev., 78,* 290 (1950).

Some other noteworthy examples of accidental discoveries made in the course of exploratory experiments in cosmic rays were the discovery of penetrating showers by L. Fussell[19] and the discovery of the μ-meson by Anderson and Neddermeyer.[20] The latter was found when a track of a very penetrating particle showed practically no energy loss in crossing a thick lead plate. The π-meson was discovered in the course of examining nuclear plates exposed for a time on high mountains.[21] Heavy nuclei in cosmic rays were similarly found in plates sent up in balloons. The discovery of cosmic rays themselves was accidentally made during the course of investigating the cause of the apparent spontaneous discharge of electroscopes.

Exploratory experiments, it might be added, are not particularly good material for Ph.D. papers, since they frequently yield nothing of particular interest and are at best difficult to interpret. Anything of particular interest that turns up is of course followed up by experiments of the question and answer type. Frequently a question and answer experiment is purposely preceded by one of an exploratory nature for the purpose of developing a technique.

One final remark may be made on a difference in techniques for exploratory and precision experiments. Generally in a precision experiment the physicist takes great care in the mechanical and electrical design of equipment, making it produce as accurate results as he possibly can within practical limits of availability of money, time, and materials. For exploratory experiments, on the other hand, physicists often design equipment that can be made cheaply and quickly, known commonly as haywire apparatus. The uncertainty of results in such experiments and the desire to get on to something else if a particular apparatus yields little useful data is back of this tendency to use haywire equipment. The only word of caution here is not to make the apparatus so poorly that it is always breaking down or would not reveal data even if it exists. The design of exploratory experiments involves a wide knowledge of what is known in a given field, much physical intuition, and as often as not is more an art than a science.

[19] Fussell, L., "Production and Absorption of Cosmic-Ray Showers," *Phys. Rev.*, *51*, 1005 (1936).

[20] Neddermeyer, Seth H. and Anderson, Carl D., "Note on the Nature of Cosmic-Ray Particles," *Phys. Rev.*, *51*, 884 (1937).

[21] Lattes, C. M. G., Occhialini, G. P. S., and Powell, C. F., "Observations on the Tracks of Slow Mesons in Photographic Emulsions," *Nature*, *160*, 453 (1947).

INDEX